

# UC Berkeley

## UC Berkeley Electronic Theses and Dissertations

### Title

The Quantum Many-Body Problem: Methods and Analysis

### Permalink

<https://escholarship.org/uc/item/8cc708cd>

### Author

Lindsey, Michael

### Publication Date

2019

Peer reviewed|Thesis/dissertation

The Quantum Many-Body Problem: Methods and Analysis

by

Michael Lindsey

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Applied Mathematics

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Lin Lin, Chair

Professor Jeffrey Neaton

Professor James Sethian

Summer 2019

The Quantum Many-Body Problem: Methods and Analysis

Copyright 2019  
by  
Michael Lindsey

## Abstract

### The Quantum Many-Body Problem: Methods and Analysis

by

Michael Lindsey

Doctor of Philosophy in Applied Mathematics

University of California, Berkeley

Professor Lin Lin, Chair

This dissertation concerns the quantum many-body problem, which is the problem of predicting the properties of systems of several quantum particles from the first principles of quantum mechanics. Included under this umbrella are various problems of fundamental importance in quantum chemistry, condensed matter physics, and materials science. Of particular interest is the electronic structure problem, the problem of determining the state of the electrons in a system with fixed atomic nuclei. Since direct numerical solution of the many-body Schrödinger equation is intractable even for systems of moderate size, a diverse array of approximate methods has been developed. The broad goals of this dissertation are to improve the mathematical understanding of certain widely-used approximations, as well as to propose new methods. Roughly speaking, we consider three (overlapping) categories of methods: Green's function methods, embedding methods, and variational methods.

One can understand Green's function methods in terms of many-body perturbation theory, which computes series expansions of physical quantities about a non-interacting reference system. These expansions can be expressed graphically in terms of Feynman diagrams, which can in turn be reorganized, in some cases, into an expansion in terms of so-called bold diagrams. Green's function methods can be specified by choosing a subset of bold diagrams to approximate the sum. At the same time, such methods can be understood in terms of an object known as the Luttinger-Ward (LW) functional, which admits a representation in terms of the bold diagrams. Many aspects of these constructions are purely formal, and indeed the existence of the fermionic LW functional as a single-valued functional has recently been called into question. To contribute to the understanding of these issues, we provide rigorous proofs of the combinatorial construction and analytic interpretation of the bold diagrams in the simplified setting of a classical field theory. In this setting we also provide a rigorous non-perturbative construction of the LW functional via convex duality and prove several key properties, including continuity up to the boundary of its domain and asymptotics in the limit of large interaction.

Quantum embedding methods, meanwhile, view a large system as being composed of smaller fragments that are treated with high accuracy and embedded in the larger system in a mutually consistent way. Inspired by a connection between the boundary analysis of the LW functional and embedding, we perform similar analysis for the 1-RDM theory for fermionic systems, which is also developed via convex duality, illustrating a relation to fermionic embedding methods such as the density matrix embedding theory (DMET).

Another embedding method of note is the dynamical mean-field theory (DMFT), which is at the same time a Green's function method that can be understood in terms of the LW functional. DMFT relies on the solution of impurity problems, which specify the embedding of an interacting system into a non-interacting bath. Underlying DMFT is a result about the sparsity pattern of the self-energy matrix for impurity problems, which to our knowledge has not been proved in the literature. We provide a rigorous proof of this result in various classical and quantum settings. We go on to investigate the fermionic DMFT in depth, identifying the key mathematical structures that appear in the algorithmic loop for solving it and using these to prove the well-posedness of this loop, in a certain sense.

Finally, we introduce a suite of new approaches to the quantum many-body problem that provide variational lower bounds to the ground-state energy. These methods, which combine the themes of convexity and embedding, are based on novel convex relaxations of the variational principles for the ground-state energies of many-body systems. To begin, we recover a second-quantized version of the formalism of strictly correlated electrons (SCE), which yields an exact expression for the exchange-correlation functional in Kohn-Sham density functional theory in the limit of infinite Coulomb repulsion in terms of the solution of a multi-marginal optimal transport problem. We introduce a semidefinite relaxation method for approximately solving this problem and obtaining a lower bound for the ground-state energy. The ideas underlying this relaxation are generalized considerably, outside the context of SCE, to yield much tighter lower bounds, which we validate numerically for both quantum spin systems and fermionic systems. We also describe how these relaxation methods can be interpreted as embedding methods via convex duality.

## Acknowledgments

First I am pleased to acknowledge my advisor Lin Lin. He is truly a supernova of cheerful and creative energy. I continue to learn from his scientific vision and to take inspiration from his sense of mission. Meanwhile his support and encouragement have been substantial. Of course, none of the work detailed below could have happened without Lin.

For their generosity and guidance during my first year of graduate school, I am grateful to Jamie Sethian, Jeff Donatelli, Alexandre Chorin, Fei Lu, and Fraydoun Rezakhanlou. I learned something valuable from each of them during this important time in my career. In particular, my perspectives on statistical mechanics, which form a central part of this work, owe much to Fraydoun.

Among my collaborators I extend special thanks to Reinhold Schneider for his hospitality, Yanir Rubinstein for his guidance over the years, and Yu Tong for many engaging discussions and indeed for his partnership in approaching the quantum many-body problem. I am likewise thankful to Dong An and Yuehaw Khoo for introducing me to interesting problems that have led to some enjoyable and rewarding collaborations. For collaborations leading to work featured in this dissertation, I thank Yuehaw, Lin, and Reinhold once again, as well as Lexing Ying. I also gratefully acknowledge the financial support of the National Science Foundation Graduate Research Fellowship Program under grant DGE-1106400.

Finally, it is a special joy to thank Katie for being my teammate every single day.

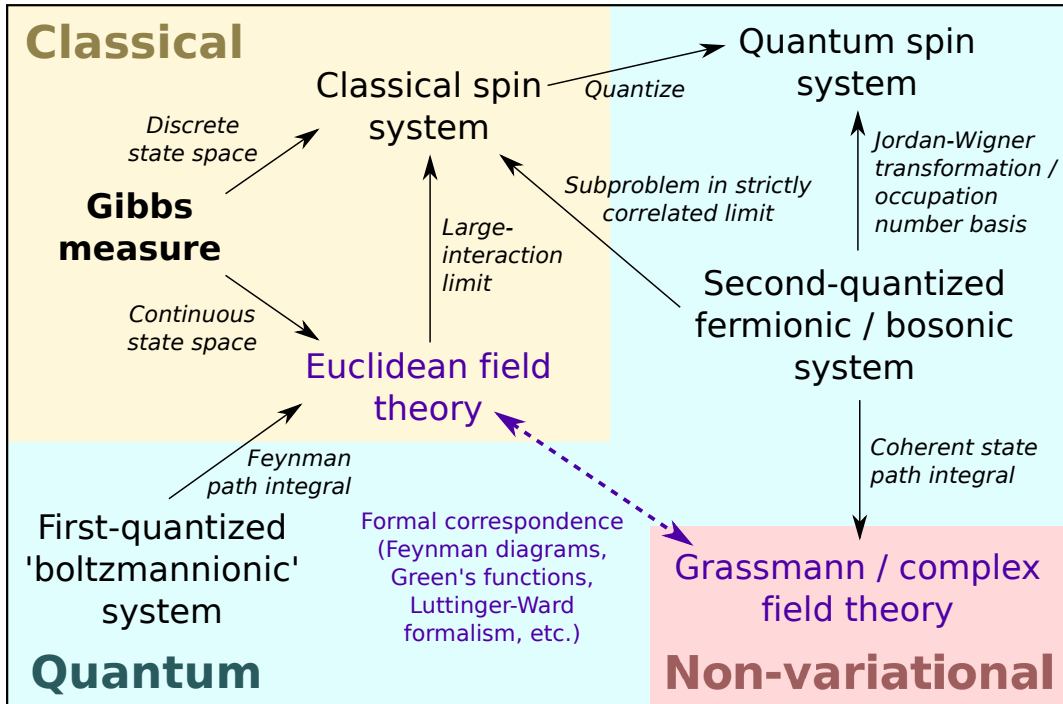


Figure 1: A map of ideas, to be interpreted somewhat loosely. The yellow and blue regions correspond to the domains of classical and quantum mechanics, respectively, and arrows indicate correspondences between settings. The purple text and arrows indicate the formal correspondence between Euclidean field theory and Grassmann / complex field theories. The red box for the Grassmann / complex field theories indicates that they do not enjoy the same variational structure as their Euclidean counterpart, i.e., there is no suitable Gibbs variational principle in this setting. By contrast, the other perspectives all enjoy suitable (classical or quantum) Gibbs variational principles.

## Preface

This work concerns statistical mechanics, broadly construed. Of course, the brevity of the term ‘statistical mechanics’ conceals the startling diversity of ideas falling under this heading, the vast majority of which shall make no appearance here. Our focus is necessarily limited to certain facets of statistical mechanics, classical and quantum, as well as some of the connections between them. Nonetheless the coherence of the ideas presented in this work is best appreciated in a broad unifying context, summarized in Figure 1. Part I below will describe the needed background for interpreting this context. Perhaps unsurprisingly, variational principles and the Legendre transform will play key unifying roles.

In fact, our guiding interest is to understand and evaluate computational approaches to the quantum many-body problem. Of special particular interest is a

fermionic many-body problem, the electronic structure problem, i.e., the problem of determining the quantum state of electrons in a system with fixed atomic nuclei, which plays a fundamental role in quantum chemistry, condensed matter, and materials science. Since direct numerical solution of the many-body Schrödinger equation is intractable even for systems of moderate size, there is a diverse hierarchy of approximate methods, trading off to varying degrees between accuracy and computational efficiency. A major goal of this work is to improve the mathematical understanding of widely-used approximations from physics and chemistry, as well as to suggest and validate new approaches.

Now even a single-minded concern for fermions is enough to motivate significant detours through the classical realm. A variety of techniques known collectively as path integrals generate correspondences from quantum-mechanical ensembles (even at zero temperature) to the Gibbs measures of classical statistical mechanics. For fermions, such a correspondence is, unfortunately, only formal. However, it gives us a perspective on some of the most widely-used methods in electronic structure, namely the *Green's function methods*, which arise from a machinery called the *Luttinger-Ward (LW) formalism*<sup>1</sup> [65] (featuring a functional of the same name), or alternatively from manipulations of the perturbatively defined *bold Feynman diagrams* [100]. The LW formalism has been widely used in both physics and chemistry [28, 45, 10, 90]. One perspective on this formalism is that of many-body perturbation theory (MBPT), which represents physical quantities of interest via perturbative expansion in the strength  $\lambda$  of the inter-particle interaction (i.e., in the case of electronic structure, the repulsion between electrons). The terms in this expansion correspond to the so-called bare Feynman diagrams. These can in turn be reorganized into a series of bold Feynman diagrams via a procedure known as ‘renormalizing the propagator,’ which involves infinite summations that are purely formal without further justification. The selection of some subset of diagrams from the bold diagrammatic series for a physical quantity known as the self-energy specifies a Green’s function method, which can be used to compute Green’s functions. Such Green’s function methods can also be formally justified via a construction known as the Luttinger-Ward (LW) functional.

In Part II, we provide the first rigorous combinatorial construction of the bold diagrams with methods applicable to rather general field theories. In Part III, by transposing to a classical statistical-mechanical setting (which we refer to as the Euclidean lattice field theory), we are able to non-perturbatively construct the LW functional in terms of the Legendre transform of a convex function, and moreover we are able provide the first rigorous interpretation for the bold diagrams as an asymptotic series for the LW functional. The construction exploits variational structure and convexity properties that are not available in the fermionic setting, though the formal perspective of the Legendre transform is still active there. In fact, the mathematical status of the LW formalism in the fermionic setting is still an open matter

---

<sup>1</sup>The Luttinger-Ward formalism is also known as the Kadanoff-Baym formalism [7] depending on the context. In this work we always use the former.



of considerable contemporary interest; see [54, 32, 103, 42].

In Part IV, we explore in the Euclidean setting the LW functional in the limit of *large*  $\lambda$ . We prove the dominance of the Hartree contribution in this regime and derive a leading-order correction defined in terms of classical spin systems. The extension of this analysis to the quantum many-body setting is still an open topic of interest.

Our construction of the Luttinger-Ward functional reveals several interesting structural properties. Of particular note is the continuity of the LW functional up to the boundary of its domain, which does not trivially follow from its definition. In fact we derive a formula that expresses the LW functional on the boundary of its domain in terms of the LW functional corresponding to a smaller physical system. This observation relates the LW functional to impurity problems, which will be featured directly in Part VI, as well as in the context of the dynamical mean-field theory (DMFT)[37, 53] in Part VII.

The appearance of the impurity problem motivates us to say something now about *embedding*, which plays a central role throughout this work. Due to the difficulty of the quantum many-body problem, it is desirable to reduce computational effort by dedicating more resources to small ‘difficult’ regions of a system where many-body effects are significant and fewer resources to ‘easier’ regions where, e.g., a single-particle picture is appropriate. While this goal seems reasonable from a high level, the right mathematical framework for accomplishing such an ‘embedding’ of a small problem into a larger one is a priori unclear. Remarkably, quantum embedding is useful not only in the scenario of a single small region of special interest, but also for extended systems in which no such region is privileged. Indeed, one can view an extended system as being divided into many smaller disjoint fragments. The goal is then to view each of these fragments as being embedded in a larger system in a mutually consistent way.

Now DMFT is in particular an embedding method (as well as a Green’s function method). It relies on the solution of impurity problems, which specify the embedding of an interacting system into a non-interacting bath. We shall keep the connection between boundary analysis and embedding in mind.

In Part V, we provide a rigorous development of the 1-RDM theory, which is in a certain sense an analog of the Euclidean LW formalism in the fermionic setting. But rather than highlighting the aforementioned field-theoretic and diagrammatic analogies between the Euclidean and fermionic LW formalisms, the 1-RDM theory represents the analogy of the *convex analysis* that underlies the Euclidean LW theory. Motivated by our boundary analysis for the LW functional, we carry out a similar analysis for the 1-RDM theory. This analysis naturally leads us to prove what we term the ‘embedding lemma,’ which underlies the justification of complete active space methods, as well as the density matrix embedding theory (DMET) [49], in the quantum chemistry literature.

With that we leave the realm of convex analysis to continue our investigation of impurity problems. In Part VI, we prove that for various impurity models, in both

classical and quantum settings, the self-energy matrix is a sparse matrix with a sparsity pattern determined by the impurity sites. In the quantum setting, such a sparsity pattern has been known since Feynman [35]. Indeed, it underlies several numerical methods for solving impurity problems, as well as many approaches to more general quantum many-body problems, such as the dynamical mean field theory. The sparsity pattern is easily motivated by a formal perturbative expansion using Feynman diagrams. However, to the extent of our knowledge, a rigorous proof has not appeared in the literature. In the classical setting, analogous considerations lead to a perhaps less-known result, i.e., that the precision matrix of a Gibbs measure of a certain kind differs only by a sparse matrix from the precision matrix of a corresponding Gaussian measure. Our argument for this result mainly involves elementary algebraic manipulations and is in particular non-perturbative. Nonetheless, the proof is robustly adapted to various settings of interest in physics, including quantum systems (both fermionic and bosonic) at zero and finite temperature, non-equilibrium systems, and superconducting systems.

The sparsity result for impurity problems is at the heart of DMFT, which we investigate in depth in Part VII. After illustrating the connection to the LW formalism via the setting of classical field theory, we turn to identifying the key mathematical structures in the algorithmic loop of DMFT for fermionic systems. In particular, we provide rigorous proof of the well-posedness of the DMFT loop in a certain sense, which should be compared with [52]. Moreover, the mathematical framework presented in this part offers a perspective on hybridization fitting, a preprocessing step for impurity problems in which the influence of the bath on the fragment is approximated by a smaller, more computationally tractable bath. This perspective has been reflected in the recent work [73].

In the final two parts of this dissertation, we discuss a suite of approaches to the quantum many-body problem that diverge from the previous parts in many formal respects yet are united to them by the recurring themes of convexity and embedding. The point of departure is again the convex variational principle for quantum many-body problems, and here we introduce computationally tractable *convex relaxations* of these variational principles which provide lower bounds to the ground state energy.

In Part VIII, the electronic many-body problem is the focus, and the relaxation under consideration really consists of two successive relaxations. The first of these relaxations is motivated via its asymptotic tightness in the limit of infinitely strong Coulomb repulsion, called the limit of *strictly correlated electrons (SCE)*. The idea of SCE [95, 94, 20, 66, 26] originally arose in first quantization as a limit in which the exchange-correlation functional in Kohn-Sham density functional theory [44, 50] admits exact expression in terms of the solution of an (albeit costly) multimarginal optimal transport (MMOT) [80] problem with pairwise cost. Our relaxation can be considered as a second-quantized analog of this development and indeed also may be exactly solved via an MMOT problem with pairwise cost, though this problem is importantly different from its first-quantized analog. Thinking of second-quantized

SCE as a *model*, we then propose a means to approximately solve it via a novel semidefinite relaxation of the MMOT problem.

In Part IX, we broaden and strengthen the approach of Part VIII, bypassing the SCE step and directly formulating a framework of new convex relaxations of the variational principle for fermionic systems, as well as quantum spin systems. We also describe how these relaxations can be interpreted as embedding methods, suggesting a pathway for fast algorithms, and unifying, for perhaps the first time, variational and embedding-based approaches to the quantum many-body problem.

Please note that Part II is based on [61] (joint work with Lin Lin), Part III is based on [62] (joint work with Lin Lin), Part VI is based on [63] (joint work with Lin Lin), and Part VIII is based on [47] (joint work with Yuehaw Khoo, Lin Lin, and Lexing Ying). Also, much of section 7 of Part I is based on the appendices of [63]. Finally, Part VII is based on joint work in preparation with Lin Lin and Reinhold Schneider, and Part IX is based on joint work in preparation with Lin Lin.

# Table of Contents

<b>Acknowledgments</b>	<b>i</b>
<b>Preface</b>	<b>ii</b>
<b>I Background</b>	<b>1</b>
<b>1 Classical statistical mechanics</b>	<b>1</b>
1.1 Gibbs measures . . . . .	1
1.2 Gibbs variational principle . . . . .	2
<b>2 Quantum statistical mechanics</b>	<b>3</b>
2.1 Quantum state spaces and Hamiltonians . . . . .	3
2.2 Quantum Gibbs states . . . . .	4
<b>3 First quantization and electronic structure</b>	<b>5</b>
3.1 Identical particles . . . . .	7
3.2 Bases for boltzmannionic, fermionic, and bosonic state spaces . . . . .	8
3.3 Feynman path integral for boltzmannions . . . . .	8
<b>4 Second quantization</b>	<b>11</b>
4.1 The Fock space . . . . .	11
4.2 The creation and annihilation operators . . . . .	12
4.3 Second-quantized operators . . . . .	13
4.4 The Jordan-Wigner transformation . . . . .	17
4.5 Canonical transformations and noninteracting problems . . . . .	18
4.6 Second-quantized model Hamiltonians . . . . .	19
<b>5 Fermionic and bosonic statistical mechanics</b>	<b>19</b>
5.1 The zero-temperature ensemble . . . . .	19
5.2 The finite-temperature ensemble . . . . .	20
<b>6 The coherent state path integral</b>	<b>22</b>
6.1 The bosonic coherent state path integral . . . . .	22
6.1.1 Bosonic coherent states . . . . .	23
6.1.2 Complex coordinates . . . . .	24
6.1.3 The resolution of identity . . . . .	25
6.1.4 Path integral . . . . .	26
6.1.5 Path integral in frequency space . . . . .	28
6.2 The fermionic coherent state path integral . . . . .	29
6.2.1 Grassmann numbers . . . . .	30
6.2.2 Fermionic coherent states . . . . .	30

6.2.3	Grassmann integration . . . . .	31
6.2.4	The resolution of identity . . . . .	33
6.2.5	Path integral . . . . .	35
<b>7</b>	<b>Green's functions</b>	<b>37</b>
7.1	Motivation via functional derivatives . . . . .	37
7.2	Green's functions and the self-energy at zero temperature . . . . .	39
7.3	Green's functions and the self-energy at finite temperature . . . . .	40
7.4	Matsubara Green's functions and self-energy . . . . .	42
<b>II</b>	<b>Bold Feynman diagrams</b>	<b>45</b>
<b>1</b>	<b>Introduction</b>	<b>45</b>
1.1	Contributions . . . . .	46
1.2	Related work . . . . .	47
1.3	Perspectives . . . . .	48
1.4	Outline . . . . .	49
<b>2</b>	<b>Preliminaries</b>	<b>49</b>
2.1	Notation . . . . .	50
2.2	Scaling relation . . . . .	52
2.3	Wick theorem . . . . .	53
<b>3</b>	<b>Feynman diagrams</b>	<b>53</b>
3.1	Motivation . . . . .	55
3.2	Labeled and unlabeled diagrams . . . . .	57
3.3	Feynman rules for $Z$ . . . . .	62
3.4	Comments on other interactions . . . . .	64
3.5	Feynman rules for $\Omega$ . . . . .	65
3.6	Feynman rules for $G$ . . . . .	67
3.7	Why we do not use fully unlabeled Green's function diagrams . . . . .	72
3.8	Feynman rules for $\Sigma$ . . . . .	72
<b>4</b>	<b>Bold diagrams</b>	<b>77</b>
4.1	Skeleton decomposition . . . . .	80
4.2	Motivating examples . . . . .	82
4.3	Ways of producing a self-energy diagram from its skeleton . . . . .	85
4.4	Understanding automorphisms in terms of the skeleton . . . . .	87
4.5	The action of $\text{Aut}(\Gamma_s)$ on $I(\Gamma, \Gamma_s)$ . . . . .	90
4.6	Bold diagrammatic expansion for the self-energy . . . . .	91
4.7	Green's function methods . . . . .	93
4.7.1	The GW approximation . . . . .	94

4.8	A preview of the Luttinger-Ward formalism . . . . .	95
4.8.1	$\Phi$ -derivability . . . . .	97
4.9	Quantities that <i>do not</i> admit a bold diagrammatic expansion . . . . .	99
<b>5</b>	<b>Numerical experiments</b>	<b>100</b>
<b>III</b>	<b>The Luttinger-Ward formalism</b>	<b>104</b>
<b>1</b>	<b>Introduction</b>	<b>104</b>
1.1	Contributions . . . . .	105
1.2	Outline . . . . .	106
<b>2</b>	<b>Preliminaries</b>	<b>106</b>
2.1	Notation and quantities of interest . . . . .	106
2.2	Interaction growth conditions . . . . .	109
2.3	Measures and entropy: notation and facts . . . . .	110
<b>3</b>	<b>Luttinger-Ward formalism</b>	<b>113</b>
3.1	Variational formulation of the free energy . . . . .	113
3.2	The Luttinger-Ward functional and the Dyson equation . . . . .	115
3.3	Transformation rule for the LW functional . . . . .	117
3.4	Impurity problems and the projection rule . . . . .	118
3.5	Continuous extension of the LW functional to the boundary . . . . .	121
<b>4</b>	<b>Bold diagrammatic expansion for the generalized Coulomb interaction</b>	<b>123</b>
4.1	Existence of asymptotic series . . . . .	124
4.2	Relating the LW and self-energy expansions . . . . .	125
4.3	Diagram-free discussion of results from Part II . . . . .	126
4.4	Derivation of self-energy bold diagrams . . . . .	127
4.5	Caveat concerning truncation of the bold diagrammatic expansion . . . . .	131
<b>5</b>	<b>Proof of the continuous extension of the LW functional</b>	<b>133</b>
5.1	Proof setup . . . . .	133
5.2	Lower bound . . . . .	134
5.3	Upper bound . . . . .	137
5.4	Dual perspective on continuous extension . . . . .	139
5.5	Counterexample of weak but not strong growth . . . . .	140
<b>IV</b>	<b>Large-interaction Luttinger-Ward asymptotics</b>	<b>142</b>
<b>1</b>	<b>Introduction</b>	<b>142</b>

<b>2</b>	<b>Preliminaries</b>	<b>142</b>
<b>3</b>	<b>Heuristic asymptotics</b>	<b>144</b>
<b>4</b>	<b>Spin system duality</b>	<b>147</b>
<b>5</b>	<b>Rigorous asymptotics</b>	<b>149</b>
<b>6</b>	<b>Gradient asymptotics</b>	<b>161</b>
<b>7</b>	<b>Polarization relation</b>	<b>163</b>
<b>8</b>	<b>Asymptotic polarization and the limit of <math>M_\lambda[G]</math></b>	<b>165</b>
<b>V</b>	<b>The 1-RDM theory and fermionic embedding</b>	<b>169</b>
<b>1</b>	<b>Introduction</b>	<b>169</b>
<b>2</b>	<b>The fermionic 1-RDM theory</b>	<b>169</b>
2.1	The noninteracting case . . . . .	175
2.2	Comparison with the Luttinger-Ward formalism . . . . .	176
<b>3</b>	<b>The embedding lemma and boundary analysis</b>	<b>177</b>
<b>4</b>	<b>The zero-temperature limit</b>	<b>180</b>
<b>5</b>	<b>Two-body embedding Hamiltonian</b>	<b>181</b>
<b>VI</b>	<b>Classical and quantum impurity problems</b>	<b>184</b>
<b>1</b>	<b>Introduction</b>	<b>184</b>
1.1	Other formal perspectives . . . . .	187
1.2	Outline . . . . .	189
<b>2</b>	<b>The classical impurity problem</b>	<b>190</b>
<b>3</b>	<b>The quantum impurity problem</b>	<b>192</b>
3.1	Zero temperature . . . . .	194
3.2	Finite temperature . . . . .	196
3.3	Arbitrary contour . . . . .	197
3.4	Anomalous setting . . . . .	202

<b>VII</b>	<b>The dynamical mean-field theory</b>	<b>209</b>
<b>1</b>	<b>Introduction</b>	<b>209</b>
1.1	Outline . . . . .	209
<b>2</b>	<b>Warm-up: DMFT in Euclidean field theory</b>	<b>209</b>
2.1	Formulation of cluster DMFT . . . . .	209
2.2	Algorithmic approach . . . . .	211
2.3	Luttinger-Ward Perspective . . . . .	211
2.4	Diagrammatic corrections to DMFT . . . . .	212
2.5	Overlapping DMFT . . . . .	212
2.6	Diagrammatic corrections to overlapping DMFT . . . . .	214
<b>3</b>	<b>Fermionic DMFT</b>	<b>214</b>
3.1	Basic definitions . . . . .	214
3.2	The impurity problem and the hybridization . . . . .	217
3.2.1	Bath fitting . . . . .	219
3.2.2	Chain Hamiltonian perspective . . . . .	219
3.3	The (cluster) dynamical mean field theory . . . . .	220
3.3.1	The self-consistency condition . . . . .	221
3.3.2	The DMFT loop . . . . .	222
3.3.3	The DMFT loop with bath fitting . . . . .	223
3.3.4	The DMFT loop with chain truncation . . . . .	223
<b>4</b>	<b>Mathematical structures</b>	<b>224</b>
4.1	Function spaces of interest . . . . .	224
4.2	Structure of Green's functions . . . . .	226
4.3	Pointwise inversion relations . . . . .	229
4.4	Structure of the self-energy and the Hartree-Fock contribution . . . . .	233
<b>5</b>	<b>Application to DMFT</b>	<b>237</b>
5.1	Well-posedness of DMFT loop . . . . .	237
5.2	Sparsity pattern of the hybridization for local models . . . . .	238
5.3	Definition and well-posedness of HF+DMFT loop . . . . .	238
<b>VIII</b>	<b>Strictly correlated electrons in second quantization</b>	<b>240</b>
<b>1</b>	<b>Introduction</b>	<b>240</b>
1.1	Contributions . . . . .	241
1.2	Related work . . . . .	242
1.3	Outline . . . . .	243



<b>2</b>	<b>Preliminaries</b>	<b>243</b>
2.1	Density functional theory in second quantization . . . . .	243
2.2	Strictly correlated electron limit . . . . .	246
2.2.1	The Kohn-Sham SCE equations . . . . .	247
2.2.2	The SCE energy and potential . . . . .	249
<b>3</b>	<b>Convex relaxation</b>	<b>251</b>
3.1	Primal problem . . . . .	251
3.2	Dual problem . . . . .	253
3.3	Strong duality and dual attainment . . . . .	255
<b>4</b>	<b>Tighter lower bound via 3-marginals</b>	<b>261</b>
<b>5</b>	<b>General MMOT with pairwise cost</b>	<b>262</b>
<b>6</b>	<b>Numerical results</b>	<b>263</b>
6.1	One-dimensional spinless model . . . . .	264
6.2	Two-dimensional spinful model . . . . .	265
<b>7</b>	<b>Conclusion</b>	<b>267</b>

## IX Variational embedding for quantum spins and fermions 269

<b>1</b>	<b>Introduction</b>	<b>269</b>
1.1	Outline . . . . .	269
<b>2</b>	<b>Quantum spins</b>	<b>270</b>
2.1	Preliminaries . . . . .	270
2.2	Local consistency constraints . . . . .	271
2.3	Global semidefinite constraints and the two-marginal SDP . . . . .	272
2.4	Abstract perspective on the global semidefinite constraints . . . . .	274
2.5	Higher constraints . . . . .	275
2.6	Cluster perspective . . . . .	275
2.7	Overlapping clusters . . . . .	276
2.8	Translation-invariant setting . . . . .	277
2.8.1	Periodicity constraints . . . . .	278
<b>3</b>	<b>Fermions</b>	<b>278</b>
3.1	Preliminaries . . . . .	278
3.2	Abstract perspective . . . . .	280
3.3	The two-cluster-marginal SDP . . . . .	282
3.4	Exactness for noninteracting problems . . . . .	284
3.5	Concrete perspective . . . . .	284

<b>4</b>	<b>Numerical results</b>	<b>287</b>
4.1	Transverse-field Ising model . . . . .	287
4.2	Anti-ferromagnetic Heisenberg model . . . . .	288
4.3	Hubbard model . . . . .	288
<b>5</b>	<b>Duality and the effective Hamiltonian perspective</b>	<b>290</b>
5.1	The quantum Kantorovich problem . . . . .	290
5.2	Partial duality . . . . .	291
5.3	Computational significance of partial duality . . . . .	293
5.4	Full duality . . . . .	294
	<b>Bibliography</b>	<b>296</b>
	<b>Appendix A Proof of Isserlis-Wick theorem</b>	<b>304</b>
	<b>Appendix B Proof of skeleton decomposition</b>	<b>306</b>
	<b>Appendix C Definitions and results from convex analysis</b>	<b>309</b>
	C.1 Convex sets and functions . . . . .	309
	C.2 First-order properties of convex functions . . . . .	310
	C.3 The convex conjugate . . . . .	311
	C.4 Sequences of convex functions . . . . .	312
	<b>Appendix D Classical results on weak convergence of probability mea- sures</b>	<b>314</b>
	<b>Appendix E Proofs of lemmas from Part III</b>	<b>315</b>
	E.1 Lemma 8 . . . . .	315
	E.2 Lemma 9 . . . . .	315
	E.3 Fact 11 . . . . .	316
	E.4 Lemma 2 . . . . .	316
	E.5 Lemma 16 . . . . .	317
	E.6 Lemma 19 . . . . .	318
	E.7 Lemma 20 . . . . .	318
	E.8 Lemma 21 . . . . .	319
	E.9 Lemma 36 . . . . .	319
	E.10 Lemma 37 . . . . .	320
	E.11 Lemma 44 . . . . .	322
	E.12 Lemma 50 . . . . .	322
	<b>Appendix F The noninteracting partition function</b>	<b>325</b>
	<b>Appendix G The Kadanoff-Baym contour</b>	<b>327</b>

# Part I

## Background

This part is dedicated to the presentation of the background material that we will use to introduce and connect the developments outlined in the Preface. The reader should beware that the presentation is not necessary standard in all points. Instead it aims to provide a unified perspective on the themes that tie this work together. More broadly I hope that it will serve as an invitation to the subjects that are active in this work and the connections that animate them. Useful resources for further study include [77, 1, 36].

### 1 Classical statistical mechanics

#### 1.1 Gibbs measures

We consider statistical-mechanical models with discrete sets of sites, indexed by  $i = 1, \dots, N$ . Each site has a local state space  $X_i$ , and for simplicity we assume that  $X_i = X$  for all  $i$ . Then the global state space is defined by  $\mathcal{X} := \bigoplus_i X_i$ . For now let us further assume that the local state space  $X$  is finite. Then given a Hamiltonian  $H : \mathcal{X} \rightarrow \mathbb{R}$  and an inverse temperature  $\beta \in (0, \infty)$ , the primary object of interest is the Gibbs measure defined by the probability mass function

$$\mu_\beta(x) = \frac{1}{Z[\beta]} e^{-\beta H(x)},$$

where  $Z[\beta] = \sum_{x \in \mathcal{X}} e^{-\beta H(x)}$  is the partition function, a normalization constant chosen to ensure that the right-hand side indeed defines a probability measure. Notice that in the zero-temperature limit (i.e., as  $\beta \rightarrow \infty$ ), the Gibbs measure concentrates around the minimizer(s) of  $H$ , hence statistical mechanics at zero temperature recover the general problem of optimization.

More generally, we can lump  $\beta$  into our definition of  $H$  and think of  $H = H_A$  itself as being parametrized by some data  $A$ . Then our Gibbs measure is likewise parametrized by  $A$  via

$$\mu_A(x) = \frac{1}{Z[A]} e^{-H_A(x)},$$

where  $Z[A] = \sum_{x \in \mathcal{X}} e^{-H_A(x)}$ .

A standard example is the ferromagnetic Ising model with external magnetic field, specified by a choice of graph structure for the sites (e.g., a  $d$ -dimensional lattice), local state space  $X = \{-1, 1\}$ , parameters  $A = (\beta, \mu)$ , and Hamiltonian

$$H_{\beta, \mu} = - \sum_{i \sim j} \sigma_i \sigma_j - \mu \sum_i \sigma_i,$$

where the summation over  $i \sim j$  indicates summation over all pairs of indices that are adjacent in the graph. There are many other related models with local state space  $X = \{-1, 1\}$ , known as classical spin systems.

These considerations carry over naturally to the continuous setting, e.g.,  $X_i = \mathbb{R}$ , which may be called ‘Euclidean (alternatively, classical or statistical) lattice field theory.’ Here the Gibbs measure is defined

$$d\mu_A(x) = \frac{1}{Z[A]} e^{-H_A(x)} dx,$$

where  $Z[A] = \int_{\mathbb{R}^N} e^{-H_A(x)} dx$ .

The major example of interest in this work is specified by taking the parameter  $A$  to be a real-symmetric  $N \times N$  matrix and

$$H_A(x) = \frac{1}{2} x^T A x + U(x), \tag{1.1}$$

where  $U(x)$  is thought of as a fixed ‘interaction,’ representing a deviation from Gaussianity, which on its own is trivial to understand. Of particular interest is the interaction form

$$U(x) = \sum_{ij} v_{ij} x_i^2 x_j^2,$$

which we call the *generalized Coulomb interaction* via its formal analogy to the Coulomb interaction of electronic structure, which is reflected in an analogy at the level of Feynman diagrams. Note that this class of models includes as a special case the lattice  $\phi^4$  model, specified by a diagonal kernel  $v_{ij} = \lambda \delta_{ij}$ .

## 1.2 Gibbs variational principle

The partition function, or equivalently the free energy  $\Omega[A] := -\log Z[A]$ , naturally encodes a great deal of information, as we shall see in our discussion of the Luttinger-Ward formalism. In the continuous setting, as we shall verify in Part III the free energy satisfies the Gibbs variational principle

$$\Omega[A] = \inf_{\mu} \left[ \int H_A(x) d\mu(x) - S(\mu) \right],$$

where the infimum is taken over a suitable class of probability measures on  $\mathbb{R}^N$  and is in fact attained by the Gibbs measure  $\mu = \mu_A$ . Here  $S$  is the differential entropy, defined

$$S(\mu) = - \int \log \frac{d\mu}{d\lambda} d\mu$$

for all  $\mu$  absolutely continuous with respect to the Lebesgue measure  $\lambda$  (and defined  $S(\mu) = -\infty$  for  $\mu$  otherwise). Note that  $\frac{d\mu}{d\lambda}$  is the Radon-Nikodym derivative, i.e., the probability density function for  $\mu$ .

A suitable analogous variational principle is available in the discrete setting:

$$\Omega[A] = \inf_{\mu} \left[ \sum_{x \in \mathcal{X}} H_A(x) \mu(x) - S(\mu) \right],$$

where  $S$  is the Shannon entropy, defined

$$S(\mu) = - \sum_{x \in \mathcal{X}} \mu(x) \log \mu(x),$$

with the convention  $0 \log 0 = 0$ .

## 2 Quantum statistical mechanics

What does it mean to ‘quantize’ a classical system? There are two aspects of the procedure: (1) upgrading the classical state space to a corresponding quantum state space and (2) choosing a Hamiltonian. The first point is straightforward, though the second point is more subtle. In elementary quantum mechanics, the procedure known as canonical quantization [36] produces a quantum Hamiltonian from the symplectic structure of a classical Hamiltonian dynamical system. However, in general classical statistical mechanics, the Hamiltonian is merely a function on states and there are no accompanying dynamics. In the setting of quantum spin systems, for example, there are many Hamiltonians of interest that are thought of as phenomenological models for interesting physics and are not derived by ‘quantizing’ classical Hamiltonians.

### 2.1 Quantum state spaces and Hamiltonians

To illustrate the first point, we describe the relation between classical and quantum spin systems. For each site  $i$ , the local quantum state space is given by  $Q_i = \mathbb{C}^{X_i}$ . Note that with the standard inner product, as a complex Hilbert space  $Q_i \simeq L^2(X_i)$ . For example, if  $X_i = \{-1, 1\}$ , then  $Q_i \simeq \mathbb{C}^2$ ; this is the important case of spin- $\frac{1}{2}$ . As we shall see later, bosonic systems in second quantization can be understood as spin systems in this sense with classical state space given by the nonnegative integers, i.e.,  $X_i = \mathbb{N}_0 := \{0, 1, 2, \dots\}$ . Moreover, there is a correspondence between fermionic systems in second quantization and quantum spin- $\frac{1}{2}$  systems via the Jordan-Wigner transformation, but this correspondence is not canonical.

Now the global quantum state space is defined as  $\mathcal{Q} := \bigotimes_i Q_i \simeq \mathbb{C}^{\mathcal{X}}$ , i.e., in the spin- $\frac{1}{2}$  case, we have  $\mathcal{Q} \simeq \bigotimes_i \mathbb{C}^2 \simeq \mathbb{C}^{(2^N)} \simeq L^2(\{-1, 1\}^N)$ . Thus each state  $|\psi\rangle \in \mathcal{Q}$  can be thought of as a  $\mathbb{C}$ -valued function  $\psi(x) = \psi(x_1, \dots, x_N)$ , i.e., a *wavefunction*. Here  $x_i \in \{-1, 1\}$ .

In the setting of first quantization, as discussed in section 3 below, we may think of  $i$  as an index for our *particles*, each with local classical state space  $X_i = \mathbb{R}^d$ , where

$d$  is the physical dimension. With  $N$  particles, the global classical state space is  $\mathcal{X} = \bigoplus_{i=1}^N \mathbb{R}^d = (\mathbb{R}^d)^N$ , hence (ignoring the spin degree of freedom quantum particles) the global quantum state space is  $\mathcal{Q} = L^2((\mathbb{R}^d)^N)$ , whose elements are functions of the form  $\psi(x) = \psi(x_1, \dots, x_N)$ , where  $x_i \in \mathbb{R}^d$  for all  $i$ . This is the ‘original’ wavefunction of elementary quantum mechanics, i.e., the wavefunction appearing in the many-body Schrödinger equation.

Now a Hamiltonian in the quantum setting is a Hermitian operator  $\hat{H} : \mathcal{Q} \rightarrow \mathcal{Q}$ . For future reference, we let the space of Hermitian operators on a vector space  $V$  be denoted by  $\mathbf{H}(V)$ , so  $\hat{H} \in \mathbf{H}(\mathcal{Q})$ . Since  $\mathcal{Q} \simeq \mathbb{C}^{\mathcal{X}}$ ,  $\mathcal{H}(\mathcal{Q})$  may alternatively be thought of as the set of complex Hermitian matrices  $(H(x, y)) \in \mathbb{C}^{\mathcal{X} \times \mathcal{X}}$ . Note that restriction to diagonal  $\hat{H}$  recovers the notion of a classical Hamiltonian  $H : \mathcal{X} \rightarrow \mathbb{R}$ .

We will discuss Hamiltonians in first quantization in section 3 below; here we discuss several examples in the quantum spin- $\frac{1}{2}$  setting. To this end, first recall the Pauli matrices:

$$\sigma^x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma^y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma^z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

which, together with the identity  $I_2$ , form a basis for  $\mathcal{H}(\mathbb{C}^2)$ . Now let  $\sigma_i^{x/y/z} \in \mathcal{H}(\bigotimes_i \mathbb{C}^2) \simeq \bigotimes_i \mathcal{H}(\mathbb{C}^2)$  be obtained by tensoring a copy of  $\sigma^{x/y/z}$  for the  $i$ -th site with the identity  $I_2$  on all the other sites, i.e., in matrix form

$$\sigma^{x/y/z}(x, y) = \sigma^{x/y/z}(x_i, y_i) \prod_{j \neq i} \delta_{x_i, y_i}.$$

(Note: the  $x/y/z$  notation for the Pauli matrices is unrelated to the notation  $x, y$  for the classical state space elements.)

Given a graph structure on the site indices, we may define two model Hamiltonians of interest—the transverse-field Ising (TFI) Hamiltonian and anti-ferromagnetic Heisenberg (AFH) Hamiltonian—as follows:

$$\hat{H}_{\text{TFI}} = -h \sum_i \sigma_i^x - \sum_{\langle i, j \rangle} \sigma_i^z \sigma_j^z$$

$$\hat{H}_{\text{AFH}} = \sum_{i \sim j} [\sigma_i^x \sigma_j^x + \sigma_i^y \sigma_j^y + \sigma_i^z \sigma_j^z].$$

In the TFI Hamiltonian,  $h$  is a scalar parameter. These Hamiltonians may be used to define quantum statistical-mechanical ensembles as we shall describe presently.

## 2.2 Quantum Gibbs states

The quantum analog of a probability measure is a density operator, i.e., a positive semidefinite Hermitian operator  $\rho : \mathcal{Q} \rightarrow \mathcal{Q}$  of unit trace. Let the space of density

operators on  $\mathcal{Q}$  be denoted  $\mathbf{D}(\mathcal{Q})$ , so in fact  $\rho \in \mathbf{D}(\mathcal{Q})$ . Via diagonalization, a density operator can be thought of as a choice of orthonormal basis, *plus* a probability measure over basis elements. Hence quantum ‘probability’ can be thought of in fact as a *generalization* of a classical probability on  $\mathcal{X}$ , which is recovered in the case of a diagonal density operator. Likewise, Hermitian operators  $\hat{O} \in \mathbf{H}(\mathcal{Q})$  generalize random variables  $\mathcal{X} \rightarrow \mathbb{R}$ , and the ‘quantum expectation’ is given by the trace  $\text{Tr}[\hat{O}\rho]$ . Physically, this value is the expected value of a quantum measurement of the Hermitian operator  $\hat{O}$  on a quantum system in state  $\rho$ .

A quantum Gibbs state is defined in terms of a Hamiltonian, i.e., an operator  $\hat{H} \in \mathbf{H}(\mathcal{Q})$ , possibly parametrized as  $\hat{H}[A]$ . Note that restriction to diagonal  $\hat{H}$  recovers the notion of a classical Hamiltonian  $H : \mathcal{X} \rightarrow \mathbb{R}$ . Now the quantum Gibbs operator is defined

$$\rho[A] = \frac{1}{Z[A]} \exp(-\hat{H}[A]),$$

where ‘exp’ denotes the operator exponential and  $Z[A] := \text{Tr} \left[ \exp(-\hat{H}[A]) \right]$ . Likewise we define the free energy  $\Omega[A] = -\log Z[A]$ , which (as we shall verify in Part V) satisfies the quantum Gibbs variational principle

$$\Omega[A] = \inf_{\rho \in \mathcal{D}(\mathcal{Q})} \left[ \text{Tr} \left( \hat{H}[A]\rho \right) - S(\rho) \right],$$

where the infimum is attained by  $\rho = \rho[A]$ , and  $S$  here denotes (with meaning clear from context) the von Neumann entropy

$$S(\rho) = -\text{Tr}[\rho \log \rho].$$

Here ‘log’ is the operator logarithm. Note that the von Neumann entropy recovers the Shannon entropy in the case of diagonal  $\rho$ .

Let us explicitly focus on an inverse temperature parameter  $\beta$ , i.e., define

$$\rho[\beta] = \frac{1}{Z[\beta]} \exp(-\beta \hat{H}),$$

for fixed  $\hat{H} \in \mathbf{H}(\mathcal{Q})$ . It can be verified by diagonalization that as  $\beta \rightarrow \infty$ , if  $\hat{H}$  has a unique (normalized) ground state (i.e., eigenvector with minimal eigenvalue)  $|\Phi_0\rangle \in \mathcal{Q}$ , then  $\rho[\beta] \rightarrow |\Phi_0\rangle\langle\Phi_0|$ . Hence quantum statistical mechanics at zero temperature recovers the problem of finding the ground state of a quantum many-body Hamiltonian, which is the quantum analog of optimization. Incidentally, any density operator of rank 1 is known as a pure state

### 3 First quantization and electronic structure

As mentioned earlier in section 2, the quantum state space for an  $N$ -particle system in first quantization is  $L^2((\mathbb{R}^d)^N) \simeq \bigotimes_{i=1}^N L^2(\mathbb{R}^d)$ , where  $d$  is the physical dimension in which the particles live and we have ignored the spin degree of freedom for

simplicity. (Note that for, e.g., spin- $\frac{1}{2}$  particles, the relevant Hilbert space is simply  $L^2((\mathbb{R}^d)^N; \mathbb{C}^2)$ , and our discussion can be extrapolated to this setting with minor modifications.) For notational clarity, here we shall use boldface to indicate elements  $\mathbf{x} = (x_1, \dots, x_N) \in (\mathbb{R}^d)^N$ , where the  $x_i \in \mathbb{R}^d$ .

Consider a classical Hamiltonian dynamical system [36] specified by the Hamiltonian  $H = H(\mathbf{x}, \mathbf{p})$  of position-momentum coordinates:

$$H(\mathbf{x}, \mathbf{p}) = \frac{1}{2} \sum_{i=1}^N |p_i|^2 + \sum_{i=1}^N V^{(1)}(x_i) + \frac{1}{2} \sum_{i \neq j} V^{(2)}(x_i - x_j).$$

Such a Hamiltonian specifies the classical dynamics of  $N$  particles that experience the same external potential  $V^{(1)}$  and interact via the pairwise potential  $V^{(2)}$ , as well as kinetic energy (for which the mass of the particles is scaled to unity).

Then canonical quantization (see, e.g., [36] for a discussion of deeper principles underlying this procedure) yields the Hamiltonian  $\hat{H}$

$$\hat{H} = -\frac{1}{2} \sum_{i=1}^N \Delta_{x_i} + \sum_{i=1}^N V^{(1)}(x_i) + \frac{1}{2} \sum_{i \neq j} V^{(2)}(x_i - x_j), \quad (3.1)$$

where  $\Delta_{x_i} = \sum_{j=1}^d \frac{\partial^2}{\partial (x_i)_j^2}$  is the Laplace operator for the  $i$ -th particle slot. (Note carefully that  $\sum_i \Delta_{x_i}$  is of course *not* an operator on  $L^2((\mathbb{R}^d)^N)$ . Hence it may seem more appropriate to look for wavefunctions in  $H^1((\mathbb{R}^d)^N)$ . But really the  $L^2$  inner product is the inner product we want. This is one motivation for the notion of a ‘rigged Hilbert space’ [27], but we sweep such analytical difficulties under the rug, as we shall discuss first quantization mostly in passing to second quantization.)

Now a Hamiltonian  $\hat{H}$  of the form (3.1) can model the physics of (relatively light) particles in the presence of the ionic potential

$$V^{(1)}(x) = - \sum_I \frac{Z_I}{|x - R_I|}$$

induced by fixed (relatively heavy) atomic nuclei, indexed by  $I$ , of charges  $Z_I$  and positions  $R_I \in \mathbb{R}^d$ , as well as a pairwise interaction specified by  $V^{(2)}$ ; hence we have implicitly assumed the Born-Oppenheimer approximation [16], in which the positions of the atomic nuclei are fixed for the computation of the quantum state of the remainder of the system. Identifying our quantum particles as electrons and specifying a repulsive pairwise Coulomb interaction

$$V^{(2)}(x_1, x_2) = \frac{1}{|x_1 - x_2|},$$

we arrive at the *electronic structure* problem, modulo one important caveat, toward which we now turn.



### 3.1 Identical particles

The caveat is the notion of *identical particles*, which is active in the case of electrons. To motivate this requires some further background.

Here we follow Dirac's notation, i.e., denoting wavefunctions  $\psi(\mathbf{x})$  and their adjoints via  $|\psi\rangle$  and  $\langle\psi|$ , respectively, we say that a wavefunction  $|\psi\rangle \in \mathcal{Q}$  is normalized if it satisfies

$$1 = \langle\psi|\psi\rangle = \int_{(\mathbb{R}^d)^N} |\psi(\mathbf{x})|^2 d\mathbf{x}.$$

Hence  $\rho = |\psi\rangle\langle\psi|$  defines a density operator in the sense of section 2. For a set  $S \in (\mathbb{R}^d)^N$ , the characteristic function of this set  $\chi_S(\mathbf{x})$  also defines a diagonal operator on  $\mathcal{Q}$  via pointwise multiplication. Then

$$\text{Tr}[\chi_S \rho] = \int_S |\psi(\mathbf{x})|^2 d\mathbf{x}.$$

Hence we interpret  $|\psi(\mathbf{x})|^2$  as the probability density function for locating our  $N$  particles at positions  $x_1, \dots, x_N$ , respectively.

For the purpose of this work, it can be taken as a fact of nature (though deeper justification can be made through quantum field theory; see, e.g., [36] that particles of certain species are identical, or indistinguishable, in the sense that

$$|\psi(\mathbf{x})|^2 = |\psi(\sigma \cdot \mathbf{x})|^2,$$

for all permutations  $\sigma \in S_N$ , where  $\sigma$  acts via  $[\sigma \cdot \mathbf{x}]_i = x_{\sigma(i)}$ .

Let us examine the consequence of such a condition. For such  $\psi$  and fixed  $\mathbf{x}$  such that  $\psi(\mathbf{x}) \neq 0$ , it must be the case that  $\psi(\sigma \cdot \mathbf{x}) = u_{\psi, \mathbf{x}}(\sigma)\psi(\mathbf{x})$  for some unique  $u_{\psi, \mathbf{x}}(\sigma) \in S^1$ , where  $S^1 \subset \mathbb{C}$  denotes the unit circle as a subset of the complex plane. This condition defines a map  $u_{\psi, \mathbf{x}} : S_N \rightarrow S^1$ , evidently a group homomorphism. It can be shown that there are only two such homomorphisms: the trivial homomorphism  $u_{\psi, \mathbf{x}}(\sigma) = 1$  and the signature homomorphism  $u_{\psi, \mathbf{x}}(\sigma) = \text{sgn}(\sigma)$ , which returns  $\pm 1$  for even/odd permutations, respectively.

Under the reasonable assumption that  $u_{\psi, \mathbf{x}}(\sigma)$  should depend continuously on  $\mathbf{x}$ , we arrive at two possibilities for  $\psi$ : either  $\psi(\sigma \cdot \mathbf{x}) = \psi(\mathbf{x})$  for all  $\mathbf{x}$ , i.e.,  $\psi$  is *symmetric*, or  $\psi(\sigma \cdot \mathbf{x}) = \text{sgn}(\sigma)\psi(\mathbf{x})$  for all  $\mathbf{x}$ , i.e.,  $\psi$  is *antisymmetric*. The former is the case of bosons and the latter of fermions. The subspace of symmetric functions is denoted  $\text{Sym}^N(L^2(\mathbb{R}^d)) \subset \bigotimes_{i=1}^N L^2(\mathbb{R}^d)$ , and the subspace of antisymmetric functions by  $\Lambda^N(L^2(\mathbb{R}^d)) \subset \bigotimes_{i=1}^N L^2(\mathbb{R}^d)$ . These are the quantum state spaces for  $N$ -particle systems of bosons, and fermions, respectively.

By contrast, certain quantum (composite) particle such as atomic nuclei can be modeled as 'boltzmannions' (note: the terminological usage is not universal), which are distinguishable and retain the full state space  $\bigotimes_{i=1}^N L^2(\mathbb{R}^d)$ .

### 3.2 Bases for boltzmannionic, fermionic, and bosonic state spaces

Given an orthonormal basis  $\{\phi_p\}_{p \in \mathcal{B}}$  for  $L^2(\mathbb{R}^d)$ , one can construct corresponding orthogonal bases for  $\bigotimes_{i=1}^N L^2(\mathbb{R}^d)$ ,  $\Lambda^N(L^2(\mathbb{R}^d))$ , and  $\text{Sym}^N(L^2(\mathbb{R}^d))$ . Of course, technically one needs a complete orthonormal *sequence* to exhaust all of  $L^2(\mathbb{R}^d)$ , but in practice one may also consider truncated bases and the relevant Galerkin projections of operators. Many basis sets adapted for electronic structure have been introduced in the quantum chemistry literature [102]. We will be somewhat casual about this point in the discussion and maintain notation that is agnostic with respect to it. In other words the basis index set  $\mathcal{B}$  can be either  $\{1, 2, 3, \dots\}$  or  $\{1, \dots, M\}$  for some finite  $M$ . Let  $\mathcal{H} \subset L^2(\mathbb{R}^d)$  be the (completion of the) span of  $\{\phi_p\}_{p \in \mathcal{B}}$ . This is our single-particle Hilbert space, after possible truncation, and the corresponding boltzmannionic, fermionic, and bosonic state spaces are denoted  $\bigotimes_{i=1}^N \mathcal{H}$ ,  $\Lambda^N(\mathcal{H})$ , and  $\text{Sym}^N(\mathcal{H})$ , respectively.

First, observe or recall that

$$\{\phi_{i_1} \otimes \cdots \otimes \phi_{i_N} : i_k \in \mathcal{B} \ \forall k = 1, \dots, N\}$$

is the standard induced orthonormal basis for  $\bigotimes_{i=1}^N \mathcal{H}$ . To construct the other bases, we must first introduce some new notation.

For  $f_1, \dots, f_N \in L^2(\mathbb{R}^d)$ , define

$$\bigodot_{i=1}^N f_i := f_1 \odot \cdots \odot f_N := \sum_{\sigma \in S_N} f_{\sigma(1)} \otimes \cdots \otimes f_{\sigma(N)} \in \bigodot_{i=1}^N \mathcal{H}$$

and

$$\bigwedge_{i=1}^N f_i := f_1 \wedge \cdots \wedge f_N := \sum_{\sigma \in S_N} \text{sgn}(\sigma) f_{\sigma(1)} \otimes \cdots \otimes f_{\sigma(N)} \in \bigwedge_{i=1}^N \mathcal{H}.$$

Then it is not hard to see that

$$\{\phi_{i_1} \odot \cdots \odot \phi_{i_N} : i_1 \leq i_2 \leq \cdots \leq i_N, \ i_k \in \mathcal{B} \ \forall k = 1, \dots, N\}$$

and

$$\{\phi_{i_1} \wedge \cdots \wedge \phi_{i_N} : i_1 < i_2 < \cdots < i_N, \ i_k \in \mathcal{B} \ \forall k = 1, \dots, N\}$$

form bases for  $S^N(\mathcal{H})$  and  $\Lambda^N(\mathcal{H})$ , respectively. (Note that  $\phi_{i_1} \wedge \cdots \wedge \phi_{i_N} \equiv 0$  if  $i_k = i_l$  for some  $k \neq l$ .)

### 3.3 Feynman path integral for boltzmannions

We will now describe how the Feynman path integral [34] can be used to transpose a boltzmannionic quantum-statistical ensemble to the setting of classical Gibbs measures. The discussion of bases in the preceding section 3.2 will not feature here.

Consider a Hamiltonian of the form

$$\hat{H} = -\frac{1}{2}\Delta + V(\mathbf{x})$$

on the boltzmannionic state space  $\mathcal{Q} = L^2((\mathbb{R}^d)^N)$ , where  $\Delta := \sum_{i=1}^N \Delta_{x_i}$ . Hence this form recovers (3.1) as a special case. Note that by identifying  $(\mathbb{R}^d)^N \simeq \mathbb{R}^K$ , where  $K = Nd$ , then the boltzmannionic many-particle system appears as nothing but a single-particle system in higher dimension. We shall accept this simplification going forward, denoting  $\mathbf{x} = (x_1, \dots, x_K)$ , where the  $x_k \in \mathbb{R}$ .

Consider the position operators  $\mathbf{X} = (X_1, \dots, X_K)$ , which are the diagonal multipliers specified by  $X_k(\mathbf{x}, \mathbf{x}) = x_k$ , and the momentum operators  $\mathbf{P} = (P_1, \dots, P_K)$ , defined by  $P_k := -i\partial_{x_k}$ . Note that  $-\Delta = \sum_{k=1}^K P_k^2$ , and  $V = V(\mathbf{X})$  in the continuous operator calculus. We shall denote by  $|\mathbf{x}\rangle$  the state of definite position  $\mathbf{x} = (x_1, \dots, x_K) \in \mathbb{R}^K$ , which is the simultaneous eigenstate of the position operators  $(X_1, \dots, X_K)$  with eigenvalues  $(x_1, \dots, x_K)$ . Meanwhile, let  $|\mathbf{p}\rangle$  denote the state of definite momentum  $\mathbf{p} = (p_1, \dots, p_K)$ , which is the simultaneous eigenstate of the position operators  $(P_1, \dots, P_K)$  with eigenvalues  $(p_1, \dots, p_K)$ . As (generalized) functions we have  $|\mathbf{x}\rangle = \delta_{\mathbf{x}}$  and  $|\mathbf{p}\rangle = e^{-i\mathbf{p}\cdot\mathbf{x}}$ . Note that  $|\mathbf{x}\rangle, |\mathbf{p}\rangle \notin L^2(\mathbb{R}^K)$ , and the right technical notion of ‘eigenfunction’ is a subtle matter. Here the theory of rigged Hilbert spaces [27] can come to the rescue, but our discussion is purely formal, and we shall elide such difficulties.

Now, as the sets  $\{|\mathbf{x}\rangle\}$  and  $\{|\mathbf{p}\rangle\}$  of eigenfunctions can each be formally viewed as an orthonormal basis, and we have the completeness relations

$$\text{Id}_{L^2(\mathbb{R}^K)} = \int_{\mathbb{R}^K} |\mathbf{x}\rangle\langle\mathbf{x}| d\mathbf{x}, \quad \text{Id}_{L^2(\mathbb{R}^K)} = \int_{\mathbb{R}^K} |\mathbf{p}\rangle\langle\mathbf{p}| d\mathbf{p},$$

and it follows that

$$\text{Id}_{L^2(\mathbb{R}^K)} = \int_{\mathbb{R}^K \times \mathbb{R}^K} d\mathbf{x} d\mathbf{p} |\mathbf{x}\rangle\langle\mathbf{x}|\mathbf{p}\rangle\langle\mathbf{p}| = \int_{\mathbb{R}^K \times \mathbb{R}^K} d\mathbf{x} d\mathbf{p} e^{-i\mathbf{p}\cdot\mathbf{x}} |\mathbf{x}\rangle\langle\mathbf{p}|. \quad (3.2)$$

Now we consider the partition function

$$Z[\beta] = \text{Tr} \left[ e^{-\beta\hat{H}} \right],$$

which we shall expand into the path integral via insertion of the completeness relations. To wit

$$\begin{aligned} Z[\beta] &= \int d\mathbf{x}_{(0)} \langle\mathbf{x}_{(0)}|e^{-\beta\hat{H}}|\mathbf{x}_{(0)}\rangle \\ &= \int d\mathbf{x}_{(0)} d\mathbf{p}_{(0)} \langle\mathbf{x}_{(0)}|\mathbf{p}_{(0)}\rangle\langle\mathbf{p}_{(0)}|e^{-\beta\hat{H}}|\mathbf{x}_{(0)}\rangle \\ &= \int d\mathbf{x}_{(0)} d\mathbf{p}_{(0)} e^{-i\mathbf{p}_{(0)}\cdot\mathbf{x}_{(0)}} \langle\mathbf{p}_{(0)}|e^{-\beta\hat{H}}|\mathbf{x}_{(0)}\rangle \end{aligned}$$

$$\begin{aligned}
&= \int d\mathbf{x}_{(0)} d\mathbf{p}_{(0)} e^{-i\mathbf{p}_{(0)} \cdot \mathbf{x}_{(0)}} \langle \mathbf{p}_{(0)} | e^{-\frac{1}{M}\beta\hat{H}} \dots e^{-\frac{1}{M}\beta\hat{H}} | \mathbf{x}_{(0)} \rangle \\
&= \int \prod_{m=0}^{M-1} d\mathbf{x}_{(m)} d\mathbf{p}_{(m)} e^{-\sum_{m=0}^{M-1} i\mathbf{p}_{(m)} \cdot \mathbf{x}_{(m)}} \\
&\quad \langle \mathbf{p}_{(0)} | e^{-\frac{1}{M}\beta\hat{H}} | \mathbf{x}_{(M-1)} \rangle \dots \langle \mathbf{p}_{(1)} | e^{-\frac{1}{M}\beta\hat{H}} | \mathbf{x}_{(0)} \rangle,
\end{aligned}$$

where in the last step we have inserted (3.2) between every pair of  $e^{-\frac{1}{M}\beta\hat{H}}$  operators. Now for  $M$  large, one has formally

$$e^{-\frac{1}{M}\beta\hat{H}} = e^{\frac{1}{M}\frac{1}{2}\sum_k P_k^2} e^{\frac{1}{M}V(\mathbf{x})} [\text{Id} + O(M^{-2})],$$

hence in the large  $M$  limit, one can replace

$$\begin{aligned}
\langle \mathbf{p}_{(m)} | e^{-\frac{1}{M}\beta\hat{H}} | \mathbf{x}_{(m-1)} \rangle &\approx \langle \mathbf{p}_{(m)} | e^{-\frac{\beta}{M}\frac{1}{2}\sum_k P_k^2} e^{-\frac{\beta}{M}V(\mathbf{x})} | \mathbf{x}_{(m-1)} \rangle \\
&= e^{-\frac{\beta}{M}(\frac{1}{2}|\mathbf{p}_{(m)}|^2 + V(\mathbf{x}_{(m-1)}))} \langle \mathbf{p}_{(m)} | \mathbf{x}_{(m-1)} \rangle \\
&= e^{-\frac{\beta}{M}(\frac{1}{2}|\mathbf{p}_{(m)}|^2 + V(\mathbf{x}_{(m-1)}))} e^{i\mathbf{p}_{(m)} \cdot \mathbf{x}_{(m-1)}}
\end{aligned}$$

where we interpret  $m$  modulo  $M$ , so

$$\begin{aligned}
Z[\beta] &= \lim_{M \rightarrow \infty} \int \prod_{m=0}^{M-1} d\mathbf{x}_{(m)} d\mathbf{p}_{(m)} e^{-\sum_{m=0}^{M-1} i\mathbf{p}_{(m)} \cdot (\mathbf{x}_{(m)} - \mathbf{x}_{(m-1)}) - \frac{\beta}{M} \sum_{m=0}^{M-1} [\frac{1}{2}|\mathbf{p}_{(m)}|^2 + V(\mathbf{x}_{(m)})]} \\
&= \int D\mathbf{x}_{\text{per}}(\cdot) D\mathbf{p}(\cdot) e^{-\int_0^\beta [\frac{1}{2}|\mathbf{p}(\tau)|^2 + i\mathbf{p}(\tau) \cdot \partial_\tau \mathbf{x}(\tau) + V(\mathbf{x}(\tau))] d\tau},
\end{aligned}$$

where the limit is understood (for now) only formally and  $D\mathbf{x}_{\text{per}}(\cdot)$  is thought of as the infinite-dimensional Lebesgue measure  $\prod_{\tau \in [0, \beta]} d\mathbf{x}(\tau)$  on *periodic* paths, i.e., paths satisfying  $\mathbf{x}(0) = \mathbf{x}(\beta)$ . Meanwhile,  $D\mathbf{p}(\cdot)$  can be understood as the infinite-dimensional Lebesgue measure  $\prod_{\tau \in [0, \beta]} d\mathbf{p}(\tau)$ , and here via the construction the periodicity requirement is relaxed. We integrate out the  $\mathbf{p}(\tau)$  path via an the formula for Gaussian integrals (formally ‘extrapolated’ to our infinite-dimensional setting):

$$\int D\mathbf{p}(\cdot) e^{-\frac{1}{2}\int_0^\beta |\mathbf{p}(\tau)|^2 d\tau - i\int_0^\beta \mathbf{p}(\tau) \cdot \partial_\tau \mathbf{x}(\tau) d\tau} = \left[ \lim_{M \rightarrow \infty} (2\pi)^{M/2} \right] e^{-\frac{1}{2}\int_0^\beta |\partial_\tau \mathbf{x}(\tau)|^2 d\tau}.$$

The ‘infinite preconstant’  $\lim_{M \rightarrow \infty} (2\pi)^{M/2}$  can be ignored as a physically insignificant contribution to the partition function, or, as one prefers, it can be formally lumped into the measure  $D\mathbf{x}(\cdot)$ , yielding

$$Z[\beta] = \int_{\mathbf{x}(0)=\mathbf{x}(\beta)} D\mathbf{x}(\cdot) e^{-\frac{1}{2}\int_0^\beta [|\partial_\tau \mathbf{x}(\tau)|^2 + V(\mathbf{x}(\tau))] d\tau},$$

where we introduce the notation  $\int D\mathbf{x}_{\text{per}}(\cdot) \dots = \int_{\mathbf{x}(0)=\mathbf{x}(\beta)} D\mathbf{x}(\cdot) \dots$ . Notice from the expression that this quantity can be viewed as a partition function for a Gibbs measure on the space of functions  $[0, \beta] \rightarrow \mathbb{R}$ , i.e., as an infinite-dimensional limit of the setting discussed above in section 1. In fact, this notion can be made precise via the Wiener measure; see, e.g., [34].

## 4 Second quantization

### 4.1 The Fock space

Second quantization considers an enlargement, called the Fock space, of any individual fermionic/bosonic  $N$ -particle state space. Indeed, the fermionic/bosonic Fock space can be defined as the (completion of the) direct sum of *all* fermionic/bosonic  $N$ -particle state spaces:

$$\mathcal{F}_f := \overline{\Lambda(\mathcal{H})}, \quad \mathcal{F}_b := \overline{\text{Sym}(\mathcal{H})},$$

where  $\Lambda(\mathcal{H}) := \bigoplus_{N=0}^{\infty} \Lambda^N(\mathcal{H})$  and  $\text{Sym}(\mathcal{H}) := \bigoplus_{N=0}^{\infty} \text{Sym}^N(\mathcal{H})$  are the exterior and symmetric algebras, respectively. For simplicity we shall further denote  $\mathcal{F}_f^{(N)} := \Lambda^N(\mathcal{H})$  and  $\mathcal{F}_b^{(N)} := \text{Sym}^N(\mathcal{H})$ . Note that we have conflated the notions of the exterior and symmetric algebras (technically defined as quotients of the tensor algebra) with the equivalent notions, respectively, of the antisymmetric and symmetric subspaces of the tensor algebra.

Even if one is only interested in a definite particle number, a simplified picture of fermionic/bosonic quantum mechanics emerges from this transformation. Moreover, the Fock space allows one to consider states of indefinite particle number and to understand the physics of the system in terms of the creation and annihilation of particles. This perspective is fundamental to quantum field theory and can in fact ought to be viewed as more fundamental than the first-quantized perspective; however, as the names suggest, first quantization preceded second quantization historically.

Now bases for  $\mathcal{F}_f$  and  $\mathcal{F}_b$  are given, respectively, by

$$\{\phi_{i_1} \wedge \cdots \wedge \phi_{i_N} : i_1 < i_2 < \cdots < i_N, N = 0, 1, 2, \dots\}$$

and

$$\{\phi_{i_1} \odot \cdots \odot \phi_{i_N} : i_1 \leq i_2 \leq \cdots \leq i_N, N = 0, 1, 2, \dots\}.$$

By convention,  $\Lambda^0(\mathcal{H}) = \text{Sym}^0(\mathcal{H}) = \mathbb{C}$ , and the basis element in the case  $N = 0$  (i.e., the ‘empty’ wedge product) is called the vacuum state, denoted by  $|-\rangle$ . Note carefully the distinction between the vacuum state and the zero vector, denoted 0.

For  $\mathbf{n} = (n_p)_{p \in \mathcal{B}} \in \{0, 1, 2, \dots\}^{\mathcal{B}} = (\mathbb{N}_0)^{\mathcal{B}}$ , one defines

$$|\mathbf{n}\rangle_f := \bigwedge_p (\phi_p)^{\wedge n_p}, \quad |\mathbf{n}\rangle_b := \bigodot_p (\phi_p)^{\odot n_p},$$

where, e.g., we denote  $(\phi_p)^{\odot n_p} = \underbrace{\phi_p \odot \cdots \odot \phi_p}_{n_p \text{ times}}$ , and if  $n_p = 0$  then the factor is omitted. Then observe that

$$\{|\mathbf{n}\rangle_f : \mathbf{n} \in \{0, 1\}^{\mathcal{B}}\}, \quad \{|\mathbf{n}\rangle_b : \mathbf{n} \in (\mathbb{N}_0)^{\mathcal{B}}\}$$

are alternative representations of the same bases for  $\mathcal{F}_f$  and  $\mathcal{F}_b$ , respectively. We refer to these bases as the occupation number bases, because for an element  $|\mathbf{n}\rangle$ ,  $n_p$  indicates the number of particles occupying the  $p$ -th state  $\phi_p$ .

Now any element  $|\psi\rangle \in \mathcal{F}_{f/b}$  can be written  $\sum_{\mathbf{p} \in \mathcal{B}} \psi(\mathbf{n}) |\mathbf{n}\rangle$ , hence can equivalently be viewed as a function  $\mathbf{n} \mapsto \psi(\mathbf{n})$ , with  $\psi \in L^2(\{0, 1\}^{\mathcal{B}})$  or  $\psi \in L^2((\mathbb{N}_0)^{\mathcal{B}})$ . Recall that in first quantization,  $|\psi(\mathbf{x})|^2$  indicates the likelihood of finding particles at positions  $(x_1, \dots, x_N)$ . Second quantization turns this conceit on its head; for second-quantized wavefunction  $\psi$ ,  $|\psi(\mathbf{n})|^2$  indicates the probability of finding, for each  $p \in \mathcal{B}$ ,  $n_p$  particles in state  $\phi_p$ . Hence the the basis functions  $\phi_p$  are the ‘sites’ of our model, as viewed through the lens of statistical mechanics.

Hence  $\mathcal{F}_f \simeq (\mathbb{C}^2)^{\mathcal{B}}$  and  $\mathcal{F}_b \simeq (\mathbb{C}^{\mathbb{N}_0})^{\mathcal{B}}$  via the correspondence(s)  $|\mathbf{n}\rangle \leftrightarrow e_{n_1} \otimes e_{n_2} \otimes \dots$ , where the  $e_k \in \mathbb{C}^2$  are the (zero-indexed) standard unit vectors. Hence fermionic and bosonic ensembles can be viewed as quantum spin systems in the sense of section 2. In the case of fermions, as we shall see, this isomorphism is not canonical in that it depends on the numerical ordering of the basis functions  $\phi_p$ . We shall examine further difficulties in section 4.4 below.

## 4.2 The creation and annihilation operators

All operators on the Fock space can be written in terms of the so-called creation and annihilation operators, denoted (respectively) by  $c_p^\dagger$  and  $c_p$  in the fermionic case and by  $b_p^\dagger$  and  $b_p$  in the bosonic case. When the context is clear it is also common to use  $a_p^\dagger$  and  $a_p$  for either case.

Now we define the fermionic creation operator  $c_p^\dagger$  via its action on a basis for  $\mathcal{F}_f$ :

$$c_p^\dagger [\phi_{i_1} \wedge \dots \wedge \phi_{i_N}] = \phi_p \wedge \phi_{i_1} \wedge \dots \wedge \phi_{i_N}.$$

For the  $N = 0$  case we understand this to mean  $c_p^\dagger |-\rangle = \phi_p$ .

Meanwhile, the annihilation operator  $c_p$  can be defined as the formal adjoint of  $c_p^\dagger$ , and it can be shown without difficulty that

$$c_p [\phi_p \wedge \phi_{i_1} \wedge \dots \wedge \phi_{i_N}] = \begin{cases} \phi_{i_1} \wedge \dots \wedge \phi_{i_N}, & p \neq i_k \ \forall k = 1, \dots, N \\ 0, & \text{otherwise.} \end{cases}$$

Moreover  $c_p |-\rangle = 0$  for all  $p$ .

Meanwhile for  $i_1, \dots, i_N \in \mathcal{B}$  and  $n_p := |\{k : i_k = p\}|$ , we define

$$b_p^\dagger [\phi_{i_1} \odot \dots \odot \phi_{i_N}] = \sqrt{n_p + 1} \phi_p \odot \phi_{i_1} \odot \dots \odot \phi_{i_N}$$

and extend by linearity. The formal adjoint  $b_p$  satisfies

$$b_p [\phi_p \odot \phi_{i_1} \odot \dots \odot \phi_{i_N}] = \sqrt{n_p + 1} \phi_{i_1} \odot \dots \odot \phi_{i_N},$$

where still  $n_p = |\{k : i_k = p\}|$ , and  $b_p |-\rangle = 0$  for all  $p$ .

It is not hard to verify that

$$\{c_p^\dagger, c_q^\dagger\} = \{c_p, c_q\} = 0, \quad \{c_p, c_q^\dagger\} = \delta_{ij},$$

where  $\{\cdot, \cdot\}$  denotes the anticommutator, and

$$[b_p^\dagger, b_q^\dagger] = [b_p, b_q] = 0, \quad [b_p, b_q^\dagger] = \delta_{ij},$$

where  $[\cdot, \cdot]$  denote the commutator. These are the canonical (anti)commutation relations, known for short as the CAR/CCR.

More abstractly, the set of operators  $\text{End}(\mathcal{F}_{\text{f/b}})$  can be viewed as the star-algebra  $\mathcal{A}_{\text{f/b}}$  generated by the  $a_p^\dagger$ , subject to the CAR/CCR. (A star-algebra is just an algebra with a star (or adjoint) operation satisfying certain predictable axioms; see, e.g., [18]) Meanwhile, roughly speaking, the Fock space can be thought of abstractly as the orbit of a vacuum state  $|-\rangle$  under an action of  $\mathcal{A}_{\text{f/b}}$ , with equivalences defined via the CCR/CAR and the relations  $a_p|-\rangle = 0$ .

Indeed, note that in both the fermionic and bosonic cases, we can write

$$|\mathbf{n}\rangle = (a_1^\dagger)^{n_1} (a_2^\dagger)^{n_2} \dots |-\rangle.$$

Hence all objects of interest to us (i.e., Hamiltonians and wavefunctions) can be understood purely in terms of the algebra  $\mathcal{A}_{\text{f/b}}$ , together with the vacuum state, satisfying certain algebraic relations.

To conclude this section, we define the number operators  $\hat{n}_p = a_p^\dagger a_p$  and the total number operator  $\hat{N} = \sum_p \hat{n}_p$ . Observe that  $|\mathbf{n}\rangle$  is an eigenstate of  $\hat{n}_p$  with eigenvalue  $n_p$  for all  $p$ . Moreover, the  $N$ -particle subspaces  $\mathcal{F}_{\text{f/b}}^{(N)}$  are precisely the  $N$ -eigenspaces of  $\hat{N}$ .

### 4.3 Second-quantized operators

Observe that an operator of the form (3.1) has the essential structure

$$\hat{H} = \sum_{k=1}^N \hat{O}_k^{(1)} + \sum_{k \neq l}^N \hat{O}_{kl}^{(2)},$$

where  $\hat{O}_k^{(1)}$  is a one-body operator on  $\mathcal{Q}$  obtained by tensoring a copy of some operator  $\hat{O}^{(1)}$  on  $\mathcal{H}$  for site  $i$  with copies of the identity for all other tensor factors  $1, \dots, N$ , and  $\hat{O}_{kl}^{(2)}$  is a two-body operator on  $\mathcal{Q}$  obtained by tensoring a copy of some operator  $\hat{O}^{(2)}$  on  $\mathcal{H} \otimes \mathcal{H}$  for sites  $k, l$  with copies of the identity for all other tensor factor  $1, \dots, N$ .

We will show how to write such operators (which preserve the (anti)symmetry of the wavefunction) in terms of the creation and annihilation operators.

**Lemma 1.** After restriction to  $\mathcal{F}_{\text{f/b}}^{(N)}$ , we have

$$\sum_{k=1}^N \hat{O}_k^{(1)} = \sum_{p,q \in \mathcal{B}} O_{pq} a_p^\dagger a_q,$$

where  $O_{pq} = \langle \phi_p | \hat{O}^{(1)} | \phi_q \rangle$ .

*Proof.* Though the fermionic and bosonic cases are very similar, it is less confusing to treat them separately. Let us consider the case of fermions first.

We check the claimed operator equality by checking on the basis element  $\phi_{i_1} \wedge \cdots \wedge \phi_{i_N}$ , and we begin by applying the first-quantized operator  $\sum_{k=1}^N \hat{O}_k^{(1)}$  as follows, where for simplicity we write  $\hat{O} = \hat{O}^{(1)}$ :

$$\begin{aligned} & \left( \sum_k \hat{O}_k^{(1)} \right) [\phi_{i_1} \wedge \cdots \wedge \phi_{i_N}] \\ &= \sum_k \sum_{\sigma \in S_N} \text{sgn}(\sigma) \hat{O}_k^{(1)} [\phi_{i_{\sigma(1)}} \otimes \cdots \otimes \phi_{i_{\sigma(N)}}] \\ &= \sum_{\sigma \in S_N} \sum_k \text{sgn}(\sigma) [\phi_{i_{\sigma(1)}} \otimes \cdots \otimes (\hat{O} \phi_{i_{\sigma(k)}}) \otimes \cdots \otimes \phi_{i_{\sigma(N)}}] \\ &\stackrel{(\star)}{=} \sum_k \sum_{\sigma \in S_N} \text{sgn}(\sigma) [\phi_{i_{\sigma(1)}} \otimes \cdots \otimes (\hat{O} \phi_{i_k}) \otimes \cdots \otimes \phi_{i_{\sigma(m)}}] \\ &= \sum_{k=1}^N \phi_{i_1} \wedge \cdots \wedge (\hat{O} \phi_{i_k}) \wedge \cdots \wedge \phi_{i_N}. \end{aligned}$$

In the step  $(\star)$  we changed the inner summation variable  $k$  according to  $k \mapsto \sigma^{-1}(k)$  and then exchanged the sums.

Now we will show that we obtain the same expression by applying  $\sum_{p,q \in \mathcal{B}} O_{pq} a_p^\dagger a_q$ . First observe that according to the definition of  $O_{pq}$ , we have  $\hat{O} \phi_q = \sum_p O_{pq} \phi_p$ . Then compute:

$$\begin{aligned} \sum_{pq} O_{pq} a_p^\dagger a_q [\phi_{i_1} \wedge \cdots \wedge \phi_{i_N}] &= \sum_p \sum_{k=1}^N O_{pi_k} a_p^\dagger a_{i_k} [\phi_{i_1} \wedge \cdots \wedge \phi_{i_N}] \\ &= \sum_p \sum_{k=1}^N (-1)^{k-1} O_{pi_k} a_p^\dagger \underbrace{[\phi_{i_1} \wedge \cdots \wedge \phi_{i_N}]}_{\text{factor } k \text{ omitted}} \\ &= \sum_{k=1}^N (-1)^{k-1} \sum_p O_{pi_k} [\phi_p \wedge \underbrace{\phi_{i_1} \wedge \cdots \wedge \phi_{i_N}}_{\text{factor } k \text{ omitted}}] \\ &= \sum_{k=1}^N (-1)^{k-1} [(\hat{O} \phi_{i_k}) \wedge \underbrace{\phi_{i_1} \wedge \cdots \wedge \phi_{i_N}}_{\text{factor } k \text{ omitted}}] \end{aligned}$$



$$= \sum_{k=1}^N \phi_{i_1} \wedge \cdots \wedge (\hat{O}\phi_{i_k}) \wedge \cdots \wedge \phi_{i_N}.$$

This completes the proof for fermions. For bosons, the proof of the identity

$$\left( \sum_k \hat{O}_k^{(1)} \right) [\phi_{i_1} \odot \cdots \odot \phi_{i_N}] = \sum_{k=1}^N \phi_{i_1} \odot \cdots \odot (\hat{O}\phi_{i_k}) \odot \cdots \odot \phi_{i_N}$$

is identical to the proof of the first identity above, up to the removal of  $\text{sgn}(\sigma)$  from the computation. For the second identity, fix  $i_1, \dots, i_N$  and define  $n_q := |\{k : i_k = q\}|$  for all  $q$ . Then compute:

$$\begin{aligned} \sum_{pq} O_{pq} a_p^\dagger a_q [\phi_{i_1} \odot \cdots \odot \phi_{i_N}] &= \sum_p \sum_{k=1}^N \frac{1}{n_{i_k}} O_{pi_k} a_p^\dagger a_{i_k} [\phi_{i_1} \odot \cdots \odot \phi_{i_N}] \\ &= \sum_p \sum_{l=1}^N \frac{1}{\sqrt{n_{i_k}}} O_{pi_k} a_p^\dagger \underbrace{[\phi_{i_1} \odot \cdots \odot \phi_{i_N}]}_{\text{factor } k \text{ omitted}} \\ &= \sum_{l=1}^N \sum_p O_{pi_k} [\phi_p \odot \underbrace{\phi_{i_1} \odot \cdots \odot \phi_{i_N}}_{\text{factor } k \text{ omitted}}] \\ &= \sum_{l=1}^N (\hat{O}\phi_{i_k}) \odot \underbrace{\phi_{i_1} \odot \cdots \odot \phi_{i_N}}_{\text{factor } k \text{ omitted}} \\ &= \sum_{l=1}^N \phi_{i_1} \odot \cdots \odot (\hat{O}\phi_{i_k}) \odot \cdots \odot \phi_{i_N}. \end{aligned}$$

□

**Lemma 2.** *After restriction to  $\Lambda^N(\mathcal{H})$  or  $\text{Sym}^N(\mathcal{H})$  according to the whether the case is fermionic or bosonic, we have*

$$\sum_{k \neq l} \hat{O}_{kl}^{(2)} = \sum_{p,q,r,s \in \mathcal{B}} O_{pq,rs} a_p^\dagger a_q^\dagger a_s a_r,$$

where  $O_{pq,rs} = \langle \phi_p \phi_q | \hat{O}^{(2)} | \phi_r \phi_s \rangle$ . (Here, e.g.,  $|\phi_r \phi_s\rangle$  denotes  $\phi_r \otimes \phi_s$ .)

*Proof.* We shall only give the proof for fermions; the bosonic proof follows by making similar changes as those made in Lemma 1 above. First note that we can assume without loss of generality that  $\hat{O}^{(2)} = \hat{A} \otimes \hat{B}$ , so  $O_{pq,rs} = A_{pr} B_{qs}$ , where  $A_{pr} = \langle \phi_p | \hat{A} | \phi_q \rangle$  and  $B_{qs} = \langle \phi_q | \hat{B} | \phi_s \rangle$ . (The general case follows from linear combination of such operators.)

As in the proof of Lemma 1, fix a basis element  $\phi_{i_1} \wedge \cdots \wedge \phi_{i_N}$  and compute

$$\begin{aligned}
& \sum_{k \neq l}^N \hat{O}_{kl}^{(2)} [\phi_{i_1} \wedge \cdots \wedge \phi_{i_N}] \\
&= \sum_k \sum_{\sigma \in S_N} \text{sgn}(\sigma) \hat{O}_{kl}^{(2)} [\phi_{i_{\sigma(1)}} \otimes \cdots \otimes \phi_{i_{\sigma(N)}}] \\
&= \sum_{\sigma \in S_N} \sum_{k \neq l} \text{sgn}(\sigma) [\phi_{i_{\sigma(1)}} \otimes \cdots \otimes (\hat{A}\phi_{i_{\sigma(k)}}) \otimes \cdots \otimes (\hat{B}\phi_{i_{\sigma(l)}}) \otimes \cdots \otimes \phi_{i_{\sigma(N)}}] \\
&\stackrel{(\star)}{=} \sum_{k \neq l} \sum_{\sigma \in S_N} \text{sgn}(\sigma) [\phi_{i_{\sigma(1)}} \otimes \cdots \otimes (\hat{A}\phi_{i_k}) \otimes \cdots \otimes (\hat{B}\phi_{i_l}) \otimes \cdots \otimes \phi_{i_{\sigma(N)}}] \\
&= \sum_{k \neq l} \phi_{i_1} \wedge \cdots \wedge (\hat{A}\phi_{i_k}) \wedge \cdots \wedge (\hat{B}\phi_{i_l}) \wedge \cdots \wedge \phi_{i_N}.
\end{aligned}$$

In the step  $(\star)$  we changed the inner summation variable  $k \neq l$  according to  $(k, l) \mapsto (\sigma^{-1}(k), \sigma^{-1}(l))$  then exchanged the sums.

Finally, we apply the second-quantized operator to the fixed basis element  $\phi_{i_1} \wedge \cdots \wedge \phi_{i_N}$ . For visual clarity, we first introduce some auxiliary notation. For  $k \neq l$ , let  $\varepsilon_{kl}$  be the sign of the permutation that permutes  $i_k, i_l$  to the first two spots of  $(i_1, \dots, i_N)$  without changing the rest of the ordering. Hence  $\varepsilon_{kl} = (-1)^{k-1}(-1)^{l-1}$  if  $k > l$  and  $\varepsilon_{kl} = -(-1)^{k-1}(-1)^{l-1}$  if  $k < l$ . Then compute

$$\begin{aligned}
& \sum_{pqrs} O_{pq,rs} a_p^\dagger a_q^\dagger a_s a_r [\phi_{i_1} \wedge \cdots \wedge \phi_{i_N}] \\
&= \sum_{pqr} \sum_{k=1}^N (-1)^{k-1} O_{pq,i_k s} a_p^\dagger a_q^\dagger a_s \underbrace{[\phi_{i_1} \wedge \cdots \wedge \phi_{i_N}]}_{\text{factor } k \text{ omitted}} \\
&= \sum_{pq} \sum_{k \neq l} \varepsilon_{kl} O_{pq,i_k i_l} a_p^\dagger a_q^\dagger \underbrace{[\phi_{i_1} \wedge \cdots \wedge \phi_{i_N}]}_{\text{factors } k, l \text{ omitted}} \\
&= \sum_{pq} \sum_{k \neq l} \varepsilon_{kl} O_{pq,i_k i_l} [\phi_p \wedge \phi_q \wedge \underbrace{\phi_{i_1} \wedge \cdots \wedge \phi_{i_N}}_{\text{factors } k, l \text{ omitted}}] \\
&= \sum_{k \neq l} \varepsilon_{kl} \sum_{pq} A_{pi_k} B_{qi_l} [\phi_p \wedge \phi_q \wedge \underbrace{\phi_{i_1} \wedge \cdots \wedge \phi_{i_N}}_{\text{factors } k, l \text{ omitted}}] \\
&= \sum_{k \neq l} \varepsilon_{kl} [(\hat{A}\phi_{i_k}) \wedge (\hat{B}\phi_{i_l}) \wedge \underbrace{\phi_{i_1} \wedge \cdots \wedge \phi_{i_N}}_{\text{factors } k, l \text{ omitted}}] \\
&= \sum_{k \neq l} \phi_{i_1} \wedge \cdots \wedge (\hat{A}\phi_{i_k}) \wedge \cdots \wedge (\hat{B}\phi_{i_l}) \wedge \cdots \wedge \phi_{i_N}
\end{aligned}$$

Note that throughout the computation, the sign factor isn't really doing anything but hanging out and waiting to help us anticommute things back to the middle of the wedge product.  $\square$

More generally, one may consider  $m$ -body operators for  $m \leq N$  with notation analogous to the above. Then the following general result should be apparent from the proofs of Lemmas 1 and 2.

**Lemma 3.** *After restriction to  $\Lambda^N(\mathcal{H})$  or  $\text{Sym}^N(\mathcal{H})$  according to the whether the case is fermionic or bosonic, we have*

$$\sum_{k_1, \dots, k_m \text{ distinct}} \hat{O}_{k_1 \dots k_m}^{(m)} = \sum_{p_1, \dots, p_m, q_1, \dots, q_m} O_{p_1 \dots p_m, q_1 \dots q_m} a_{p_1}^\dagger \dots a_{p_m}^\dagger a_{q_m} \dots a_{q_1},$$

where  $O_{p_1 \dots p_m, q_1 \dots q_m} = \langle \phi_{p_1} \dots \phi_{p_m} | \hat{O}^{(m)} | \phi_{q_1} \dots \phi_{q_m} \rangle$ .

## 4.4 The Jordan-Wigner transformation

In this section we will first focus on the fermionic case. We have already seen how  $\mathcal{F}_f$  can be put into correspondence with  $(\mathbb{C}^2)^\mathcal{B}$  via the  $|\mathbf{n}\rangle \leftrightarrow e_{n_1} \otimes e_{n_2} \otimes \dots$ , which puts the occupation number basis for  $\mathcal{F}_f$  in correspondence with the standard basis of  $(\mathbb{C}^2)^\mathcal{B}$ . (Recall that here the  $e_k \in \mathbb{C}^2$  are the *zero-indexed* standard unit vectors.)

One verifies that under this isomorphism  $c_p^\dagger$  (abusing notation slightly by overloading the notation for  $c_p^\dagger$ ) can be written

$$c_p^\dagger = \underbrace{\sigma^z \otimes \dots \otimes \sigma^z}_{p-1 \text{ factors}} \otimes \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \otimes I_2 \otimes I_2 \otimes \dots,$$

from which the corresponding formula for  $c_p$  is apparent. This transformation from fermionic creation and annihilation operators to quantum spin- $\frac{1}{2}$  operators (or vice versa) is known as the Jordan-Wigner transformation.

Observe that, due to the that a reordering of the index  $p$  *does not* commute with the corresponding reordering of the tensor factors in  $c_p^\dagger$ . Hence our representation of fermionic operators in terms of quantum spin operators depends importantly on the choice of ordering of the basis. Moreover, a one-body Hermitian operator such as  $a_p^\dagger a_q + a_q^\dagger a_p$ , as might appear in a second-quantized Hamiltonian, acts nontrivially on *all spins* between the indices  $p$  and  $q$ , inclusive. In particular, any physical locality of the fermionic Hamiltonian may be destroyed by Jordan-Wigner transformation.

By contrast the bosonic creation operator can be written as an operator on  $(\mathbb{C}^{\mathbb{N}_0})^\mathcal{B}$  via

$$b_p^\dagger = \underbrace{\text{Id} \otimes \dots \otimes \text{Id}}_{p-1 \text{ factors}} \otimes \begin{pmatrix} 0 & 0 & 0 & 0 & \dots \\ 1 & 0 & 0 & 0 & \dots \\ 0 & \sqrt{2} & 0 & 0 & \dots \\ 0 & 0 & \sqrt{3} & 0 & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix} \otimes \text{Id} \otimes \text{Id} \otimes \dots$$

It is evidently more natural to view bosonic operators in this way, which is independent of the basis ordering and which preserves physical locality.

## 4.5 Canonical transformations and noninteracting problems

We now discuss how a change of the orthonormal basis for  $\mathcal{H}$  in first quantization may be understood as a transformation (namely, a canonical transformation) of the creation and annihilation operators in second quantization. Consider a basis  $\{\tilde{\phi}_p\}$  for  $\mathcal{H}$ , written in terms of the original basis  $\{\phi_p\}$  via

$$\tilde{\phi}_p = \sum_q U_{pq} \phi_q,$$

where  $U = (U_{pq})$  is unitary. Consider the fermionic case for concreteness (the bosonic case is almost identical), and recall that for  $\Phi \in \mathcal{F}_f = \Lambda(\mathcal{H})$ , the action of the creation operator  $c_p^\dagger$  is given by

$$c_p^\dagger \Phi = \phi_p \wedge \Phi.$$

Now let  $\tilde{c}_p^\dagger$  denote the set of creation operators induced by the basis  $\{\tilde{\phi}_p\}$ . Then

$$\tilde{c}_p^\dagger \Phi = \tilde{\phi}_p \wedge \Phi = \left( \sum_q U_{pq} \phi_q \right) \wedge \Phi = \sum_q U_{pq} c_p^\dagger \Phi$$

for all  $\Phi$ , hence  $\tilde{c}_p^\dagger = \sum_q U_{pq} c_p^\dagger$ . By similar reasoning for bosons and conjugation, we obtain the general formulas

$$\tilde{a}_p^\dagger = \sum_q U_{pq} a_p^\dagger, \quad a_p = \sum_q \bar{U}_{pq} a_q$$

for canonical transformation. The canonical transformation can be thought of in terms of the CCR/CAR alone, without any reference to a first-quantized setting. (Indeed, this is a more fundamental point of view, physically.)

The canonical transformation allows us to completely solve so-called *noninteracting* systems, specified by Hamiltonians of the form

$$\hat{H} = \sum_{pq} A_{pq} a_p^\dagger a_q$$

because after a suitable canonical transformation, we can assume that  $A$  is diagonal, i.e., we can assume  $\hat{H} = \sum_p u_p \hat{n}_p$ , so the states decouple (as the  $\hat{n}_p$  commute). Such systems are derived from first-quantized Hamiltonians that lack any many-body terms. In the context of many-body physics, noninteracting systems may be thought of as ‘trivial’ and can often be viewed as a building block or point of departure for methods designed for many-body systems.

## 4.6 Second-quantized model Hamiltonians

Second quantization allows us to consider—in addition to Hamiltonians derived from first quantization via the choice of an orbital basis—various model problems that may capture physical phenomenology of interest.

Of particular note is the fermionic Hubbard model, whose states we enumerate via the orbital-spin index  $(i, \sigma)$ , where  $i = 1, \dots, M$ ,  $\sigma = \uparrow, \downarrow$ .

$$\hat{H} = -t \sum_{ij\sigma} A_{ij} a_{i\sigma}^\dagger a_{j\sigma} + U \sum_i \hat{n}_{i\uparrow} \hat{n}_{i\downarrow},$$

where  $A_{ij}$  is the adjacency matrix of a graph with vertex set  $\{1, \dots, M\}$ , e.g., a square lattice.

More generally, one can consider a ‘generalized Coulomb model’ of the form

$$\hat{H} = \sum_{ij\sigma} h_{ij} a_{i\sigma}^\dagger a_{j\sigma} + \sum_{ij\sigma\tau} U_{ij} \hat{n}_{i\sigma} \hat{n}_{j\tau},$$

which includes in particular the Hubbard model and variants with longer-range interactions. In fact, via certain choices of orbital bases such as the recently introduced Gausslets [105], electronic structure problems in the continuum can be mapped to second-quantized Hamiltonians of this form.

Recall that in electronic structure, the most general Hamiltonian of interest (arising from an arbitrary choice of orbital basis) can be written

$$\hat{H} = \sum_{pq} A_{pq} a_p^\dagger a_q + \sum_{pqrs} U_{pqrs} a_p^\dagger a_q^\dagger a_s a_r. \quad (4.1)$$

## 5 Fermionic and bosonic statistical mechanics

In this section we adopt the notation  $\zeta = +1, -1$  to indicate the bosonic and fermionic cases, respectively. Moreover, we consider Fock spaces with finitely many states  $d = |\mathcal{B}|$ . We indicate these parameters in the notation via  $\mathcal{F}_{\zeta,d}$ . Moreover, we let  $\mathcal{F}_{\zeta,d}^{(N)}$  indicate the  $N$ -particle sector of the Fock space, i.e., the  $N$ -eigenspace of the total number operator  $\hat{N}$ .

### 5.1 The zero-temperature ensemble

At zero temperature, typically one first fixes a particle number  $N$ , and attention is restricted to the  $N$ -particle subspace. Let  $|\Psi_0^{(N)}\rangle \in \mathcal{F}_{\zeta,d}^{(N)}$  be the  $N$ -particle ground state of  $\hat{H}$ , i.e., the minimal normalized eigenvector of  $\hat{H}$  considered as an operator on the  $N$ -particle subspace. The role of the density operator is assumed by the orthogonal projector  $|\Psi_0^{(N)}\rangle\langle\Psi_0^{(N)}|$  onto the ground state  $|\Psi_0^{(N)}\rangle$ , i.e., the statistical

average of a linear operator  $\hat{A}$  (with respect to the  $N$ -particle *canonical ensemble*) is given by

$$\langle \hat{A} \rangle_N = \langle \Psi_0^{(N)} | \hat{A} | \Psi_0^{(N)} \rangle.$$

## 5.2 The finite-temperature ensemble

At inverse temperature  $\beta \in (0, \infty)$ , the *partition function* is defined by

$$Z := \text{Tr}[e^{-\beta(\hat{H}-\mu\hat{N})}].$$

where ‘Tr’ indicates the Fock space trace. Here  $\mu \in \mathbb{R}$  is the *chemical potential*, but before commenting on its role, some further elaboration on the trace is owed in the bosonic case, in which the Fock space is infinite-dimensional.

By assumption,  $\hat{H}$  conserves particle number, i.e., it maps  $\mathcal{F}_{\zeta,d}^{(N)}$  to itself for all  $N$ . Thus  $e^{-\beta(\hat{H}-\mu\hat{N})}$  does as well and can be viewed as a positive-definite operator on each  $\mathcal{F}_{\zeta,d}^{(N)}$ . The trace can then be constructed as

$$\text{Tr}[e^{-\beta(\hat{H}-\mu\hat{N})}] = \sum_{N=0}^{\infty} \text{Tr}_N[e^{-\beta(\hat{H}-\mu\hat{N})}] = \sum_{N=0}^{\infty} e^{\beta\mu N} \text{Tr}_N[e^{-\beta\hat{H}}],$$

where ‘Tr<sub>N</sub>’ indicates the trace on the  $N$ -particle subspace. Since each of the summands is positive,  $\text{Tr}[e^{-\beta(\hat{H}-\mu\hat{N})}] \in (0, +\infty]$  is well-defined.

More generally, the trace is defined for all operators in the trace class of  $\mathcal{F}_{\zeta,d}$ , i.e., the set of bounded linear operators  $\hat{O} : \mathcal{F}_{\zeta,d} \rightarrow \mathcal{F}_{\zeta,d}$  for which

$$\sum_{\mathbf{n} \in \mathcal{N}_{\zeta}^d} \langle \mathbf{n} | (\hat{O}^\dagger \hat{O})^{1/2} | \mathbf{n} \rangle < +\infty,$$

in which case

$$\text{Tr}[\hat{O}] = \sum_{\mathbf{n} \in \mathcal{N}_{\zeta}^d} \langle \mathbf{n} | \hat{O} | \mathbf{n} \rangle.$$

See, e.g., [89] for further details on trace class operators.

Now since the partition function can be viewed as a normalization factor, the scenario  $Z = +\infty$  is to be avoided. It is now that we turn to the chemical potential. We can view  $Z$  as defined above as a function of  $\mu$ . Evidently  $\mu \mapsto Z(\mu)$  is non-decreasing.

First we want to rule out the case that  $Z \equiv +\infty$ . Unfortunately, this case cannot be ruled out without further assumptions! To see why, suppose that  $d = 1$  (so write  $a = a_1$ ), and let  $\hat{H} = -a^\dagger a - a^\dagger a^\dagger a a = -a^\dagger a a^\dagger a = -\hat{N}^2$ . Then

$$\text{Tr}[e^{-\beta(\hat{H}-\mu\hat{N})}] = \sum_{N=0}^{\infty} e^{\beta(N^2+\mu N)} \text{Tr}_N \left[ \text{Id}_{\mathcal{F}_{\zeta,d}^{(N)}} \right] = \sum_{N=0}^{\infty} e^{\beta(N^2+\mu N)} \binom{N+d-1}{d-1} = +\infty,$$

for all  $\mu \in \mathbb{R}$ .

We conclude that such a choice of  $\hat{H}$  is *unphysical*, and to rule out such pathologies, we adopt the following:

**Assumption 4.** *We assume, in the case of bosons, that there exist some positive integer  $N_0$  and some  $\mu \in \mathbb{R}$  such that  $\hat{H} - \mu\hat{N} \succeq 0$  as an operator on all  $N$ -particle subspaces for  $N \geq N_0$ . (It is equivalent to require that there exist  $N_0, \mu$  such that  $\hat{U} - \mu\hat{N} \succeq 0$  on all  $N$ -particle subspaces for  $N \geq N_0$ .)*

This condition is satisfied in particular if  $\hat{U}$  is a two-body interaction as in (4.1) such that  $\tilde{U}_{ik,jl} := U_{kj,il}$ , interpreted as a  $d^2 \times d^2$  matrix, is positive semidefinite. Indeed, in this case,  $\hat{U}$  is equal to (up to a single-body term)

$$\frac{1}{2} \sum_{ijkl} \tilde{U}_{ik,jl} \left[ a_i^\dagger a_k \right]^\dagger \left[ a_j^\dagger a_l \right] \succeq 0.$$

If the  $U_{ij,kl}$  are derived from a real-space two-body potential  $v$  that is a positive semidefinite function (i.e., has nonnegative Fourier transform), then it follows from Lemma 2 that the matrix  $(\tilde{U}_{ik,jl})$  is positive definite as desired. Note that it is possible for  $v$  to be positive definite but take negative values at long ranges, i.e.,  $v$  can act attractively at long range.

Now that we have argued that Assumption 4 is natural, let us see how it guarantees that the set  $\text{dom } Z := \{\mu : Z(\mu) < +\infty\}$  is non-empty. Indeed, choose  $\mu'$  negative enough such that  $\hat{H} - \mu'\hat{N} \succeq 0$  as an operator on all  $N$ -particle subspaces, and let  $\mu = \mu' - \delta$ , where  $\delta > 0$ . Then

$$\text{Tr}[e^{-\beta(\hat{H} - \mu\hat{N})}] \leq \sum_{N=0}^{\infty} \text{Tr}_N[e^{-\beta\delta\hat{N}}] = \sum_{N=0}^{\infty} e^{-\beta\delta N} \binom{N+d-1}{d-1}.$$

Now the binomial coefficient in the last expression is  $O(N^{d-1})$  as  $N \rightarrow +\infty$ , so the sum converges.

We will always assume in the finite-temperature setting that  $\mu \in \text{int dom } Z$ . It can be shown that if  $\hat{U} = 0$ , then  $\text{dom } Z = \{\mu : h \succ \mu I_d\}$ . Moreover, if there exist  $N_0, \delta > 0$  such that  $\hat{U} \succeq \delta\hat{N}^2$  on all  $N$ -particle subspaces for  $N \geq N_0$  (which holds in particular if  $\hat{U}$  is a two-body interaction as in (4.1) where the  $d^2 \times d^2$  matrix  $\tilde{U}_{ki,jl} := U_{ij,kl}$  is *positive definite*), then  $\text{dom } Z = \mathbb{R}$ .

Notice that if  $\text{dom } Z$  is open, then since  $Z$  is increasing we can write  $\text{dom } Z = (-\infty, \mu_c)$  for some  $\mu_c$  possibly infinite. If  $\mu_c < +\infty$ , then by Fatou's lemma we have that  $\liminf_{\mu \rightarrow \mu_c^-} Z(\mu) \geq Z(\mu_c) = +\infty$ , so  $Z(\mu) \rightarrow +\infty$  as  $\mu \rightarrow \mu_c^-$ . (And in any case it follows from the definition of  $Z$  that  $Z(\mu) \rightarrow +\infty$  as  $\mu \rightarrow +\infty$ , so we can write more compactly that  $Z(\mu) \rightarrow +\infty$  as  $\mu \rightarrow \mu_c$ , no matter whether  $\mu_c$  is finite or infinite.)

The *grand canonical ensemble* is defined by the *density operator*

$$\rho := Z^{-1} e^{-\beta(\hat{H} - \mu\hat{N})},$$

and the statistical average of an operator  $\hat{A}$  with respect to the grand canonical ensemble is denoted

$$\langle \hat{A} \rangle_{\beta, \mu} = \text{Tr}[\hat{A}\rho]$$

whenever  $\hat{A}\rho$  is in the trace class. Note that if  $\hat{A}$  conserves particle number then

$$\text{Tr}[\hat{A}\rho] = \sum_{N=0}^{\infty} \text{Tr}_N[\hat{A}\rho] = Z^{-1} \sum_{N=0}^{\infty} e^{\beta\mu N} \text{Tr}_N[\hat{A}e^{-\beta\hat{H}}]$$

is defined under the condition that the sum is absolutely convergent, which holds in particular if the norm of  $\hat{A}$  as an operator on the  $N$ -particle subspace grows only polynomially with  $N$ , via the assumption that  $\mu \in \text{int dom } Z$ . When the context is clear we simply write  $\langle \cdot \rangle$ .

Of particular interest is the *expected particle number*

$$\langle \hat{N} \rangle = \frac{\sum_{N=0}^{\infty} N e^{\beta\mu N} \text{Tr}_N[e^{-\beta\hat{H}}]}{\sum_{N=0}^{\infty} e^{\beta\mu N} \text{Tr}_N[e^{-\beta\hat{H}}]}.$$

Observe that  $\langle \hat{N} \rangle_{\beta, \mu} \rightarrow 0$  as  $\mu \rightarrow -\infty$ . Also note that if  $\text{dom } Z = \mathbb{R}$ , then  $\langle \hat{N} \rangle_{\beta, \mu} \rightarrow +\infty$ . Defining the *free energy*  $\Omega(\mu) := \beta^{-1} \log Z(\mu)$ , we see that  $\langle \hat{N} \rangle_{\beta, \mu} = \Omega'(\mu)$ .

It is not hard to check that  $\Omega$  is (strictly) convex, i.e.,  $\langle \hat{N} \rangle_{\beta, \mu}$  is increasing in  $\mu$  for  $\mu \in \text{int dom } Z$ . Recall that if  $\text{dom } Z = (0, \mu_c)$ , then  $Z(\mu) \rightarrow +\infty$  as  $\mu \rightarrow \mu_c$ , hence  $\Omega(\mu) \rightarrow +\infty$  as  $\mu \rightarrow \mu_c$ . If  $\mu_c < +\infty$ , it follows that  $\Omega'(\mu) \rightarrow +\infty$  as  $\mu \rightarrow \mu_c^-$ . (Otherwise, since  $\Omega'$  is increasing, it approaches a finite limit  $\mu \rightarrow \mu_c^-$ . But in this case it would follow from the fundamental theorem of calculus that  $\Omega$  approaches a finite limit as well: contradiction.) In summary we have established that if  $\text{dom } Z$  is open, then  $Z(\mu) \rightarrow +\infty$  as  $\mu \rightarrow \mu_c$ , no matter whether  $\mu_c$  is finite or infinite. It follows that in this case  $\mu \mapsto \langle \hat{N} \rangle_{\beta, \mu}$  is a bijection from  $\text{dom } Z = (-\infty, \mu_c)$  to  $(0, +\infty)$ . Thus one can *select* the chemical potential  $\mu$  to yield a predetermined expected particle number.

## 6 The coherent state path integral

There is a path integral expansion of the partition function in second quantization that is similar in spirit to the original Feynman path integral of section 3.3. It is simplest to treat the bosonic case first because the fermionic path integral formalism requires the introduction of new abstractions.

### 6.1 The bosonic coherent state path integral

We let  $\mathbf{b} = (b_p)_{p \in \mathcal{B}}$  denote the vector of annihilation operators and likewise use bold notation throughout to denote vectors indexed by the state index set  $\mathcal{B}$ . (Note: we



also retain some bold notation from section 3.3.) We say that a Hamiltonian  $\hat{H}$  is written in normal-ordered form if it is a polynomial  $\hat{H} = H(\mathbf{b}^\dagger, \mathbf{b})$  of the creation and annihilation operators such that, for each monomial term, all creation operators appear to the left of all annihilation operators, e.g.,  $\sum_{ij} A_{ij} b_i^\dagger b_j$ . Without loss of generality we shall assume that our second-quantized Hamiltonian  $\hat{H}$  is of this form. For an operator  $\hat{O}$ , viewed symbolically as a polynomial of creation and annihilation operators, we write  $:\hat{O}:$  for the normal-ordered symbolic operator obtained by formally commuting creation and annihilation operators, e.g.,  $:bb^\dagger: = b^\dagger b \neq bb^\dagger$ .

For inspiration we recapitulate the essential points of the derivation of the Feynman path integral. Recall that we considered a Hamiltonian of the form  $\hat{H} = H_{\text{kin}}(\mathbf{P}) + V(\mathbf{X})$ , where  $\mathbf{P}$  and  $\mathbf{X}$  were the momentum and position operators, respectively. Then for momentum and position eigenstates  $|\mathbf{p}\rangle$  and  $|\mathbf{x}\rangle$  we have

$$\langle \mathbf{p} | \hat{H} | \mathbf{x} \rangle = \left( \overline{H_{\text{kin}}(\mathbf{p})} + V(\mathbf{x}) \right) \langle \mathbf{p} | \mathbf{x} \rangle.$$

Using this observation, together with resolutions of the identity in terms of the momentum and position eigenstates, we derived the path integral.

Loosely following such a recipe, we are inspired to consider *eigenstates of the annihilation operators*, which will be known as the coherent states. Let  $|z\rangle$  be such an eigenstate for  $\mathbf{z} = (z_p)_{p \in \mathcal{B}} \in \mathbb{C}^{\mathcal{B}}$ , satisfying  $b_p |z\rangle = z_p |z\rangle$ . Then for normal-ordered  $\hat{H} = H(\mathbf{b}^\dagger, \mathbf{b})$ , we have

$$\langle w | \hat{H} | z \rangle = H(\bar{w}, z) \langle w | z \rangle.$$

Then if we can construct coherent states and determine a resolution of the identity in terms of them, we will be in good shape.

### 6.1.1 Bosonic coherent states

Consider the case of a bosonic system of a single state, i.e.,  $|\mathcal{B}| = 1$ . Then we want to find

$$|z\rangle = \sum_{n=0}^{\infty} \lambda_n |n\rangle$$

such that  $b|z\rangle = z|z\rangle$ , i.e.,

$$\sum_{n=0}^{\infty} z \lambda_n |n\rangle = \sum_{n=1}^{\infty} \lambda_n \sqrt{n} |n-1\rangle = \sum_{n=0}^{\infty} \lambda_{n+1} \sqrt{n+1} |n\rangle.$$

By equating corresponding terms we conclude that we must have

$$\lambda_{n+1} = \frac{z}{\sqrt{n+1}} \lambda_n,$$

and choosing  $\lambda_0 = 1$ , we obtain  $\lambda_n = \frac{z^n}{\sqrt{n!}}$ . Therefore

$$|z\rangle = \sum_{n=0}^{\infty} \frac{z^n}{\sqrt{n!}} |n\rangle = \sum_{n=0}^{\infty} \frac{z^n}{n!} (b^\dagger)^n |-\rangle,$$

or, more succinctly,

$$|z\rangle = e^{zb^\dagger} |-\rangle.$$

More generally, we may derive

$$|\mathbf{z}\rangle = e^{\mathbf{z}\cdot\mathbf{b}^\dagger} |-\rangle = \sum_{\mathbf{n}\in(\mathbb{N}_0)^\mathcal{B}} \frac{\mathbf{z}^\mathbf{n}}{\sqrt{\mathbf{n}!}} |\mathbf{n}\rangle,$$

where we interpret  $\mathbf{z}^\mathbf{n} = \prod_p z_p^{n_p}$  and  $\mathbf{n}! = \prod_p n_p!$ . This is the general formula for the bosonic coherent state, indexed by  $\mathbf{z} \in \mathbb{C}^\mathcal{B}$ .

Now for  $\mathbf{w}, \mathbf{z} \in \mathbb{C}$ ,

$$\langle \mathbf{w} | \mathbf{z} \rangle = \sum_{\mathbf{n}} \frac{\overline{\mathbf{w}}^\mathbf{n} \mathbf{z}^\mathbf{n}}{\mathbf{n}!} = \prod_{p \in \mathcal{B}} \sum_{n=0}^{\infty} \frac{(\overline{w}_p z_p)^n}{n!} = \prod_{p \in \mathcal{B}} e^{w_p z_p} = e^{\mathbf{w}^* \cdot \mathbf{z}}.$$

In particular, the coherent states are *not* orthogonal.

Nonetheless, we can write a resolution of the identity in terms of the coherent states. To wit, we have in the case  $|\mathcal{B}| = 1$  that

$$\text{Id} = \int \frac{d\bar{z} dz}{2\pi i} e^{-|z|^2} |z\rangle \langle z|, \quad (6.1)$$

as we shall confirm below. We write the integration in this particular way to lay some conceptual groundwork for the fermionic case and to clarify certain analogies. For the reader unfamiliar with such notation, we shall record presently the relevant details.

### 6.1.2 Complex coordinates

To interpret the integration we view  $\frac{1}{2i} d\bar{z} dz = \frac{1}{2i} d\bar{z} \wedge dz$  as a differential form. Writing  $z = x + iy$  for  $x, y \in \mathbb{R}$ , we see that

$$\frac{1}{2i} d\bar{z} \wedge dz = \frac{1}{2i} (dx - idy) \wedge (dx + idy) = dx \wedge dy,$$

i.e., the integration measure is the standard Lebesgue measure on  $\mathbb{R}^2$ , identified with  $\mathbb{C}$  via the decomposition into real and imaginary parts. Now for a smooth (not necessarily analytic) function  $f : \mathbb{C} \rightarrow \mathbb{C}$ ,  $f_z = \partial_z f = \frac{\partial f}{\partial z}$  and  $f_{\bar{z}} = \partial_{\bar{z}} f = \frac{\partial f}{\partial \bar{z}}$  are defined by the formula

$$df = f_z dz + f_{\bar{z}} d\bar{z}.$$

By substituting the formulas  $dz = dx + idy$  and  $d\bar{z} = dx - idy$ , one obtains the concrete identities

$$f_z = \frac{1}{2}(f_x - if_y), \quad f_{\bar{z}} = \frac{1}{2}(f_x + if_y),$$

where  $f_x$  and  $f_y$  are the standard partial derivatives. By construction

$$\int f_z d\bar{z} dz = 0, \quad \int f_{\bar{z}} d\bar{z} dz = 0 \quad (6.2)$$

for any  $f$  (with derivatives of sufficient decay, e.g., of the Schwartz class). Indeed, to see this, note, e.g., that  $f_z d\bar{z} dz = -df \wedge d\bar{z} = -d(f d\bar{z})$ , hence the claim follows by Stokes' theorem.

It is natural to consider  $f(z) = g(z, \bar{z})$ , where  $g : \mathbb{C}^2 \rightarrow \mathbb{C}$  is analytic. (In fact it is not hard to check that any real-analytic function  $\mathbb{R}^2 \rightarrow \mathbb{C}$  can be written this way.) For example, choosing  $g(z, w) = zw$  yields  $f(z) = |z|^2$ . In this case  $f_z(z) = g_z(z, \bar{z})$  and  $f_{\bar{z}}(z) = g_w(z, \bar{z})$ . Roughly speaking, we can think of  $z, \bar{z}$  algebraically as independent variables and compute the derivatives  $f_z$  and  $f_{\bar{z}}$  via the application of the usual symbolic rules to any given formula for  $f$ . In our example  $f(z) = |z|^2$ , this means that  $f_z = \bar{z}$  and  $f_{\bar{z}} = z$ . To confirm this claim, one observes (by writing difference quotients) that  $f_x(z) = g_z(z, \bar{z}) + g_w(z, \bar{z})$  and  $f_y(z) = i(g_z(z, \bar{z}) - g_w(z, \bar{z}))$ , hence  $f_z(z) = g_z(z, \bar{z})$  and  $f_{\bar{z}}(z) = g_w(z, \bar{z})$ , as desired.

### 6.1.3 The resolution of identity

Consider  $f(z) = e^{-|z|^2}$ , which can be written  $f(z) = g(z, \bar{z})$ , where  $g(z, w) = e^{-zw}$ . By inductively applying (6.2) to derivatives  $\partial_z^i \partial_{\bar{z}}^j f$ , one can show that

$$\int z^m \bar{z}^n e^{-|z|^2} d\bar{z} dz = 0, \quad \text{if } m \neq n.$$

One also has the elementary identity

$$\int e^{-|z|^2} \frac{d\bar{z} dz}{2\pi i} = \frac{1}{\pi} \int e^{-(x^2+y^2)} dx dy = 1.$$

This identity is the motivation for the normalization of the measure  $\frac{d\bar{z} dz}{2\pi i}$  and shall be directly analogized later on in the fermionic setting. Now by using (6.2) and induction once again, one derives that in turn

$$\int |z|^{2m} e^{-|z|^2} \frac{d\bar{z} dz}{2\pi i} = m!$$

Hence in summary

$$\frac{1}{m!} \int z^m \bar{z}^n e^{-|z|^2} \frac{d\bar{z} dz}{2\pi i} = \delta_{mn}. \quad (6.3)$$

Via polynomial approximation, (6.3) tells us how to integrate arbitrary functions against the measure  $e^{-|z|^2} \frac{d\bar{z} dz}{2\pi i}$ ; hence (6.3) can be thought of as an *algebraic* specification of this measure. It is this sense that can be extrapolated to the fermionic setting.

Now we have the tools needed to verify the resolution of the identity (6.1), for which it suffices to apply  $\langle -|b^m$  from the left, as follows:

$$\langle -|b^m \int \frac{dz d\bar{z}}{2\pi i} e^{-|z|^2} |z\rangle \langle z| = \int \frac{d\bar{z} dz}{2\pi i} e^{-|z|^2} z^m \langle -|z\rangle \langle z|$$

$$\begin{aligned}
&= \int \frac{d\bar{z} dz}{2\pi i} e^{-|z|^2} z^m \langle z | \\
&= \int \frac{d\bar{z} dz}{2\pi i} e^{-|z|^2} z^m \langle - | \sum_{n=0}^{\infty} \frac{\bar{z}^n}{n!} b^n \\
&= \langle - | \sum_{n=0}^{\infty} b^n \frac{1}{m!} \int z^m \bar{z}^n e^{-|z|^2} \frac{d\bar{z} dz}{2\pi i} \\
&= \langle - | b^m
\end{aligned}$$

Note the essential features of the derivation: the eigenfunction property of the coherent state  $|z\rangle$ , the normalization  $\langle -|z\rangle = 1$ , and the integration identity (6.3). The reader should keep these features in mind for the fermionic setting.

It is straightforward to likewise verify the more general resolution of identity for arbitrary  $|\mathcal{B}| \geq 1$ :

$$\text{Id} = \int \left[ \prod_{p \in \mathcal{B}} \frac{d\bar{z}_p dz_p}{2\pi i} \right] e^{-|\mathbf{z}|^2} |\mathbf{z}\rangle \langle \mathbf{z}| = \int d(\bar{\mathbf{z}}, \mathbf{z}) e^{-|\mathbf{z}|^2} |\mathbf{z}\rangle \langle \mathbf{z}|, \quad (6.4)$$

where we introduce the formal notation  $d(\bar{\mathbf{z}}, \mathbf{z}) := \prod_{p \in \mathcal{B}} \frac{d\bar{z}_p dz_p}{2\pi i}$ . The relevant integration identity is simply the product measure version of (6.3):

$$\frac{1}{\prod_p m_p!} \int \left( \prod_{p \in \mathcal{B}} z^{m_p} \bar{z}^{n_p} \right) e^{-|\mathbf{z}|^2} d(\bar{\mathbf{z}}, \mathbf{z}) = \prod_{p \in \mathcal{B}} \delta_{m_p n_p}. \quad (6.5)$$

### 6.1.4 Path integral

As suggested above, we expand the partition function (temporarily lumping the chemical potential contribution into the Hamiltonian  $\hat{H}$ ) as:

$$\begin{aligned}
Z &= \text{Tr} \left[ e^{-\beta \hat{H}} \right] \\
&= \text{Tr} \left[ e^{-\beta \hat{H}} \text{Id} \right] \\
&= \int d(\bar{\mathbf{z}}_{(0)}, \mathbf{z}_{(0)}) e^{-|\mathbf{z}_{(0)}|^2} \langle \mathbf{z}_{(0)} | e^{-\beta \hat{H}} | \mathbf{z}_{(0)} \rangle \\
&= \int d(\bar{\mathbf{z}}_{(0)}, \mathbf{z}_{(0)}) e^{-|\mathbf{z}_{(0)}|^2} \langle \mathbf{z}_{(0)} | e^{-\frac{1}{M} \beta \hat{H}} \dots e^{-\frac{1}{M} \beta \hat{H}} | \mathbf{z}_{(0)} \rangle \\
&= \int \left[ \prod_{m=0}^{M-1} d(\bar{\mathbf{z}}_{(m)}, \mathbf{z}_{(m)}) \right] e^{-\sum_{m=0}^{M-1} |\mathbf{z}_{(m)}|^2} \langle \mathbf{z}_{(0)} | e^{-\frac{1}{M} \beta \hat{H}} | \mathbf{z}_{(M-1)} \rangle \dots \langle \mathbf{z}_{(1)} | e^{-\frac{1}{M} \beta \hat{H}} | \mathbf{z}_{(0)} \rangle.
\end{aligned}$$

Now for  $M$  large, one hopes that  $e^{-\frac{1}{M} \beta \hat{H}} = :e^{-\frac{1}{M} \beta \hat{H}}: + O(M^{-2})$ , allowing us to substitute

$$\begin{aligned}
\langle \mathbf{z}_{(m+1)} | e^{-\frac{1}{M}\beta\hat{H}} | \mathbf{z}_{(m)} \rangle &\approx \langle \mathbf{z}_{(m+1)} | :e^{-\frac{1}{M}\beta\hat{H}}: | \mathbf{z}_{(m)} \rangle \\
&= e^{-\frac{1}{M}\beta H(\bar{\mathbf{z}}_{(m+1)}, \mathbf{z}_{(m)})} \langle \mathbf{z}_{(m+1)} | \mathbf{z}_{(m)} \rangle \\
&= e^{-\frac{1}{M}\beta H(\bar{\mathbf{z}}_{(m+1)}, \mathbf{z}_{(m)})} e^{\mathbf{z}_{(m+1)}^* \mathbf{z}_{(m)}}.
\end{aligned}$$

Indeed  $e^{-\frac{1}{M}\beta\hat{H}} = 1 - \frac{1}{M}\beta\hat{H} + O(M^{-2})$ , and by assumption  $:\hat{H}: = \hat{H}$ , so indeed our hope is justified (though a complete justification of the path integral would require significant further argument). Proceeding, we compute (interpreting the index  $m$  modulo  $M$ ):

$$\begin{aligned}
Z &= \lim_{M \rightarrow \infty} \int \left[ \prod_{m=0}^{M-1} d(\bar{\mathbf{z}}_{(m)}, \mathbf{z}_{(m)}) \right] e^{-\sum_{m=0}^{M-1} [|\mathbf{z}_{(m+1)}|^2 - \mathbf{z}_{(m+1)}^* \mathbf{z}_{(m)}] - \frac{\beta}{M} \sum_{m=0}^{M-1} H(\bar{\mathbf{z}}_{(m+1)}, \mathbf{z}_{(m)})} \\
&= \lim_{M \rightarrow \infty} \int \left[ \prod_{m=0}^{M-1} d(\bar{\mathbf{z}}_{(m)}, \mathbf{z}_{(m)}) \right] e^{-\sum_{m=0}^{M-1} [\mathbf{z}_{(m+1)}^* (\mathbf{z}_{(m+1)} - \mathbf{z}_{(m)})] - \frac{\beta}{M} \sum_{m=0}^{M-1} H(\bar{\mathbf{z}}_{(m+1)}, \mathbf{z}_{(m)})} \\
\text{“=”} & \int D_{\text{per}} [\bar{\mathbf{z}}(\cdot), \mathbf{z}(\cdot)] e^{-\int_0^\beta [\mathbf{z}(\tau)^* \partial_\tau \mathbf{z}(\tau) + H(\bar{\mathbf{z}}(\tau), \mathbf{z}(\tau))] d\tau}
\end{aligned}$$

where  $D_{\text{per}} [\bar{\mathbf{z}}(\cdot), \mathbf{z}(\cdot)]$  is formally the infinite-dimensional Lebesgue measure (properly normalized) on periodic paths  $\mathbf{z} : [0, \beta] \rightarrow \mathbb{C}$ . Here the “=” indicates that the expression in the last line of the display is only formal and ought to be more rigorously understood as a limit as  $M \rightarrow \infty$ . Nonetheless, the formal perspective offers significant insight!

Then by replacing  $\hat{H} \leftarrow \hat{H} - \mu \hat{N}$  and noting that  $\hat{N}(\bar{\mathbf{z}}, \mathbf{z}) = |\mathbf{z}|^2$ , we obtain the path integral formulation of the partition function

$$Z = \int D_{\text{per}} [\bar{\mathbf{z}}(\cdot), \mathbf{z}(\cdot)] e^{-S(\bar{\mathbf{z}}, \mathbf{z})},$$

where the action  $S$  is defined by

$$S(\bar{\mathbf{z}}, \mathbf{z}) := \int_0^\beta [\mathbf{z}(\tau)^* (\partial_\tau - \mu) \mathbf{z}(\tau) + H(\bar{\mathbf{z}}(\tau), \mathbf{z}(\tau))] d\tau$$

If we write  $\hat{H}$  as a sum of a noninteracting part  $\hat{H}_0$  and an interaction  $\hat{U} = U(\mathbf{b}^\dagger, \mathbf{b})$ , i.e.,

$$\hat{H} = \hat{H}_0 + \hat{U} = \sum_{p,q} h_{pq} b_p^\dagger b_q + \hat{U},$$

then we can write

$$S(\bar{\mathbf{z}}, \mathbf{z}) = S_0(\bar{\mathbf{z}}, \mathbf{z}) + S_{\text{int}}(\bar{\mathbf{z}}, \mathbf{z}),$$

where

$$S_0(\bar{\mathbf{z}}, \mathbf{z}) := \int_0^\beta \mathbf{z}(\tau)^* (\partial_\tau + h - \mu) \mathbf{z}(\tau) d\tau, \quad S_{\text{int}}(\bar{\mathbf{z}}, \mathbf{z}) = \int_0^\beta U(\bar{\mathbf{z}}(\tau), \mathbf{z}(\tau)).$$

In particular, for a general two-body interaction

$$\hat{U} = \sum_{p,q,r,s} U_{pqrs} a_p^\dagger a_q^\dagger a_s a_r, \quad (6.6)$$

we have

$$S_{\text{int}}(\bar{\mathbf{z}}, \mathbf{z}) = \sum_{pqrs} U_{pqrs} \int_0^\beta \bar{z}_p(\tau) \bar{z}_q(\tau) z_s(\tau) z_r(\tau) d\tau,$$

and for the generalized Coulomb interaction  $\hat{U} = \sum_{pq} v_{pq} \hat{n}_p \hat{n}_q$  (with  $v_{pp} = 0$ ), which corresponds to the choice  $U_{pqrs} = v_{pq} \delta_{pr} \delta_{qs}$ , we have

$$S_{\text{int}}(\bar{\mathbf{z}}, \mathbf{z}) = \sum_{pq} v_{pq} \int_0^\beta |z_p(\tau)|^2 |z_q(\tau)|^2 d\tau.$$

Observe, at this point, the formal similarity of the path integral to the Euclidean field theory presented in section 1. There is, however, a crucial difference. The contribution of the term  $\int_0^\beta \mathbf{z}(\tau)^* \partial_\tau \mathbf{z}(\tau) d\tau$  to the action includes an imaginary part, hence the path integral cannot be interpreted as a Gibbs measure, even in an infinite-dimensional sense. (This scenario should be contrasted with that of the Feynman path integral of section 3.3.)

### 6.1.5 Path integral in frequency space

Since the action in the path integral is time-translation-invariant and our paths are periodic, it makes sense to consider our paths in the frequency space. To begin we define frequency representations  $\hat{\bar{\mathbf{z}}}(\omega_n)$  and  $\hat{\mathbf{z}}(\omega_n)$  of the periodic complex paths on the Matsubara frequencies  $\omega_n = 2n\pi/\beta$  (where  $n \in \mathbb{Z}$ ):

$$\mathbf{w}(\omega_n) = \frac{1}{\sqrt{\beta}} \int_0^\beta \mathbf{z}(\tau) e^{i\omega_n \tau} d\tau,$$

so

$$\bar{\mathbf{w}}(\omega_n) = \frac{1}{\sqrt{\beta}} \int_0^\beta \bar{\mathbf{z}}(\tau) e^{-i\omega_n \tau} d\tau,$$

and we have

$$\bar{\mathbf{z}}(\tau) = \frac{1}{\sqrt{\beta}} \sum_n \bar{\mathbf{w}}(\omega_n) e^{i\omega_n \tau}, \quad \mathbf{z}(\tau) = \frac{1}{\sqrt{\beta}} \sum_n \mathbf{w}(\omega_n) e^{-i\omega_n \tau}.$$

Then we convert our action to the frequency representation by computing

$$\int_0^\beta \mathbf{z}^*(\tau) \partial_\tau \mathbf{z}(\tau) d\tau = \frac{1}{\beta} \sum_{nm} -i\omega_m \mathbf{w}^*(\omega_n) \mathbf{w}(\omega_m) \underbrace{\int_0^\beta e^{i(\omega_n - \omega_m)\tau} d\tau}_{=\beta \delta_{nm}}.$$

$$= \sum_n (-i\omega_n) \mathbf{w}^*(\omega_n) \mathbf{w}(\omega_n),$$

and

$$\int_0^\beta \mathbf{z}^*(\tau)(h - \mu)\mathbf{z}(\tau) d\tau = \sum_n \mathbf{w}^*(\omega_n)(h - \mu)\mathbf{w}(\omega_n).$$

Finally, for the two-body interaction (6.6) we compute

$$\begin{aligned} S_{\text{int}}(\bar{\mathbf{z}}, \mathbf{z}) &:= \int_0^\beta \sum_{ijkl} U_{ijkl} \bar{z}_i(\tau) \bar{z}_j(\tau) \xi_l(\tau) \xi_k(\tau) d\tau \\ &= \frac{1}{\beta^2} \sum_{ijkl} U_{ijkl} \sum_{mnpq} \bar{w}_i(\omega_m) \bar{w}_j(\omega_n) w_l(\omega_q) w_k(\omega_p) \int_0^\beta e^{i(\omega_m + \omega_n - \omega_p - \omega_q)\tau} d\tau \\ &= \frac{1}{\beta} \sum_{ijkl} \sum_{mnpq} U_{ijkl} \delta_{m+n, p+q} \bar{w}_i(\omega_m) \bar{w}_j(\omega_n) w_l(\omega_q) w_k(\omega_p) \\ &=: \hat{S}_{\text{int}}(\bar{\mathbf{w}}, \mathbf{w}). \end{aligned}$$

Since the transformations  $\mathbf{z} \mapsto \hat{\mathbf{z}}$  is a unitary change of variables, we have that

$$Z = \int \hat{D}[\bar{\mathbf{w}}(\cdot), \mathbf{w}(\cdot)] e^{-\hat{S}(\bar{\mathbf{w}}, \mathbf{w})}$$

where  $\hat{D}[\bar{\mathbf{w}}(\cdot), \mathbf{w}(\cdot)]$  is understood as the infinite-dimensional Grassmann Lebesgue measure  $\prod_{n \in \mathbb{Z}} d(\bar{\mathbf{w}}, \mathbf{w})$ , and

$$\hat{S}(\bar{\mathbf{w}}, \mathbf{w}) = \hat{S}_0(\bar{\mathbf{w}}, \mathbf{w}) + \hat{S}_{\text{int}}(\bar{\mathbf{w}}, \mathbf{w})$$

with

$$\hat{S}_0(\bar{\mathbf{w}}, \mathbf{w}) := \sum_n \mathbf{w}^*(\omega_n) (-i\omega_n + h - \mu) \mathbf{w}(\omega_n).$$

## 6.2 The fermionic coherent state path integral

When we try to mimic the derivation of the bosonic coherent states we immediately run into a difficulty. Indeed, consider the case of a single-state fermionic system, i.e.,  $|\mathcal{B}| = 1$ , and suppose that  $|\mathbf{z}\rangle$  is an eigenstate of the annihilation operators  $c_p$  with corresponding eigenvalues  $z_p$ . Then  $c_p c_q |\mathbf{z}\rangle = z_q z_p |\mathbf{z}\rangle$ , but also  $c_p c_q |\mathbf{z}\rangle = -c_q c_p |\mathbf{z}\rangle = -z_p z_q |\mathbf{z}\rangle$ , so  $z_p z_q = -z_q z_p$ . In particular, it follows that  $z_p = 0$  for all  $p$ , hence apparently any coherent state is in the null space of all of the annihilation operators. But the only state satisfying this property is the vacuum state! Clearly this won't do.

### 6.2.1 Grassmann numbers

To find our coherent states, we have to expand the space of numbers in which we look for eigenvalues. In particular, following the above reasoning, we want our eigenvalues  $\xi_i$  to satisfy  $\xi_i \xi_j = -\xi_j \xi_i$ , i.e., we want them to anticommute. This motivates the introduction of the algebra  $\mathcal{G} = \mathcal{G}(\mathcal{B})$  of Grassmann numbers (also known as supernumbers [30]), which can be identified with the exterior algebra  $\Lambda(\mathbb{C}^{\mathcal{B}})$  via the isomorphism

$$e_{p_1} \wedge \cdots \wedge e_{p_m} \mapsto \xi_{p_1} \cdots \xi_{p_m},$$

with the additional stipulation that  $\Lambda^0(\mathbb{C}^{\mathcal{B}}) \simeq \mathbb{C}$  corresponds to the complex part of a Grassmann number, sometimes referred to as the ‘body.’ More concretely, a Grassmann number  $z$  can be written uniquely as

$$z = z_B + z_S = \sum_m \sum_{p_1 < \cdots < p_m} c_{p_1 \cdots p_m} \xi_{p_1} \cdots \xi_{p_m},$$

where  $z_B$  is the complex part or body, and  $z_S$  is the rest, i.e., the ‘soul.’

### 6.2.2 Fermionic coherent states

In fact, we will always consider an extension  $\mathcal{G}^*$  of this algebra that allows us to take adjoints, and the extended algebra (itself a Grassmann algebra) will in fact be a star-algebra. Concretely, the extension is achieved by considering the Grassmann algebra generated by an enlarged set  $\{\xi_p, \bar{\xi}_p\}_{p \in \mathcal{B}}$  of anticommuting symbols; hence our algebra will be isomorphic to  $\Lambda(\mathbb{C}^{\mathcal{B} \sqcup \mathcal{B}})$ . Moreover, the adjoint ‘\*’ is defined via  $(c \xi_{p_1} \cdots \xi_{p_m})^* = \bar{c} \bar{\xi}_{p_m} \cdots \bar{\xi}_{p_1}$ . (Note that the notation for  $\mathcal{G}$  and  $\mathcal{G}_*$  is nonstandard.)

Moreover, we can consider Grassmann numbers as multipliers on Fock space operators as in the expression  $\xi_i c_i^\dagger$ . Mathematically we are considering  $\xi_i c_i^\dagger$  as an element of the extended star-algebra  $\text{End}_{\mathcal{G}^*}(\mathcal{F}) = \mathcal{G}^* \otimes_{\mathbb{C}} \text{End}_{\mathbb{C}}(\mathcal{F})$ , where multiplication is defined, for  $z \in \mathcal{G}^*$  and by  $(z_1 \otimes \hat{O}_1)(z_2 \otimes \hat{O}_2) = z_1 z_2 \otimes \hat{O}_1 \hat{O}_2$ , or, for short,  $(z_1 \hat{O}_1)(z_2 \hat{O}_2) = z_1 z_2 \hat{O}_1 \hat{O}_2$ . Moreover, the adjoint is defined by  $(z \hat{O})^\dagger = z^* \hat{O}^\dagger$ . Likewise we can extend the Fock space via the Grassmann algebra as  $\mathcal{G}^* \otimes_{\mathbb{C}} \mathcal{F}$  to consider elements such as  $\xi_p |\mathbf{n}\rangle$ , and we can extend the dual space to consider elements such as  $\langle \mathbf{n} | \bar{\xi}_p$ .

For the most part, such technicalities need not be emphasized. Nonetheless, we have given some indication of the mathematical structures in order to reassure the readers that the Grassmann numbers and all accompanying manipulations can in fact be backed by honest mathematical definition.

Finally, by analogy with the bosonic case, we define the fermionic coherent state

$$|\boldsymbol{\xi}\rangle = e^{\boldsymbol{\xi} \cdot \mathbf{c}^\dagger} |-\rangle.$$

It is important to note that unlike complex numbers, the Grassmann generators  $\xi_i$  should not be thought of as variables with indeterminate numerical value; rather they



are mere symbols constrained to satisfy certain algebraic relations. In the development of the theory of bosonic coherent states, we attempted to emphasize the role of  $z_i, \bar{z}_i$  as mere symbols and of integration as a recipe for assigning numerical values to algebraic expressions in these symbols. To transfer our developments to the fermionic setting, we will likewise only need a recipe for ‘integrating’ (or assigning numerical value to) elements of the Grassmann algebra.

Now observe that the via the anticommutation of both the Grassmann generators and the fermionic creation operators we have that that  $\xi_i c_i^\dagger$  all commute within  $\text{End}_{\mathcal{G}^*}(\mathcal{F})$ , much like the analogous bosonic operators  $z_i b_i^\dagger$ , where  $z_i \in \mathbb{C}$ . Hence

$$e^{\boldsymbol{\xi} \cdot \mathbf{c}^\dagger} = e^{\sum_p \xi_p c_p^\dagger} = \prod_p e^{\xi_p c_p^\dagger},$$

where the order in the product of the latter expression can be arbitrary. Now  $e^{\xi_p c_p^\dagger} = 1 + \xi_p c_p^\dagger$ ; note that the Taylor series series terminates abruptly because  $\xi_p^2 = 0$ .

Now let us verify that  $|\boldsymbol{\xi}\rangle \in \mathcal{G}^* \otimes_{\mathbb{C}} \mathcal{F}$  are eigenfunctions for the annihilation operators  $c_p$  with eigenvalues  $\xi_p \in \mathcal{G}^*$ :

$$\begin{aligned} c_p |\boldsymbol{\xi}\rangle &= c_p (1 + \xi_p c_p^\dagger) \prod_{q \neq p} e^{\xi_q c_q^\dagger} |-\rangle \\ &= (c_p + \xi_p c_p c_p^\dagger) \prod_{q \neq p} e^{\xi_q c_q^\dagger} |-\rangle \\ &= (c_p + \xi_p (1 - c_p^\dagger c_p)) \prod_{q \neq p} e^{\xi_q c_q^\dagger} |-\rangle \\ &= \xi_p \prod_{q \neq p} e^{\xi_q c_q^\dagger} |-\rangle + (1 - c_p^\dagger) c_p \prod_{q \neq p} e^{\xi_q c_q^\dagger} |-\rangle. \end{aligned}$$

Now notice that  $\prod_{q \neq p} e^{\xi_q c_q^\dagger} |-\rangle$  is in the zero-eigenspace of  $\hat{n}_p$ , hence  $c_p \prod_{q \neq p} e^{\xi_q c_q^\dagger} |-\rangle = 0$ , and the second term in the last display is zero. Moreover  $\xi_p = \xi_p (1 + \xi_p c_p^\dagger)$ , so we have derived

$$c_p |\boldsymbol{\xi}\rangle = \xi_p |\boldsymbol{\xi}\rangle,$$

as desired. Note that the adjoint coherent state is given by

$$\langle \boldsymbol{\xi} | = |\boldsymbol{\xi}\rangle^* = \langle - | e^{\bar{\boldsymbol{\xi}} \cdot \mathbf{c}_p}.$$

### 6.2.3 Grassmann integration

Now in order to formulate a resolution of identity, we need an integration formula.<sup>2</sup> First let us focus on the case of  $|\mathcal{B}| = 1$ , i.e., the case of  $\mathcal{G}^* = \langle \bar{\xi}, \xi \rangle$ , where we use angle

---

<sup>2</sup>We will only consider integration on algebras with adjoint symbols in order to emphasize the analogy with the bosonic case, but in fact Grassmann integration can also be defined without difficulty on any Grassmann algebra. However, the evenness that accompanies the adjoint structure makes some aspects of the theory more elegant because even elements of the Grassmann algebra are commute with all elements of the algebra.

brackets to indicate the Grassmann algebra generated by the anticommuting symbols therein contained. The motivation here will be to analogize (6.5) from the bosonic setting. Since  $\xi^m = \bar{\xi}^n = 0$  for  $m, n \geq 2$ , it is sufficient to *define* an integration functional  $\mathcal{I} : \mathcal{G}^* \rightarrow \mathbb{C}$  via

$$\delta_{mn} = \mathcal{I} \left[ \xi^m \bar{\xi}^n e^{-\bar{\xi}\xi} \right] =: \int \xi^m \bar{\xi}^n e^{-\bar{\xi}\xi} d(\bar{\xi}, \xi)$$

for  $m, n \in \{0, 1\}$ . Note that

$$\begin{aligned} e^{-\bar{\xi}\xi} &= 1 - \bar{\xi}\xi = 1 + \xi\bar{\xi} \\ \xi e^{-\bar{\xi}\xi} &= \xi(1 - \bar{\xi}\xi) = \xi \\ \bar{\xi} e^{-\bar{\xi}\xi} &= \bar{\xi}(1 - \bar{\xi}\xi) = \bar{\xi} \\ \xi\bar{\xi} e^{-\bar{\xi}\xi} &= \xi\bar{\xi}(1 - \bar{\xi}\xi) = \xi\bar{\xi}, \end{aligned}$$

so it follows that one could otherwise define  $\mathcal{I}$  via

$$\mathcal{I}[1] = \mathcal{I}[\xi] = \mathcal{I}[\bar{\xi}] = 0, \quad \mathcal{I}[\xi\bar{\xi}] = 1.$$

To define an integration on a more general algebra  $\mathcal{G}^*(\mathcal{B}) = \langle \{\bar{\xi}_p, \xi_p\}_{p \in \mathcal{B}} \rangle$ , we seek to analogize the bosonic integration formula via

$$\int \left( \prod_{p \in \mathcal{B}} \xi^{m_p} \bar{\xi}^{n_p} \right) e^{-\xi^* \xi} d(\bar{\xi}, \xi) := \mathcal{I} \left[ \left( \prod_{p \in \mathcal{B}} \xi^{m_p} \bar{\xi}^{n_p} \right) e^{-\xi^* \xi} \right] := \prod_{p \in \mathcal{B}} \delta_{m_p n_p} \quad (6.7)$$

for  $m_p, n_p \in \{0, 1\}$ . Sometimes we will alternatively write  $d(\bar{\xi}, \xi) = \prod_{p \in \mathcal{B}} d(\bar{\xi}_p, \xi_p)$  to denote the multivariate Grassmann integration ‘measure.’ We will sometimes write  $d(\bar{\xi}, \xi)$  immediately after the integration sign  $\int$ , as in  $\int d(\bar{\xi}, \xi) \cdots$ , but the meaning is unchanged. Note carefully that the definition makes sense regardless of the ordering of  $p \in \mathcal{B}$  in the product because whenever  $m_p + n_p$  is odd for some  $p$ , the result of the integration is defined to be zero; meanwhile, whenever  $m_p + n_p$  is even,  $\xi^{m_p} \bar{\xi}^{n_p}$  commutes with all elements of  $\mathcal{G}^*$ . One verifies that our definition of  $\mathcal{I} : \mathcal{G}^*(\mathcal{B}) \rightarrow \mathbb{C}$  is equivalent to the definition

$$\mathcal{I} \left[ \prod_{p \in \mathcal{B}} \xi_p \bar{\xi}_p \right] = 1, \quad \mathcal{I} [\text{any other monomial}] = 0.$$

We may also consider partial integration  $\mathcal{I}_{\mathcal{S}} : \mathcal{G}^*(\mathcal{B}) \rightarrow \mathcal{G}^*(\mathcal{B} \setminus \mathcal{S})$ , defined by

$$\mathcal{I}_{\mathcal{S}} [f((\bar{\xi}_p, \xi_p)_{p \in \mathcal{S}}) g((\bar{\xi}_p, \xi_p)_{p \notin \mathcal{S}})] = \mathcal{I} [f((\bar{\xi}_p, \xi_p)_{p \in \mathcal{S}})] g((\bar{\xi}_p, \xi_p)_{p \notin \mathcal{S}})$$

for polynomials  $f, g$ . The left-hand side may alternatively be denoted by

$$\int f((\bar{\xi}_p, \xi_p)_{p \in \mathcal{S}}) g((\bar{\xi}_p, \xi_p)_{p \notin \mathcal{S}}) \left[ \prod_{p \in \mathcal{S}} d(\bar{\xi}_p, \xi_p) \right].$$

There is in fact another perspective on the definition of the Grassmann integration; we may view it as an attempt to analogize (6.2). To pursue such an analogy, we need to define a suitable notion of Grassmann differentiation, i.e., linear operators  $\partial_{\xi_p}, \partial_{\bar{\xi}_p} : \mathcal{G}^* \rightarrow \mathcal{G}^*$ . These operators are determined entirely by the formulas

$$\begin{aligned} \partial_{\xi_p} [\xi_p f(\bar{\boldsymbol{\xi}}, \{\xi_p\}_{p \neq q})] &= f(\bar{\boldsymbol{\xi}}, \{\xi_p\}_{p \neq q}), & \partial_{\xi_p} [f(\bar{\boldsymbol{\xi}}, \{\xi_p\}_{p \neq q})] &= 0 \\ \partial_{\bar{\xi}_p} [\bar{\xi}_p f(\{\bar{\xi}_p\}_{p \neq q}, \boldsymbol{\xi})] &= f(\{\bar{\xi}_p\}_{p \neq q}, \boldsymbol{\xi}), & \partial_{\bar{\xi}_p} [f(\{\bar{\xi}_p\}_{p \neq q}, \boldsymbol{\xi})] &= 0. \end{aligned}$$

Note carefully, e.g., that  $\partial_{\xi_1}(\xi_2 \xi_1) = -\partial_{\xi_1}(\xi_1 \xi_2) = -\xi_2$ .

Then we can think of the stipulation that  $\mathcal{I} \left[ \left( \prod_{p \in \mathcal{B}} \xi_p \bar{\xi}_p \right) e^{-\boldsymbol{\xi}^* \boldsymbol{\xi}} \right] = 1$ , or equivalently that  $\mathcal{I} \left[ \prod_{p \in \mathcal{B}} \bar{\xi}_p \xi_p \right] = 1$ , as a kind of arbitrary ‘normalization’ of the Grassmann measure, just as in the bosonic case. Meanwhile, integration of arbitrary polynomials can then be defined via the stipulation, analogous to (6.2), that

$$\int \partial_{\xi_p} f d(\bar{\boldsymbol{\xi}}, \boldsymbol{\xi}) = \int \partial_{\bar{\xi}_p} f d(\bar{\boldsymbol{\xi}}, \boldsymbol{\xi}) = 0$$

for all polynomials  $f$ . Since every monomial in  $\mathcal{G}^*(\mathcal{B})$  besides  $\prod_{p \in \mathcal{B}} \xi_p \bar{\xi}_p$  can be written as a derivative, the integration rule introduced above follows.

#### 6.2.4 The resolution of identity

Due to the eigenfunction property of the coherent state  $|\boldsymbol{\xi}\rangle$ , the normalization  $\langle -|\boldsymbol{\xi}\rangle = 1$ , and the integration identity (6.7) (analogous to (6.5)), the proof of our resolution of identity will be analogous to the proof in the bosonic case. The resolution of identity is written

$$\text{Id}_{\mathcal{F}} = \int d(\bar{\boldsymbol{\xi}}, \boldsymbol{\xi}) e^{-\boldsymbol{\xi}^* \boldsymbol{\xi}} |\boldsymbol{\xi}\rangle \langle \boldsymbol{\xi}|, \quad (6.8)$$

and we prove by applying an arbitrary occupation number basis element  $\langle -|c_{p_m} \cdots c_{p_1}$  from the left, as

$$\begin{aligned} \langle -|c_{p_m} \cdots c_{p_1} \int d(\bar{\boldsymbol{\xi}}, \boldsymbol{\xi}) e^{-\boldsymbol{\xi}^* \boldsymbol{\xi}} |\boldsymbol{\xi}\rangle \langle \boldsymbol{\xi}| &= \int d(\bar{\boldsymbol{\xi}}, \boldsymbol{\xi}) e^{-\boldsymbol{\xi}^* \boldsymbol{\xi}} \xi_{p_1} \cdots \xi_{p_m} \langle -|\boldsymbol{\xi}\rangle \langle \boldsymbol{\xi}| \\ &= \int d(\bar{\boldsymbol{\xi}}, \boldsymbol{\xi}) e^{-\boldsymbol{\xi}^* \boldsymbol{\xi}} \xi_{p_1} \cdots \xi_{p_m} \langle -| \prod_{p \in \mathcal{B}} (1 + \bar{\xi}_p c_p) \\ &= \langle -| \int \xi_{p_1} \cdots \xi_{p_m} \prod_{p \in \mathcal{B}} (1 + \bar{\xi}_p c_p) e^{-\boldsymbol{\xi}^* \boldsymbol{\xi}} d(\bar{\boldsymbol{\xi}}, \boldsymbol{\xi}), \end{aligned}$$

where we have used the fact that  $e^{-\xi^* \xi}$  commutes with all of  $\mathcal{G}^*$ . Then note that upon expanding the product, by (6.7) the only term that survives is  $\bar{\xi}_{p_m} \cdots \bar{\xi}_{p_1} c_{p_m} \cdots c_{p_1}$ , hence

$$\begin{aligned} & \langle -|c_{p_m} \cdots c_{p_1} \int d(\bar{\xi}, \xi) e^{-\xi^* \xi} |\xi\rangle \langle \xi| \\ &= \langle -|c_{p_m} \cdots c_{p_1} \left[ \int \xi_{p_1} \cdots \xi_{p_m} \bar{\xi}_{p_m} \cdots \bar{\xi}_{p_1} e^{-\xi^* \xi} d(\bar{\xi}, \xi) \right] \\ &= \langle -|c_{p_m} \cdots c_{p_1}, \end{aligned}$$

as was to be shown. In the last step we used (6.7), together with the fact that

$$\xi_{p_1} \cdots \xi_{p_m} \bar{\xi}_{p_m} \cdots \bar{\xi}_{p_1} = (\xi_{p_1} \bar{\xi}_{p_1}) \cdots (\xi_{p_m} \bar{\xi}_{p_m}),$$

which follows from grouping the factor  $\xi_{p_m} \bar{\xi}_{p_m}$ , which commutes with the entire algebra, and moving it all the way to the right, then repeating for  $\xi_{p_{m-1}} \bar{\xi}_{p_{m-1}}$ , etc.

In order to use the resolution of identity to compute traces, it is useful to derive the following identity:

$$\langle \mathbf{m} | \xi \rangle \langle \xi | \mathbf{n} \rangle = \langle -\xi | \mathbf{n} \rangle \langle \mathbf{m} | \xi \rangle, \quad \text{when } \sum_p (m_p - n_p) \equiv 0 \pmod{2}. \quad (6.9)$$

Here we interpret  $|-\xi\rangle = e^{(-\xi) \cdot c^\dagger} |-\rangle$ , and  $\langle -\xi| = |-\xi\rangle^* = \langle -| e^{(-\bar{\xi}) \cdot c}$ . Note that  $c_p |-\xi\rangle = -\xi_p |-\xi\rangle$  for all  $p$ .

Now to prove the claim, write  $|\mathbf{m}\rangle = c_{p_1}^\dagger \cdots c_{p_M}^\dagger |-\rangle$  and  $|\mathbf{n}\rangle = c_{q_1}^\dagger \cdots c_{q_N}^\dagger |-\rangle$ , where  $M - N$  is even, and compute

$$\begin{aligned} \langle \mathbf{m} | \xi \rangle &= \langle -|c_{p_M} \cdots c_{p_1} | \xi \rangle = \xi_{p_1} \cdots \xi_{p_M} \langle -| \xi \rangle = \xi_{p_1} \cdots \xi_{p_M} \\ \langle \xi | \mathbf{n} \rangle &= \langle \xi | c_{q_1}^\dagger \cdots c_{q_N}^\dagger |-\rangle = \bar{\xi}_{q_N} \cdots \bar{\xi}_{q_1} \\ \langle -\xi | \mathbf{n} \rangle &= \langle -\xi | c_{q_1}^\dagger \cdots c_{q_N}^\dagger |-\rangle = (-\bar{\xi}_{q_N}) \cdots (-\bar{\xi}_{q_1}) = (-1)^N \bar{\xi}_{p_1} \cdots \bar{\xi}_{p_m}. \end{aligned}$$

Then

$$\begin{aligned} \langle \mathbf{m} | \xi \rangle \langle \xi | \mathbf{n} \rangle &= \xi_{p_1} \cdots \xi_{p_M} \bar{\xi}_{q_N} \cdots \bar{\xi}_{q_1} \\ &= (-1)^{MN} \bar{\xi}_{q_N} \cdots \bar{\xi}_{q_1} \xi_{p_1} \cdots \xi_{p_M} \\ &= (-1)^{MN} (-1)^N \langle -\xi | \mathbf{n} \rangle \langle \mathbf{m} | \xi \rangle. \end{aligned}$$

But

$$(-1)^{MN} = (-1)^{(N+M-N)N} = (-1)^{N^2} (-1)^{(M-N)N} = (-1)^N,$$

where we have used the facts that  $(-1)^{N^2} = (-1)^N$  and that  $M - N$  is even (so  $(M - N)N$  is even as well). The claim (6.9) follows.

From the identity (6.9), together with our resolution of identity (6.7), we may derive a formula for  $\text{Tr}(\hat{O})$  for operators  $\hat{O} = O(\mathbf{c}^\dagger, \mathbf{c})$ , where  $O$  is an *even* polynomial (as is required of physical fermionic operators):

$$\text{Tr}(\hat{O}) = \int d(\bar{\boldsymbol{\xi}}, \boldsymbol{\xi}) e^{-\boldsymbol{\xi}^* \boldsymbol{\xi}} \langle -\boldsymbol{\xi} | \hat{O} | \boldsymbol{\xi} \rangle. \quad (6.10)$$

To derive the identity, we expand as

$$\begin{aligned} \text{Tr}(\hat{O}) &= \text{Tr}(\hat{O} \text{Id}_{\mathcal{F}}) \\ &= \int d(\bar{\boldsymbol{\xi}}, \boldsymbol{\xi}) e^{-\boldsymbol{\xi}^* \boldsymbol{\xi}} \text{Tr}(\hat{O} |\boldsymbol{\xi}\rangle \langle \boldsymbol{\xi}|) \\ &= \int d(\bar{\boldsymbol{\xi}}, \boldsymbol{\xi}) e^{-\boldsymbol{\xi}^* \boldsymbol{\xi}} \sum_{\mathbf{n}} \langle \mathbf{n} | \hat{O} | \boldsymbol{\xi} \rangle \langle \boldsymbol{\xi} | \mathbf{n} \rangle. \end{aligned}$$

Now write  $\hat{O} = \sum_{\mathbf{m}', \mathbf{m}} O_{\mathbf{m}' \mathbf{m}} |\mathbf{m}'\rangle \langle \mathbf{m}|$ , where  $O_{\mathbf{m}' \mathbf{m}} = 0$  whenever  $\sum_p (m'_p - m_p)$  is odd. Then inserting this expression we obtain

$$\begin{aligned} \text{Tr}(\hat{O}) &= \int d(\bar{\boldsymbol{\xi}}, \boldsymbol{\xi}) e^{-\boldsymbol{\xi}^* \boldsymbol{\xi}} \sum_{\mathbf{n}, \mathbf{m}} O_{\mathbf{n} \mathbf{m}} \langle \mathbf{m} | \boldsymbol{\xi} \rangle \langle \boldsymbol{\xi} | \mathbf{n} \rangle \\ &= \int d(\bar{\boldsymbol{\xi}}, \boldsymbol{\xi}) e^{-\boldsymbol{\xi}^* \boldsymbol{\xi}} \sum_{\mathbf{n}, \mathbf{m}} O_{\mathbf{n} \mathbf{m}} \langle -\boldsymbol{\xi} | \mathbf{n} \rangle \langle \mathbf{m} | \boldsymbol{\xi} \rangle \\ &= \int d(\bar{\boldsymbol{\xi}}, \boldsymbol{\xi}) e^{-\boldsymbol{\xi}^* \boldsymbol{\xi}} \langle -\boldsymbol{\xi} | \hat{O} | \boldsymbol{\xi} \rangle, \end{aligned}$$

as was to be shown. (In the second equality of the last display, we used (6.9), together with the fact that  $O_{\mathbf{n} \mathbf{m}} = 0$  whenever  $\sum_p (n_p - m_p)$  is odd.)

### 6.2.5 Path integral

Expand the partition function (again temporarily lumping the chemical potential contribution into the Hamiltonian  $\hat{H}$ ) via the trace identity (6.10):

$$\begin{aligned} Z &= \text{Tr} \left[ e^{-\beta \hat{H}} \right] \\ &= \int d(\bar{\boldsymbol{\xi}}_{(0)}, \boldsymbol{\xi}_{(0)}) e^{-\boldsymbol{\xi}_{(0)}^* \boldsymbol{\xi}_{(0)}} \langle -\boldsymbol{\xi}_{(0)} | e^{-\beta \hat{H}} | \boldsymbol{\xi}_{(0)} \rangle \\ &= \int d(\bar{\boldsymbol{\xi}}_{(0)}, \boldsymbol{\xi}_{(0)}) e^{-\boldsymbol{\xi}_{(0)}^* \boldsymbol{\xi}_{(0)}} \langle -\boldsymbol{\xi}_{(0)} | e^{-\frac{1}{M} \beta \hat{H}} \dots e^{-\frac{1}{M} \beta \hat{H}} | \boldsymbol{\xi}_{(0)} \rangle \\ &= \int \left[ \prod_{m=0}^{M-1} d(\bar{\boldsymbol{\xi}}_{(m)}, \boldsymbol{\xi}_{(m)}) \right] e^{-\sum_{m=0}^{M-1} \boldsymbol{\xi}_{(m)}^* \boldsymbol{\xi}_{(m)}} \\ &\quad \langle -\boldsymbol{\xi}_{(0)} | e^{-\frac{1}{M} \beta \hat{H}} | \boldsymbol{\xi}_{(M-1)} \rangle \dots \langle \boldsymbol{\xi}_{(1)} | e^{-\frac{1}{M} \beta \hat{H}} | \boldsymbol{\xi}_{(0)} \rangle. \end{aligned}$$

Observe that in the last line, the integration takes place in the enlarged Grassmann algebra

$$\mathcal{G}_M^* := \langle \{\bar{\xi}_{(m),p}, \xi_{(m),p}\}_{p \in \mathcal{B}, m=0, \dots, M-1} \rangle.$$

Evidently, in this enlarged Grassmann algebra we shall have to compute the overlaps  $\langle \boldsymbol{\xi}_{(m)} | \boldsymbol{\xi}_{(m-1)} \rangle$ .

More generally, we compute the overlap  $\langle \boldsymbol{\theta} | \boldsymbol{\xi} \rangle$  within  $\langle \{\bar{\xi}_{,p}, \xi_{,p}, \bar{\theta}_p, \theta_p\}_{p \in \mathcal{B}} \rangle$ :

$$\langle \boldsymbol{\theta} | \boldsymbol{\xi} \rangle = e^{\boldsymbol{\theta}^* \boldsymbol{\xi}}, \quad (6.11)$$

analogously to the bosonic case. To verify this identity, first rewrite

$$|\boldsymbol{\xi}\rangle = \prod_p (1 + \xi_p c_p^\dagger) |-\rangle = \sum_{S \subset \mathcal{B}} \prod_{p \in S} (\xi_p c_p^\dagger) |-\rangle.$$

Note that the ordering of  $p$  within the product does not matter. Similarly,

$$\langle \boldsymbol{\theta} | = \sum_{S \subset \mathcal{B}} \langle - | \prod_{p \in S} (\bar{\theta}_p c_p),$$

from which it follows that

$$\begin{aligned} \langle \boldsymbol{\theta} | \boldsymbol{\xi} \rangle &= \sum_{S \subset \mathcal{B}} \prod_{p \in S} (\bar{\theta}_p \xi_p) \\ &= \prod_{p \in \mathcal{B}} (1 + \bar{\theta}_p \xi_p) \\ &= \prod_{p \in \mathcal{B}} e^{\bar{\theta}_p \xi_p} \\ &= e^{\boldsymbol{\theta}^* \boldsymbol{\xi}}, \end{aligned}$$

as was to be shown.

Now for  $M$  large, we again make use of  $e^{-\frac{1}{M}\beta\hat{H}} = :e^{-\frac{1}{M}\beta\hat{H}}: + O(M^{-2})$ , allowing us to substitute

$$\begin{aligned} \langle \boldsymbol{\xi}_{(m+1)} | e^{-\frac{1}{M}\beta\hat{H}} | \boldsymbol{\xi}_{(m)} \rangle &\approx \langle \boldsymbol{\xi}_{(m+1)} | :e^{-\frac{1}{M}\beta\hat{H}}: | \boldsymbol{\xi}_{(m)} \rangle \\ &= e^{-\frac{1}{M}\beta H(\bar{\boldsymbol{\xi}}_{(m+1)}, \boldsymbol{\xi}_{(m)})} \langle \boldsymbol{\xi}_{(m+1)} | \boldsymbol{\xi}_{(m)} \rangle \\ &= e^{-\frac{1}{M}\beta H(\bar{\boldsymbol{\xi}}_{(m+1)}, \boldsymbol{\xi}_{(m)})} e^{\boldsymbol{\xi}_{(m+1)}^* \boldsymbol{\xi}_{(m)}}. \end{aligned}$$

Proceeding, we compute, adopting the convention  $\boldsymbol{\xi}_{(M)} = -\boldsymbol{\xi}_{(0)}$ :

$$Z = \lim_{M \rightarrow \infty} \int \left[ \prod_{m=0}^{M-1} d(\bar{\boldsymbol{\xi}}_{(m)}, \boldsymbol{\xi}_{(m)}) \right] e^{-\sum_{m=0}^{M-1} [\boldsymbol{\xi}_{(m+1)}^* (\boldsymbol{\xi}_{(m+1)} - \boldsymbol{\xi}_{(m)})] - \frac{\beta}{M} \sum_{m=0}^{M-1} H(\bar{\boldsymbol{\xi}}_{(m+1)}, \boldsymbol{\xi}_{(m)})}$$

$$= \int D_{\text{a-per}} [\bar{\boldsymbol{\xi}}(\cdot), \boldsymbol{\xi}(\cdot)] e^{-\int_0^\beta [\boldsymbol{\xi}(\tau)^* \partial_\tau \boldsymbol{\xi}(\tau) + H(\bar{\boldsymbol{\xi}}(\tau), \boldsymbol{\xi}(\tau))] d\tau}$$

where  $D_{\text{a-per}} [\bar{\boldsymbol{\xi}}(\cdot), \boldsymbol{\xi}(\cdot)]$  is formally the infinite-dimensional Lebesgue measure (properly normalized) on *antiperiodic* ‘Grassmann paths’  $\boldsymbol{\xi}(\tau)$ . Again “=” indicates that the expression in the last line of the display is only formal and ought to be more rigorously understood as a limit as  $M \rightarrow \infty$ . The notion even of a Grassmann path is shakily interpreted at best; by contrast to a complex path, it cannot be interpreted as an (anti)periodic function on  $[0, \beta)$ . Rather, its meaning is only symbolic.

Then by replacing  $\hat{H} \leftarrow \hat{H} - \mu \hat{N}$  and noting that  $\hat{N}(\bar{\boldsymbol{\xi}}, \boldsymbol{\xi}) = \boldsymbol{\xi}^* \boldsymbol{\xi}$ , we obtain the path integral formulation of the partition function

$$Z = \int D_{\text{a-per}} [\bar{\boldsymbol{\xi}}(\cdot), \boldsymbol{\xi}(\cdot)] e^{-S(\bar{\boldsymbol{\xi}}, \boldsymbol{\xi})},$$

where the action  $S$  is defined by

$$S(\bar{\boldsymbol{\xi}}, \boldsymbol{\xi}) := \int_0^\beta [\boldsymbol{\xi}(\tau)^* (\partial_\tau - \mu) \boldsymbol{\xi}(\tau) + H(\bar{\boldsymbol{\xi}}(\tau), \boldsymbol{\xi}(\tau))] d\tau$$

If we write  $\hat{H}$  as a sum of a noninteracting part  $\hat{H}_0$  and an interaction  $\hat{U} = U(\mathbf{c}^\dagger, \mathbf{c})$ , i.e.,

$$\hat{H} = \hat{H}_0 + \hat{U} = \sum_{p,q} h_{pq} c_p^\dagger c_q + \hat{U},$$

then we can write

$$S(\bar{\boldsymbol{\xi}}, \boldsymbol{\xi}) = S_0(\bar{\boldsymbol{\xi}}, \boldsymbol{\xi}) + S_{\text{int}}(\bar{\boldsymbol{\xi}}, \boldsymbol{\xi}),$$

where

$$S_0(\bar{\boldsymbol{\xi}}, \boldsymbol{\xi}) := \int_0^\beta \boldsymbol{\xi}(\tau)^* (\partial_\tau + h - \mu) \boldsymbol{\xi}(\tau) d\tau, \quad S_{\text{int}}(\bar{\boldsymbol{\xi}}, \boldsymbol{\xi}) = \int_0^\beta U(\bar{\boldsymbol{\xi}}(\tau), \boldsymbol{\xi}(\tau)).$$

Again one can observe the formal similarity of the path integral to the Euclidean field theory presented in section 1. However, the analogy is even more restricted here for obvious reasons.

This concludes our discussion of the coherent-state path integral. The use of this construction in this dissertation is limited to section 7.1 below, where we use it to motivate the connection between Green’s functions in the Euclidean lattice field theory (which will be key in Parts II, III, IV, and VI) and Green’s functions in fermionic and bosonic statistical mechanics (which will be key in Parts VI and VII).

## 7 Green’s functions

### 7.1 Motivation via functional derivatives

Before we proceed with standard definitions to many-body Green’s functions, we first offer some motivating discussion from a more general perspective. In this section we will consider  $d = |\mathcal{B}| < \infty$ .

In the setting of Euclidean field theory, our notion of the ‘Green’s function’ associated to a Gibbs measure  $d\mu(x) = e^{-H(x)} dx$  is simply the two-point correlator

$$G = \int_{\mathbb{R}^N} xx^\top d\mu(x).$$

Note that for the choice  $H_A(x)$  of (1.1), by defining the free energy

$$\Omega[A] := -\log Z[A] = \int_{\mathbb{R}^N} e^{-\frac{1}{2}x^\top Ax - U(x)} dx,$$

we can write  $G = G[A]$  as a gradient via  $G[A] = -\nabla_A \Omega[A]$ , where we define  $\nabla_A := \left( \frac{\partial}{\partial A_{ij}} + \frac{\partial}{\partial A_{ji}} \right)$ . We comment that the self-energy is defined as the difference  $\Sigma = A - G^{-1}$  so that  $\Sigma = 0$  if  $U \equiv 0$  (i.e., in the noninteracting case). Further detail is provided in Part III, which views this relation as the foundation of the so-called Luttinger-Ward formalism.

For now, let us analogize this construction to the setting of the coherent state path integral. Note that the discussion will be only informal, with rigorous definitions to follow later. For concreteness, we will stick to the fermionic case. There, the ‘quadratic part’ of the action (i.e., the analogy of  $\frac{1}{2}x^\top Ax$  in the Euclidean setting) is

$$S_0(\bar{\xi}, \xi) = \int_0^\beta \xi(\tau)^* (\partial_\tau + h - \mu) \xi(\tau) d\tau.$$

We can extend this particular action to the broadest possible parametric class of quadratic actions as

$$S_0[A](\bar{\xi}, \xi) := \int_0^\beta \xi(\tau)^* \partial_\tau \xi(\tau) d\tau + \int_0^\beta \int_0^\beta \xi(\tau')^* A(\tau', \tau) \xi(\tau) d\tau d\tau',$$

so that the action map  $(\bar{\xi}, \xi) \mapsto S_0[A](\bar{\xi}, \xi)$  is itself a functional of the Hermitian-operator-valued kernel  $(\tau', \tau) \mapsto A(\tau', \tau)$ .

Then by considering the partition function  $Z = Z[A]$  as a functional of the kernel  $A$  and defining  $\Omega[A] = \log Z[A]$ , we may in turn define a Green’s function via

$$\begin{aligned} G(\tau, \tau')[A] &:= \frac{\delta \Omega}{\delta A(\tau', \tau)}[A] \\ &= \frac{-1}{Z[A]} \int D_{\text{a-per}} [\bar{\xi}(\cdot), \xi(\cdot)] \xi(\tau) \xi(\tau')^* e^{-S_0[A](\bar{\xi}, \xi) - S_{\text{int}}(\bar{\xi}, \xi)}. \end{aligned}$$

By evaluating at a kernel of the form  $A(\tau', \tau) = (h - \mu)\delta(\tau' - \tau)$  and *reversing the steps* of the derivation of the path integral, we find that

$$G_{ij}(\tau, \tau') = \frac{-1}{Z} \text{Tr} \left[ \mathcal{T} \left\{ a_i(\tau) a_j^\dagger(\tau') \right\} e^{-\beta(\hat{H} - \mu \hat{N})} \right],$$



where  $\mathcal{T} \{a_i(\tau)a_j^\dagger(\tau')\}$  indicates the imaginary-time-ordering operator, formally defined by

$$\mathcal{T} \{a_i(\tau)a_j^\dagger(\tau')\} = \begin{cases} a_i(\tau)a_j^\dagger(\tau'), & \tau' < \tau \\ -a_j^\dagger(\tau')a_i(\tau), & \tau' \geq \tau. \end{cases}$$

In fact this matches the definition of the *Matsubara Green's function* to be given below.

## 7.2 Green's functions and the self-energy at zero temperature

For  $t \in \mathbb{R}$ , we denote the annihilation and creation operators in the Heisenberg representation by

$$a_i(t) := e^{i\hat{H}t}a_i e^{-i\hat{H}t}, \quad a_i^\dagger(t) := e^{i\hat{H}t}a_i^\dagger e^{-i\hat{H}t}.$$

Then for a zero-temperature ensemble with  $N$  particles, the *time-ordered, single-body, real-time Green's function* (which we call the *Green's function* for short) is a function  $G : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{C}^{d \times d}$  defined by

$$G_{ij}(t, t') = -i \langle \Psi_0^{(N)} | \mathcal{T} \{a_i(t)a_j^\dagger(t')\} | \Psi_0^{(N)} \rangle,$$

where  $\mathcal{T}$  is the *time-ordering operator*, formally defined by

$$\mathcal{T} \{a_i(t)a_j^\dagger(t')\} = \begin{cases} a_i(t)a_j^\dagger(t'), & t' < t \\ \zeta a_j^\dagger(t')a_i(t), & t' \geq t. \end{cases}$$

Note that  $\mathcal{T}$  is not really an operator and it is interpreted merely via the *symbolic* content of its argument.

We can write

$$G(t, t') = G^+(t, t') + G^-(t, t'),$$

where

$$\begin{aligned} iG^+(t, t') &:= \langle \Psi_0^{(N)} | a_i(t)a_j^\dagger(t') | \Psi_0^{(N)} \rangle \theta(t - t'), \\ iG^-(t, t') &:= \langle \Psi_0^{(N)} | a_j^\dagger(t')a_i(t) | \Psi_0^{(N)} \rangle (1 - \theta(t - t')), \end{aligned}$$

with

$$\theta(s) := \begin{cases} 1, & s > 0 \\ 0, & s \leq 0. \end{cases}$$

It is easy to show that  $G(t, t')$ ,  $G^+(t, t')$ , and  $G^-(t, t')$  depend only on  $t - t'$ , so we can define  $G(t) := G(t, 0)$ ,  $G^+(t) := G^+(t, 0)$ , and  $G^-(t) := G^-(t, 0)$  and consider these objects without any loss of information. It is then equivalent to consider the Fourier transforms

$$G(\omega) := \int_{\mathbb{R}} G(t) e^{i\omega t - \eta|t|} dt$$

and likewise  $G^+(\omega)$  and  $G^-(\omega)$  defined similarly, so

$$G(\omega) = G^+(\omega) + G^-(\omega).$$

Here  $\eta$  is interpreted as a positive, infinitesimally small quantity needed to ensure the convergence of the relevant integrals, and  $G(\omega)$ ,  $G^+(\omega)$ , and  $G^-(\omega)$  are not really functions, but rather distributions on  $\mathbb{R}$  defined via the limit  $\eta \rightarrow 0^+$ .

One can show that

$$G_{ij}^+(\omega) = \langle \Psi_0^{(N)} | a_i \frac{1}{\omega - (\hat{H} - E_0^{(N)}) + i\eta} a_j^\dagger | \Psi_0^{(N)} \rangle$$

and

$$G_{ij}^-(\omega) = -\zeta \langle \Psi_0^{(N)} | a_j^\dagger \frac{1}{\omega + (\hat{H} - E_0^{(N)}) - i\eta} a_i | \Psi_0^{(N)} \rangle,$$

where  $E_0^{(N)}$  is the energy of the  $N$ -particle ground state, i.e.,  $\hat{H} | \Psi_0^{(N)} \rangle = E_0 | \Psi_0^{(N)} \rangle$ .

Now we can think of  $G^\pm$  as the restriction to the real axis of the rational function  $G^\pm : \mathbb{C} \rightarrow \mathbb{C}^{d \times d}$  defined by

$$G_{ij}^+(z) := \langle \Psi_0^{(N)} | a_i \frac{1}{z - (\hat{H} - E_0^{(N)})} a_j^\dagger | \Psi_0^{(N)} \rangle$$

$$G_{ij}^-(z) := -\zeta \langle \Psi_0^{(N)} | a_j^\dagger \frac{1}{z + (\hat{H} - E_0^{(N)})} a_i | \Psi_0^{(N)} \rangle,$$

and we can define  $G(z) := G^+(z) + G^-(z)$  accordingly to be rational on  $\mathbb{C}$ .

Note that here we have left out the  $\pm i\eta$  in the denominators, which specified whether poles should be viewed as being infinitesimally above or below the real axis. This erases the distinction between the time-ordered Green's function and the advanced and retarded Green's functions, which we do not define here, though see [77] for details. In fact the distinction does not matter for our sparsity results, which applies equally well in all of these cases.

The self-energy is the rational function  $\Sigma : \mathbb{C} \rightarrow \mathbb{C}^{d \times d}$  defined by

$$\Sigma(z) := z - h - G(z)^{-1}.$$

### 7.3 Green's functions and the self-energy at finite temperature

As above, for  $t \in \mathbb{R}$ , we denote the annihilation and creation operators in the Heisenberg representation by

$$a_i(t) := e^{i\hat{H}t} a_i e^{-i\hat{H}t}, \quad a_i^\dagger(t) := e^{i\hat{H}t} a_i^\dagger e^{-i\hat{H}t}.$$

Then at finite inverse temperature  $\beta \in (0, \infty)$  and chemical potential  $\mu \in \text{int dom } Z$ , the *time-ordered, single-body, real-time Green's function* (which we call the *Green's*

function for short when the context is clear) is a function  $G : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{C}^{d \times d}$  defined by

$$G_{ij}(t, t') = -i \langle \mathcal{T} \{ a_i(t) a_j^\dagger(t') \} \rangle_{\beta, \mu}.$$

We can write

$$G(t, t') = G^+(t, t') + G^-(t, t'),$$

where

$$\begin{aligned} iG^+(t, t') &= \frac{1}{Z} \text{Tr} \left[ a_i(t) a_j^\dagger(t') e^{-\beta(\hat{H} - \mu \hat{N})} \right] \theta(t - t'), \\ iG^-(t, t') &= \frac{\zeta}{Z} \text{Tr} \left[ a_j^\dagger(t') a_i(t) e^{-\beta(\hat{H} - \mu \hat{N})} \right] (1 - \theta(t - t')), \end{aligned}$$

with

$$\theta(s) := \begin{cases} 1, & s > 0 \\ 0, & s \leq 0. \end{cases}$$

as above.

Once again it is easy to show that  $G(t, t')$ ,  $G^+(t, t')$ , and  $G^-(t, t')$  depend only on  $t - t'$ , so we can define  $G(t) := G(t, 0)$ ,  $G^+(t) := G^+(t, 0)$ , and  $G^-(t) := G^-(t, 0)$  and consider these objects without any loss of information. It is then equivalent to consider the Fourier transforms

$$G(\omega) := \int_{\mathbb{R}} G(t) e^{i\omega t - \eta|t|} dt$$

and likewise  $G^+(\omega)$  and  $G^-(\omega)$  defined similarly, so

$$G(z) = G^+(\omega) + G^-(\omega).$$

Now since  $\hat{H}$  preserves particle number, we can safely diagonalize  $\hat{H}$  as an operator on each of the  $N$ -particle subspaces separately. Then the spectrum of  $\hat{H}$  consists of the union of its spectra on the  $N$ -particle subspaces. It follows from Assumption 4 that  $\hat{H} - \mu \hat{N}$  has a ground state, i.e., that its spectrum is bounded from below, for  $\mu \in \text{int dom } Z$ . Let  $m = 0, 1, \dots$ , (terminating at  $m = 2^d$  in the case of fermions) index the spectrum of  $\hat{H}$ , and let  $|\Psi_m\rangle$  denote the  $m$ -th eigenstate. Let  $N_m$  be the particle number of  $|\Psi_m\rangle$  (which is an eigenstate of  $\hat{N}$ ), and let  $E_m$  be defined by  $\hat{H}|\Psi_m\rangle = E_m|\Psi_m\rangle$ .

One can show that

$$G_{ij}^+(\omega) = \frac{1}{Z} \sum_m e^{-\beta(E_m - \mu N_m)} \langle \Psi_m | a_i \frac{1}{\omega - (\hat{H} - E_m) + i\eta} a_j^\dagger | \Psi_m \rangle$$

and

$$G_{ij}^-(\omega) = \frac{-\zeta}{Z} \sum_m e^{-\beta(E_m - \mu N_m)} \langle \Psi_m | a_j^\dagger \frac{1}{\omega + (\hat{H} - E_m) - i\eta} a_i | \Psi_m \rangle.$$

Recall that

$$Z = \sum_m e^{-\beta(E_m - \mu N_m)}.$$

Now we can think of  $G^\pm$  as the restriction to the real axis of the rational function  $G^\pm : \mathbb{C} \rightarrow \mathbb{C}^{d \times d}$  defined by

$$\begin{aligned} G_{ij}^+(z) &:= \frac{1}{Z} \sum_m e^{-\beta(E_m - \mu N_m)} \langle \Psi_m | a_i \frac{1}{z - (\hat{H} - E_m)} a_j^\dagger | \Psi_m \rangle \\ G_{ij}^-(z) &:= \frac{-\zeta}{Z} \sum_m e^{-\beta(E_m - \mu N_m)} \langle \Psi_m | a_j^\dagger \frac{1}{z + (\hat{H} - E_m)} a_i | \Psi_m \rangle, \end{aligned}$$

and we can define  $G(z) := G^+(z) + G^-(z)$  accordingly to be rational on  $\mathbb{C}$ . Once again we have ignored the infinitesimal  $\eta$  in this definition; the same comments made in Appendix 5.1 apply here.

The self-energy is the rational function  $\Sigma : \mathbb{C} \rightarrow \mathbb{C}^{d \times d}$  defined by

$$\Sigma(z) := z - h - G(z)^{-1}.$$

## 7.4 Matsubara Green's functions and self-energy

For  $\tau \in \mathbb{R}$ , we define (abusing notation)

$$a_i(\tau) := e^{(\hat{H} - \mu \hat{N})\tau} a_i e^{-(\hat{H} - \mu \hat{N})\tau}, \quad a_i^\dagger(\tau) := e^{(\hat{H} - \mu \hat{N})\tau} a_i^\dagger e^{-(\hat{H} - \mu \hat{N})\tau}.$$

Although we have overloaded the notation, the distinction between, e.g.,  $a_i(\tau)$  and  $a_i(t)$  should be clear from context. Note carefully that  $a_i^\dagger(\tau)$  is *not* the adjoint of  $a_i(\tau)$ . This is merely a notation. The operators  $a_i(\tau)$  and  $a_i^\dagger(\tau)$  can be thought of as the imaginary-time Heisenberg representation of the annihilation and creation operators. Although the analogy with the real-time Heisenberg representation is broken by considering  $\hat{H} - \mu \hat{N}$  in place of  $\hat{H}$ , our convention is indeed the more widely used due to its naturality in the context of the imaginary-time path integral.

Then at finite inverse temperature  $\beta \in (0, \infty)$  and chemical potential  $\mu \in \text{int dom } Z$ , the *time-ordered, single-body, imaginary-time Green's function* (which we call the *Matsubara Green's function* for clarity) is a function  $G^M : [0, \beta]^2 \rightarrow \mathbb{C}^{d \times d}$  defined by

$$G_{ij}^M(\tau, \tau') = -\langle \mathcal{T} \{ a_i(\tau) a_j^\dagger(\tau') \} \rangle_{\beta, \mu},$$

where  $\mathcal{T}$  here indicates the imaginary-time-ordering operator, formally defined by

$$\mathcal{T} \{ a_i(\tau) a_j^\dagger(\tau') \} = \begin{cases} a_i(\tau) a_j^\dagger(\tau'), & \tau' < \tau \\ \zeta a_j^\dagger(\tau') a_i(\tau), & \tau' \geq \tau. \end{cases}$$

We can write

$$G^M(\tau, \tau') = G^{M,+}(\tau, \tau') + G^{M,-}(\tau, \tau'),$$

where

$$\begin{aligned} -G_{ij}^{\text{M},+}(\tau, \tau') &= \frac{1}{Z} \text{Tr} \left[ a_i(\tau) a_j^\dagger(\tau') e^{-\beta(\hat{H} - \mu \hat{N})} \right] \theta(\tau - \tau'), \\ &= \frac{1}{Z} \text{Tr} \left[ a_i e^{-(\hat{H} - \mu \hat{N})(\tau - \tau')} a_j^\dagger e^{(\tau - \tau' - \beta)(\hat{H} - \mu \hat{N})} \right] \theta(\tau - \tau'), \end{aligned}$$

and

$$\begin{aligned} -G_{ij}^{\text{M},-}(\tau, \tau') &= \frac{\zeta}{Z} \text{Tr} \left[ a_j^\dagger(\tau') a_i(\tau) e^{-\beta(\hat{H} - \mu \hat{N})} \right] (1 - \theta(\tau - \tau')). \\ &= \frac{\zeta}{Z} \text{Tr} \left[ a_j^\dagger e^{-(\hat{H} - \mu \hat{N})(\tau' - \tau)} a_i e^{(\tau' - \tau - \beta)(\hat{H} - \mu \hat{N})} \right] (1 - \theta(\tau - \tau')). \end{aligned}$$

Once again it is easy to show that  $G^{\text{M}}(\tau, \tau')$ ,  $G^{\text{M},+}(\tau, \tau')$ , and  $G^{\text{M},-}(\tau, \tau')$  depend only on  $\tau - \tau'$ . Then the full information of  $G^{\text{M}}(\tau, \tau')$  can be recovered from

$$G^{\text{M}}(\tau) := \begin{cases} G^{\text{M},+}(\tau, 0), & \tau > 0 \\ G^{\text{M},-}(0, -\tau), & \tau \leq 0, \end{cases}$$

defined for  $\tau \in (-\beta, \beta)$ . Now for  $\tau \in (0, \beta)$ , we can compute via the above formulas:

$$G^{\text{M}}(\tau - \beta) = G^{\text{M},-}(0, \beta - \tau) = \zeta G^{\text{M},+}(0, \tau) = \zeta G^{\text{M}}(\tau).$$

Therefore, by considering  $G^{\text{M}}(\tau) = G^{\text{M},+}(\tau, 0)$  only on  $(0, \beta)$ , i.e.,

$$G^{\text{M}}(\tau) = \frac{-1}{Z} \text{Tr} \left[ a_i e^{-\tau(\hat{H} - \mu \hat{N})} a_j^\dagger e^{(\tau - \beta)(\hat{H} - \mu \hat{N})} \right], \quad \tau \in (0, \beta)$$

and extending by  $\beta$ -(anti)periodicity, we can recover the full information of the Matsubara Green's function.

It is then equivalent to consider the frequency-space representation of at the Matsubara frequencies

$$\omega_n = \begin{cases} 2n\pi/\beta, & \zeta = +1 \\ (2n+1)\pi/\beta, & \zeta = -1 \end{cases}$$

for  $n \in \mathbb{Z}$ , defined via

$$G^{\text{M}}(i\omega_n) := \int_0^\beta G^{\text{M}}(\tau) e^{i\omega_n \tau} d\tau,$$

so

$$G^{\text{M}}(\tau) = \frac{1}{\beta} \sum_n G^{\text{M}}(i\omega_n) e^{-i\omega_n \tau}.$$

One can show that

$$G^{\text{M}}(i\omega_n) = G(i\omega_n + \mu),$$

where  $G$  is the rational function  $\mathbb{C} \rightarrow \mathbb{C}^{d \times d}$  defined in the preceding subsection.

The Matsubara self-energy is defined by

$$\Sigma^{\text{M}}(i\omega_n) = i\omega_n - (h - \mu) - G^{\text{M}}(i\omega_n)^{-1} = \Sigma(i\omega_n + \mu),$$

where  $\Sigma$  is the rational function  $\mathbb{C} \rightarrow \mathbb{C}^{d \times d}$  defined in the preceding subsection. Thus to study the Matsubara Green's function and self-energy it suffices to study the rational functions  $G$  and  $\Sigma$  defined earlier.

Finally, we comment that in the imaginary-time representation, we can write

$$-\partial_\tau G^{\text{M}}(\tau, \tau') - (h - \mu)G^{\text{M}}(\tau, \tau') - \int_0^\beta \Sigma^{\text{M}}(\tau, \tau'')G^{\text{M}}(\tau'', \tau') d\tau'' = I_d \delta(\tau - \tau').$$

## Part II

# Bold Feynman diagrams

## 1 Introduction

In quantum many-body physics, the computational complexity of obtaining the numerically exact solution to the many-body Schrödinger equation generally scales exponentially with respect to the number of particles in the system. Hence a direct approach to the quantum many-body problem is intractable for all but very small systems. Many-body perturbation theory (MBPT) formally treats the Coulomb interaction between electrons as a small quantity and provides useful approximations to many quantities of physical interest with significantly reduced computational cost. MBPT has been demonstrated to be quantitatively useful, and sometimes indispensable, in a wide range of scientific applications. These range from the early description of helium atoms and the uniform electron gas [33] to modern theories of photovoltaics and the optical excitation of electrons [5, 79, 11]. Even for ‘strongly correlated’ systems [37, 19] where a perturbation theory is known to be unsuitable, MBPT still provides the basic building blocks used in many successful approaches [53].

MBPT is usually formulated in the language of second quantization. For many problems of interest, the Hamiltonian can be split into a single-particle term and a two-particle term, which are respectively quadratic and quartic in the creation and annihilation operators [33]. These operators are defined on a Fock space whose dimension scales exponentially with respect to the system size. Nonetheless, in the special case of ‘non-interacting’ systems, in which the Hamiltonian contains only the quadratic term, quantities of interest can be obtained exactly. Hence the non-interacting system is naturally taken as a reference system. The remaining quartic interaction, which arises from the Coulomb interaction between electrons and makes the system ‘interacting,’ is responsible for almost all of the difficulties in quantum many-body physics. MBPT treats the quartic term as a perturbation to the non-interacting system.

Feynman diagrams arise naturally in MBPT as a bookkeeping device for the coefficients of perturbative series, though they can be further endowed with physical interpretation [33]. Initially these diagrams involve contributions from the so-called non-interacting Green’s function (alternatively known as the bare propagator) that specifies the non-interacting reference problem, as well the quartic interaction. Virtually all physical quantities of interest can be represented perturbatively via such a *bare* Feynman diagrammatic expansion. Remarkably, the bare Feynman diagrammatic expansions of certain quantities can be simplified into *bold* Feynman diagrammatic expansions [100]. Such an expansion is obtained from a bare expansion via a ‘partial resummation’ procedure, which selects certain pieces of bare diagrams and sums their contribution to infinite order. In the bold diagrams, the role of the bare propagator

is assumed by the interacting Green’s function, alternatively known as the bold propagator. This procedure, referred to as ‘dressing’ or ‘renormalizing’ the propagator, is *a priori* valid only in a formal sense. Although it may be initially motivated as an attempt to simplify the diagrammatic expansion, the passage from bare to bold diagrams has significant implications. In particular, most Green’s function methods for the theoretical and numerical investigation of quantum many-body physics, such as the self-consistent Hartree-Fock approximation, the second-order Green’s function approximation (GF2), the GW approximation [43], the dynamical mean-field theory (DMFT) [37, 85], the GW+DMFT method [12], the dynamical cluster approximation [99], and bold diagrammatic Monte Carlo methods [86, 57], can be derived via summation over some (possibly infinite) subset of the bold diagrams.

All of these methods, as well as the bold diagrammatic expansion itself, can be viewed as resting on a foundation known as the Luttinger-Ward (LW)<sup>3</sup> formalism [65] that originated in 1960. This formalism has found widespread usage in physics and chemistry [28, 45, 10, 90]. However, the LW formalism, which is based on a functional of the same name, is defined only formally. This is a serious issue both in theory and in practice. Indeed, even the very existence of the LW functional in the context of fermionic systems is under debate, with numerical evidence to the contrary appearing in the past few years [54, 32, 103, 42] in the physics community.

## 1.1 Contributions

Please note that this Part is based on [61] (joint work with Lin Lin). In this Part, we provide a self-contained explanation of MBPT in the setting of a *Gibbs model* (alternatively, following the physics literature, ‘Euclidean lattice field theory’). The perturbation theory of this model, with a specific form of quartic interaction that we refer to as the *generalized Coulomb interaction*, enjoys a correspondence with the Feynman diagrammatic expansion for the quantum many-body problem with two-body interaction [77, 2, 1]. The model is also of interest in its own right and includes, e.g., the (lattice)  $\varphi^4$  theory [2, 108], as a special case. In the setting of the Gibbs model, one is interested in the computation of expectation values with respect to possibly high-dimensional Gibbs measures. While the exact computation of such high-dimensional integrals is generally intractable, important exceptions are the Gaussian integrals, i.e., integrals for the moments of a Gaussian measure, which can be evaluated exactly. Hence the Gaussian measure plays the role of the reference system. One can construct perturbation series using Feynman diagrams, which correspond to moments of Gaussian measures, to evaluate quantities of interest.

The main contribution of this Part is the rigorous justification of the bold diagrammatic expansion in the Gibbs measure setting. Although the basic idea of the passage from bare to bold diagrams can be intuitively perceived, the validity of this

---

<sup>3</sup>The Luttinger-Ward formalism is also known as the Kadanoff-Baym formalism [7] depending on the context. In this work we always use the former.



procedure actually relies on subtle combinatorial arguments, which to the extent of our knowledge, have not appeared in the literature. We remark that the arguments appearing in this Part regarding these manipulations are just as applicable to the quantum many-body problem as they are to the Gibbs model. Furthermore, these arguments clarify why certain quantities such as the self-energy admit a bold diagrammatic expansion, while other quantities, such as the free energy, do not.

In fact, bosonic and fermionic field theories (which can in particular be derived from the non-relativistic quantum many-body problem via the coherent state path integral formalism [1, 77]) can be viewed formally as infinite-dimensional Gibbs measures over complex and Grassmann fields, respectively, in contrast to the real ‘fields’ considered in this work. The diagrammatic expansions for such theories yield propagator lines with a direction (indicated by an arrow), due to the distinction between creation and annihilation operators. In the setting of the two-body interaction, this additional structure significantly reduces the symmetry of the Feynman diagrams, and in fact the self-energy and single-particle Green’s function diagrams all have a symmetry factor of one. This greatly simplifies the proof of the bold diagrammatic expansion in these settings (although a proof of the unique skeleton decomposition as in Proposition 24 is still necessary). However, we view this simplification as largely accidental because it does not extend to interactions beyond the two-body interaction, where more sophisticated arguments are necessary (and indeed, to our knowledge, bold diagrams have not yet been considered). By contrast the tools introduced here can be applied with minimal modification to more complicated interaction forms.

As an auxiliary tool for carefully establishing diagrammatic expansions (both bare and bold), we have also found it necessary to introduce definitions of the various flavors of Feynman diagrams (as well as associated notions of isomorphism, automorphism, etc.) in a way that is new, as far as we know. Most of this perspective (which views Feynman diagrams as data structures with half-edges as the fundamental building block) is conveyed in section 3.2. We have also aimed to make this framework durable enough not just for the developments of this Part, but also to allow us to pursue further (and more sophisticated) diagrammatic manipulations, such as the development of the bold screened diagrams, Hedin’s equations [43], and the Bethe-Salpeter equation [79], in future work.

## 1.2 Related work

The construction of Feynman diagrams in the setting of the Gibbs measure is well known, particularly in quantum field theory [83, 1]. To our knowledge, this setting is mostly discussed as a prelude to the setting of quantum field theory (in particular, via the coherent state path integral, the quantum many-body setting) [1, 77], or to more general mathematical settings arising in geometry and topology [29, 84].

### 1.3 Perspectives

MBPT is known to be difficult to work with, both analytically and numerically. In fact, even learning MBPT can be difficult without a considerable amount of background knowledge in physics. Hence, more broadly, we hope that this Part will serve as an introduction to bare and bold Feynman diagrams that is self-contained, rigorous, and accessible to a mathematical audience without a background in quantum physics. The prerequisites for understanding this part of the two-part series are just multivariable calculus and some elementary combinatorics.

In fact, our perspective is that the Gibbs model can be used as a point of departure (especially for mathematicians) for the study of the many-body problem in three senses: (1) theoretically, (2) numerically, and (3) pedagogically. We shall elaborate on these three points presently.

(1) Virtually all of the important concepts of MBPT for the quantum many-body problem—such as Green’s functions, the self-energy, the bare and bold diagrams, and the Luttinger-Ward formalism, to name a few—have analogs in the setting of the Gibbs model. The same can be said of virtually all Green’s function methods, including all of the methods cited above. Furthermore, there is an analog of the impurity problem, which is fundamental in quantum embedding theory [101].

When rigorous theoretical understanding of the quantum many-body problem becomes difficult, a lateral move to the more tractable Gibbs model may yield interesting results. Headway in this direction is reported Part III, in which the Luttinger-Ward formalism is established rigorously for the first time, and further progress will be reported in future work. Moreover, studying the extent to which results *fail* to translate between settings, given the formal correspondence between the two, may yield interesting insights.

(2) Numerical methods in MBPT are known to be difficult to implement, and the calculations are often found to be difficult to converge, time-consuming, or both. Given that virtually all Green’s function methods of interest translate to the Gibbs model, this setting can serve as a sandbox for the evaluation and comparison of methods in various regimes. Indeed, one can benchmark these methods by obtaining essentially exact approximations of the relevant integrals via Monte Carlo techniques. We hope that the Gibbs model can provide new insights into a number of difficult issues in MBPT, such as the role of self-consistency in the GW method, the appropriate choice of vertex correction beyond the GW method, and the study of embedding methods.

(3) For a mathematical reader, the literature of MBPT can be difficult to digest. In our view the consideration of the Gibbs model offers perhaps the simplest introduction to the major concepts of MBPT. This Part, together with a familiarity with second quantization and the basics of many-body Green’s functions, should equip the reader to follow the development of the various approximations and methods found in the literature, by distilling these concepts via the Gibbs model. We have attempted to respect this goal by maintaining a pedagogical style of exposition, with many examples

provided throughout for concreteness.

## 1.4 Outline

In section 2 (‘Preliminaries’) we formally introduce the Gibbs model as well as its associated physical quantities such as the partition function, the free energy, and the Green’s function. Here we prove the classic formula (Theorem 2) attributed to Isserlis and Wick for computing the moments of a Gaussian measure, which is the basis for all Feynman diagrammatic expansions. We also quickly recover the Galitskii-Migdal formula (Theorem 1) from quantum many-body physics in this setting using a scaling argument.

In section 3, we introduce various flavors of Feynman diagrams and use them to compute diagrammatic expansions for the partition function (Theorem 14), the free energy (Theorem 17), and the Green’s function (Theorem 18). Then, motivated by the prospect of simplifying the perturbative computation of the Green’s function, we introduce the self-energy and the Dyson equation, which can be used to recover the Green’s function once the self-energy is known, and then compute the diagrammatic expansion of the self-energy (Theorem 20).

In section 4, the main goal is to formulate and prove the bold diagrammatic expansion of the self-energy (Theorem 32). This result is only a fact about formal power series, but in Part III, we will show that the bold diagrammatic expansion admits an analytical interpretation as an asymptotic series, in a sense that we preview in Remark 33. Theorem 32, which is a combinatorial result, is in fact used in establishing the analytical fact in Part III. In section 4.7 we provide an overview of Green’s function methods, including a diagrammatic derivation of the GW method and a discussion of a property known as  $\Phi$ -derivability. In section 4.8, we provide a preview of the Luttinger-Ward formalism from the diagrammatic perspective and explain how the LW functional relates to the free energy.

Finally, in section 5 we consider a few basic numerical experiments with Green’s function methods for the Gibbs model.

## 2 Preliminaries

Before discussing Feynman diagrams in proper, we discuss various preliminaries, including the basics of Gaussian integration. For simplicity we restrict our attention to real matrices, though analogous results can be obtained in the complex Hermitian case.

## 2.1 Notation

First we recall some basic facts from calculus. For a real symmetric positive definite matrix  $A \in \mathbb{R}^{N \times N}$ , we define

$$Z_0 := \int_{\mathbb{R}^N} e^{-\frac{1}{2}x^T A x} dx = (2\pi)^{\frac{N}{2}} (\det(A))^{-\frac{1}{2}}. \quad (2.1)$$

The two-point correlation function  $G^0$  is an  $N \times N$  matrix with entries

$$G_{ij}^0 := \frac{1}{Z_0} \int_{\mathbb{R}^N} x_i x_j e^{-\frac{1}{2}x^T A x} dx = (A^{-1})_{ij}, \quad (2.2)$$

i.e.,  $G^0 = A^{-1}$ . (We place the ‘0’ in the superscript merely to accommodate the use of indices more easily in the notation.) Note that  $G^0$  is the covariance matrix  $\mathbb{E}(XX^T) = A^{-1}$  of the  $N$ -dimensional Gaussian random variable  $X \sim \mathcal{N}(0, A^{-1})$ .

Now consider a more general  $N$ -dimensional integral, called the partition function, given by

$$Z = \int_{\mathbb{R}^N} e^{-\frac{1}{2}x^T A x - U(x)} dx, \quad (2.3)$$

where  $U(x)$  is called the interaction term. Throughout this Part, we take  $U$  to be the following quartic polynomial:

$$U(x) = \frac{1}{8} \sum_{i,j=1}^N v_{ij} x_i^2 x_j^2. \quad (2.4)$$

Without loss of generality we assume that  $v_{ij} = v_{ji}$ , since otherwise we can always replace  $v_{ij}$  by  $(v_{ij} + v_{ji})/2$  without changing the value of  $U(x)$ . The factor of 8 comes from the fact that we do not distinguish between the  $i$  and  $j$  indices (due to the symmetry of the  $v$  matrix), nor do between the two terms  $x_i x_i$  and  $x_j x_j$ . Each will contribute a symmetry factor of 2 in the developments that follow, and this convention simplifies the bookkeeping of the constants in the diagrammatic series. Quartics of the form (1.6) arise from the discretization of the  $\varphi^4$  theory [2] and moreover as a classical analog of the interaction arising in quantum many-body settings, such as the Coulomb interaction of electronic structure theory and the interaction term in simplified condensed matter models such as the Hubbard model [68]. With some abuse of notation, we will refer to the interaction (1.6) as the *generalized Coulomb interaction*.

Our results generalize quite straightforwardly to other interactions. We will comment in section 3.4 on diagrammatic developments for other interactions. But for concreteness of the example expressions and diagrams throughout, it is simpler to stick to the generalized Coulomb interaction as a reference.

Let  $\mathcal{S}^N$ ,  $\mathcal{S}_+^N$ , and  $\mathcal{S}_{++}^N$  denote respectively the sets of symmetric, symmetric positive semidefinite, and symmetric positive definite  $N \times N$  real matrices. We also

require that

$$\frac{1}{2}x^T Ax + U(x) \rightarrow +\infty, \quad \|x\| \rightarrow +\infty \quad (2.5)$$

for any  $A \in \mathcal{S}^N$  and moreover that the growth in Eq. (2.6) that the integral (1.4) is well-defined. Here  $\|\cdot\|$  is the vector 2-norm. Note that Eq. (2.6) does not require  $A$  to be positive definite. For instance, if  $N = 1$ , then Eq. (1.4) becomes

$$Z = \int_{\mathbb{R}} e^{-\frac{1}{2}ax^2 - \frac{1}{8}vx^4} dx, \quad (2.6)$$

and the expression in (2.6) is well-defined as long as  $v > 0$ , regardless of the sign of  $a$ . Nonetheless, we assume that  $A \in \mathcal{S}_{++}^N$ , as this assumption is necessary for the construction of a perturbative series in the interaction strength. In Part III, Eq. (2.6) will help us understand the behavior of bold diagrammatic methods when  $A$  is indefinite.

For general  $N \geq 1$ , there is a natural condition on the matrix  $v$  that ensures that integrals like (2.6) are convergent, namely that the matrix  $v$  is positive definite. Indeed, this assumption ensures in particular that  $U$  is a nonnegative polynomial, strictly positive away from  $x = 0$ . Since  $U$  is homogeneous quartic, it follows that  $U \geq C^{-1}|x|^4$  for some constant  $C$  sufficiently large, so for any  $A$ , the expression  $\frac{1}{2}x^T Ax + U(x)$  goes to  $+\infty$  quartically as  $\|x\| \rightarrow \infty$ . Another sufficient assumption is that the entries of  $v$  are nonnegative and moreover that the diagonal entries are strictly positive. We will explore the implication of such conditions in future work.

To simplify the notation, for any function  $f(x)$ , we define

$$\langle f \rangle = \frac{1}{Z} \int_{\mathbb{R}^N} f(x) e^{-\frac{1}{2}x^T Ax - U(x)} dx, \quad \langle f \rangle_0 = \frac{1}{Z_0} \int_{\mathbb{R}^N} f(x) e^{-\frac{1}{2}x^T Ax} dx. \quad (2.7)$$

Throughout this Part we are mostly interested in computing two quantities. The first is the *free energy*, defined as the negative logarithm of the partition function:

$$\Omega = -\log Z. \quad (2.8)$$

The second is the two-point correlation function (also called the *Green's function* by analogy with the quantum many-body literature), which is the  $N \times N$  matrix

$$G_{ij} = \frac{1}{Z} \int_{\mathbb{R}^N} x_i x_j e^{-\frac{1}{2}x^T Ax - U(x)} dx =: \langle x_i x_j \rangle. \quad (2.9)$$

It is important to recognize that

$$G \in \mathcal{S}_{++}^N. \quad (2.10)$$

In fact, as we shall see in Part III, this constraint defines the domain of ‘physical’ Green’s functions, in a certain sense. In the discussion below,  $G$  is also called the interacting Green’s function, in contrast to the *non-interacting Green’s function*  $G^0 = A^{-1}$ . The non-interacting and interacting Green’s functions are also often called the *bare* and *bold propagators*, respectively, especially in the context of diagrams.

## 2.2 Scaling relation

The homogeneity of the quartic term  $U(x)$  allows for the derivation of a *scaling relation* for the partition function. Define the  $\lambda$ -dependent partition function as

$$Z_\lambda = \int_{\mathbb{R}^N} e^{-\frac{1}{2}x^T Ax - \lambda U(x)} dx. \quad (2.11)$$

Then by an change of variable  $y = \lambda^{\frac{1}{4}}x$ , we have

$$Z_\lambda = \lambda^{-\frac{N}{4}} \int_{\mathbb{R}^N} e^{-\frac{1}{2\sqrt{\lambda}}y^T Ay - U(y)} dy. \quad (2.12)$$

The scaling relation allows us to represent other averaged quantities using the two-point correlation function. One example is given in Theorem 1, which is analogous to the computation of a quantity called the (internal) energy using the Galitskii-Migdal formula in quantum physics [68].

**Theorem 1** (Galitskii-Migdal). *The internal energy*

$$E := \left\langle \frac{1}{2}x^T Ax + U(x) \right\rangle = \frac{1}{Z} \int_{\mathbb{R}^N} \left( \frac{1}{2}x^T Ax + U(x) \right) e^{-\frac{1}{2}x^T Ax - U(x)} dx \quad (2.13)$$

can be computed using the two point correlation function  $G$  as

$$E = \frac{1}{4} \text{Tr}[AG + I], \quad (2.14)$$

where  $I$  is the  $N \times N$  identity matrix.

*Proof.* By the definition of  $G$ , we have

$$\left\langle \frac{1}{2}x^T Ax \right\rangle = \frac{1}{2} \text{Tr} [A \langle xx^T \rangle] = \frac{1}{2} \text{Tr}[AG]. \quad (2.15)$$

In order to evaluate  $\langle U(x) \rangle$ , we consider the  $\lambda$ -dependent partition function in Eq. (2.11), and we have

$$-\frac{dZ_\lambda}{d\lambda} \Big|_{\lambda=1} = \int_{\mathbb{R}^N} U(x) e^{-\frac{1}{2}x^T Ax - U(x)} dx. \quad (2.16)$$

Using the scaling relation in Eq. (2.12), we have

$$-\frac{dZ_\lambda}{d\lambda} \Big|_{\lambda=1} = \frac{N}{4} Z - \int_{\mathbb{R}^N} \frac{1}{4} y^T Ay e^{-\frac{1}{2}y^T Ay - U(y)} dy. \quad (2.17)$$

Combining Eq. (2.15) to (2.17) we have

$$E = \left\langle \frac{1}{2}x^T Ax + U(x) \right\rangle = \frac{1}{4} \langle x^T Ax + N \rangle = \frac{1}{4} \text{Tr}[AG + I].$$

□

## 2.3 Wick theorem

We first introduce the following notation. For an even number  $m$ , we denote by  $\mathcal{I}_m$  a set of integers  $\{1, \dots, m\}$ . For  $i \neq j \in \mathcal{I}_m$ , we call  $(i, j)$  a pair. A *pairing*  $\sigma$  on  $\mathcal{I}_m$  is defined to be a partition of  $\mathcal{I}_m$  into  $k$  disjoint pairs. For example, the set of all possible pairings of the set  $\mathcal{I}_4 = \{1, 2, 3, 4\}$  is  $\{(1, 2)(3, 4), (1, 3)(2, 4), (1, 4)(2, 3)\}$ . Note that a pairing  $\sigma$  can be viewed as an element of the permutation group  $\text{Sym}(\mathcal{I}_m)$ , such that  $\sigma^2 = 1$  and whose action on  $\mathcal{I}_m$  has no fixed points. Under this interpretation  $\sigma$  maps any element  $i \in \mathcal{I}_m$  to the element  $\sigma(i)$  of the pairing containing  $i$ . For a given pairing  $\sigma$ , we define the set

$$\mathcal{I}_m/\sigma := \{i \in \mathcal{I}_m \mid i < \sigma(i)\}$$

to be the collection of indices corresponding to the ‘first element’ of each pair. Denote by  $\Pi(\mathcal{I}_m)$  the set of all possible pairings. Observe that there are

$$|\Pi(\mathcal{I}_m)| = \frac{m!}{2^{m/2}(m/2)!}$$

pairings in total.

Now Wick’s theorem (Theorem 2), also known as Isserlis’ theorem [46] in probability theory, is the basic tool for deriving the Feynman rules for diagrammatic expansion. For completeness we give a proof, but since this is a classic result, it is provided in Appendix A.

**Theorem 2** (Isserlis-Wick). *For integers  $1 \leq \alpha_1, \dots, \alpha_m \leq N$ ,*

$$\langle x_{\alpha_1} \cdots x_{\alpha_m} \rangle_0 = \begin{cases} 0, & m \text{ is odd,} \\ \sum_{\sigma \in \Pi(\mathcal{I}_m)} \prod_{i \in \mathcal{I}_m/\sigma} G_{\alpha_i, \alpha_{\sigma(i)}}^0, & m \text{ is even.} \end{cases} \quad (2.18)$$

In Theorem 2, the indices  $\alpha_i$  do not need to be distinct from one another. For example, for  $N = 4$ ,

$$\langle x_1 x_2 x_3 x_4 \rangle_0 = G_{12}^0 G_{34}^0 + G_{13}^0 G_{24}^0 + G_{14}^0 G_{23}^0,$$

and

$$\langle x_1^2 x_3 x_4 \rangle_0 = G_{11}^0 G_{34}^0 + G_{13}^0 G_{14}^0 + G_{14}^0 G_{13}^0 = G_{11}^0 G_{34}^0 + 2G_{14}^0 G_{13}^0.$$

Similarly

$$\langle x_1^4 \rangle_0 = G_{11}^0 G_{11}^0 + G_{11}^0 G_{11}^0 + G_{11}^0 G_{11}^0 = 3G_{11}^0 G_{11}^0.$$

## 3 Feynman diagrams

Let us now consider the expansion of quantities such as  $Z$ ,  $\Omega$ , and  $G$  with respect to a (small) interaction term. For the case currently under consideration, in which  $U$

is of the form (1.6), the size of the interaction term is measured by the magnitude of the coefficients  $v_{ij}$ . Equivalently, we can consider a  $\lambda$ -dependent interaction as in the definition of the  $\lambda$ -dependent partition function  $Z_\lambda$  and expand in the small parameter  $\lambda$ . This motivates us to expand  $e^{-U(x)}$  using a Taylor series, i.e.

$$Z = \int_{\mathbb{R}^N} \sum_{n=0}^{\infty} \frac{1}{n!} (-U(x))^n e^{-\frac{1}{2}x^T A x} dx \sim \sum_{n=0}^{\infty} \frac{1}{n!} \int_{\mathbb{R}^N} (-U(x))^n e^{-\frac{1}{2}x^T A x} dx. \quad (3.1)$$

The ‘ $\sim$ ’ indicates that interchanging the order of integration of summation leads only to an asymptotic series expansion with respect to the interaction strength, also called the coupling constant [68]. This can be readily seen for the example with  $n = 1$  of Eq. (2.6), where

$$Z_\lambda \sim \int \sum_{n=0}^{\infty} \frac{1}{n!} \left( -\frac{1}{8} \lambda x^4 \right)^n e^{-\frac{1}{2}x^2} dx = \sum_{n=0}^{\infty} \frac{(-1)^n \lambda^n}{n!} 2^{-n+\frac{1}{2}} \Gamma\left(2n + \frac{1}{2}\right). \quad (3.2)$$

Here  $\Gamma(\cdot)$  is the Gamma-function. It is clear that the series has zero convergence radius, and the series is only an *asymptotic series* in the sense that the error of the truncation to  $n$ -th order is  $O(\lambda^{n+1})$  as  $\lambda \rightarrow 0^+$ .

One might have guessed that the radius of convergence must be zero by the following heuristic argument: evidently  $Z_\lambda = +\infty$  for any  $\lambda < 0$ , which suggests that the radius of convergence cannot be positive at  $\lambda = 0$ .

In general since  $U$  is a quartic polynomial in  $x$ , the  $n$ -th term in Eq. (1.5) can be expressed as the linear combination of a number of  $4n$ -point correlation functions for a Gaussian measure. These can be readily evaluated using the Wick theorem.

To motivate the need for Feynman diagrams, we first compute the first few terms of the expansion for the partition function ‘by hand’.

The 0-th order term in (1.5) is clearly  $Z_0$ . Using the Wick theorem, the first-order contribution to  $Z_0$  contains two terms as

$$-Z_0 \sum_{i,j} \frac{1}{8} v_{ij} \langle x_i^2 x_j^2 \rangle_0 = -Z_0 \sum_{i,j} v_{ij} \left( \frac{1}{8} G_{ii}^0 G_{jj}^0 + \frac{1}{4} G_{ij}^0 G_{ij}^0 \right). \quad (3.3)$$

The second-order contribution, however, can be seen with some effort to contain 8 distinct terms as

$$\begin{aligned} Z_0 \frac{1}{2!} \sum_{i_1, j_1, i_2, j_2} \frac{1}{8^2} v_{i_1 j_1} v_{i_2 j_2} \langle x_{i_1}^2 x_{j_1}^2 x_{i_2}^2 x_{j_2}^2 \rangle_0 \\ = Z_0 \sum_{i_1, j_1, i_2, j_2} v_{i_1 j_1} v_{i_2 j_2} \left[ \left( \frac{1}{2! \cdot 8^2} G_{i_1 i_1}^0 G_{j_1 j_1}^0 G_{i_2 i_2}^0 G_{j_2 j_2}^0 \right. \right. \\ \left. \left. + \frac{1}{2! \cdot 4^2} G_{i_1 j_1}^0 G_{i_1 j_1}^0 G_{i_2 j_2}^0 G_{i_2 j_2}^0 + \frac{1}{4 \cdot 8} G_{i_1 i_1}^0 G_{j_1 j_1}^0 G_{i_2 j_2}^0 G_{i_2 j_2}^0 \right) \right] \end{aligned}$$



$$\begin{aligned}
& + \left( \frac{1}{2! \cdot 8} G_{i_1 i_1}^0 G_{i_2 i_2}^0 G_{j_1 j_2}^0 G_{j_1 j_2}^0 + \frac{1}{2 \cdot 2} G_{i_1 j_1}^0 G_{i_2 j_2}^0 G_{i_1 i_2}^0 G_{j_1 j_2}^0 + \frac{1}{4} G_{i_1 j_1}^0 G_{i_1 i_2}^0 G_{j_1 i_2}^0 G_{j_2 j_2}^0 \right) \\
& + \left( \frac{1}{2! \cdot 8} G_{i_1 i_2}^0 G_{i_1 i_2}^0 G_{j_1 j_2}^0 G_{j_1 j_2}^0 + \frac{1}{2! \cdot 4} G_{i_1 i_2}^0 G_{j_1 i_2}^0 G_{i_1 j_2}^0 G_{j_1 j_2}^0 \right) \Big]. \quad (3.4)
\end{aligned}$$

The form in which this expression has been written (in particular, the form of the denominators of the pre-factors) will become clear later on.

Following the same principle, one can derive higher-order contributions to  $Z$ . However, the number of distinct terms in each order grows combinatorially with respect to  $n$ . The number of distinct terms, as well as the associated pre-constants, are already non-trivial in the second-order expansion. Feynman diagrams provide a graphical way to systematically organize such terms.

### 3.1 Motivation

In fact it is helpful to view  $-v_{ij}x_i^2x_j^2$  as the contraction of the fourth-order tensor  $-u_{ikjl}x_i x_j x_k x_l$ , where  $u_{ikjl} = v_{ij}\delta_{ik}\delta_{jl}$ . (Notice that  $u_{ikjl}$  is invariant under the exchange of the first two indices with one another, of the last two indices with one another, and of the first two indices with the last two indices. This yields an eightfold redundancy that will become relevant later on.) Using this insight we can expand the  $n$ -th term in the series of Eq. (1.5) as

$$\frac{Z_0}{8^n n!} \sum_{i_1, j_1, k_1, l_1, \dots, i_n, j_n, k_n, l_n=1}^N \left( \prod_{m=1}^n -v_{i_m j_m} \delta_{i_m k_m} \delta_{j_m l_m} \right) \left\langle \prod_{m=1}^n x_{i_m} x_{j_m} x_{k_m} x_{l_m} \right\rangle_0. \quad (3.5)$$

One can then use the Wick theorem to express this quantity as a sum over pairings of  $\mathcal{I}_{4n}$ . However, it is easier to represent the pairings graphically in the following way. For each  $m = 1, \dots, n$ , we draw one copy of Fig. 2 (b), i.e., a *wiggled line* known as the interaction line, with four dangling *half-edges* labeled  $i, j, k, l$ .

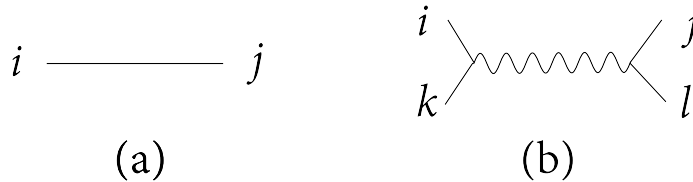


Figure 2: (a) the bare propagator,  $G_{ij}^0$ . (b) the interaction,  $-v_{ij}\delta_{ik}\delta_{jl}$ .

We can then number each interaction line as  $1, \dots, n$  and indicate this by adding an appropriate subscript to the labels  $i, j, k, l$  associated to this vertex. (For the first-order terms, since there is only one interaction line, we may skip this step.) The  $4n$  half-edges, each with a unique label, represent the set on which we consider

pairings. We depict a pairing by linking the paired half-edges with a *straight line*, which represents the bare propagator  $G^0$ . The resulting figure is a (labeled, closed) Feynman diagram of order  $n$ . An example of order 2 is depicted in Fig. 3.

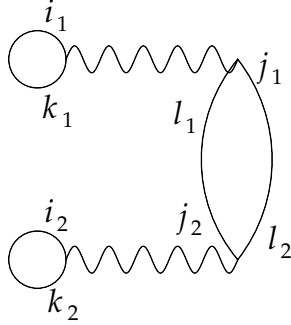


Figure 3: A labeled closed Feynman diagram of order 2.

The quantity associated via Wick's theorem with the pairing represented by such a diagram can then be computed by taking a product over all propagators and interaction lines of the associated quantities indicated in Fig. 2 (a) and (b), respectively. For instance, a line between half-edges  $i_1$  and  $k_2$  would yield the factor  $G_{i_1 k_2}^0$ . Meanwhile, the contribution of the interaction lines altogether is  $\prod_{m=1}^n v_{i_m j_m} \delta_{i_m k_m} \delta_{j_m l_m}$ . The resulting product is then summed over the indices  $i_1, j_1, k_1, l_1, \dots, i_n, j_n, k_n, l_n$ . For the example depicted Fig. 3, this procedure yields the sum

$$\sum_{i_1, j_1, k_1, l_1, i_2, j_2, k_2, l_2} v_{i_1 j_1} \delta_{i_1 k_1} \delta_{j_1 l_1} v_{i_2 j_2} \delta_{i_2 k_2} \delta_{j_2 l_2} G_{i_1 k_1}^0 G_{j_1 l_1}^0 G_{l_1 j_2}^0 G_{i_2 k_2}^0 = \sum_{i_1, j_1, i_2, j_2} v_{i_1 j_1} v_{i_2 j_2} G_{i_1 i_1}^0 G_{i_2 i_2}^0 G_{j_1 j_2}^0 G_{j_1 j_2}^0. \quad (3.6)$$

In summary, we can graphically represented the sum over pairings furnished by Wick's theorem as a sum over such diagrams. It is debatable whether we have really made any progress at this point; keeping in mind that the diagrams we have constructed distinguish labels, there are as many diagrams to sum over as there are pairings of  $\mathcal{I}_{4n}$ . Nonetheless, we can use our new perspective to group similar diagrams and mitigate the proliferation of terms at high order.

Indeed, many diagrams yield the same contribution. In Fig. 4, the labeled first-order diagrams are depicted. Fig. 4 (b) and (b') differ only by a relabeling that swaps  $j$  and  $l$  and so yield the same contribution after indices are summed over. From another point of view, after removing labels these diagrams become 'topologically equivalent', or isomorphic in some sense.

Our goal is to remove this redundancy in our summation by summing only over *unlabeled* diagrams. One expects that the 'amount' of redundancy of each unlabeled

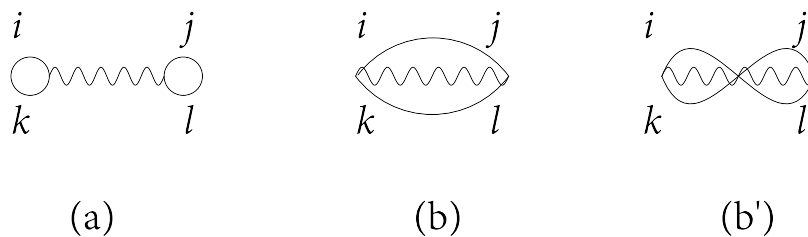


Figure 4: First order expansion for  $Z$  with labeled diagrams. (a) (b) correspond to the first and second term in Eq. (3.3). (b') gives an equivalent term to (b) and should not be counted twice.

diagram is measured by its symmetry in a certain sense. Before making these notions precise, we provide careful definitions of labeled and unlabeled closed Feynman diagrams.

### 3.2 Labeled and unlabeled diagrams

We begin with a definition of unlabeled closed Feynman diagrams, and then define labeled diagrams as unlabeled diagrams equipped with extra structure. Given  $n$  unlabeled interaction lines, each with four dangling half-edges, intuitively speaking we produce an unlabeled closed Feynman diagram by linking half-edges according to a pairing on all  $4n$  of them. By linking together the half-edges dangling from a single interaction line, one can produce only the two ‘topologically distinct’ diagrams shown in Fig. 5. By applying the linking procedure to two interaction lines, one obtains the diagrams in Fig. 6.



Figure 5: Unlabeled closed Feynman diagrams of order 1. In many-body perturbation theory, the left-hand diagram corresponds to the ‘Hartree’ term and is often referred to as the ‘dumbbell’ diagram. The right-hand diagram corresponds to the ‘Fock exchange’ term and is often referred to as the ‘oyster’ diagram.

Observe that via this linking procedure, each interaction line can be viewed as a *vertex* of degree 4 in an undirected graph with some *additional structure*, in particular a partition of the four half-edges that meet at the vertex into two pairs of half-edges (separated by the wiggled line). Half-edges from the same interaction line may be linked, so in fact the resulting graph may have self-edges (or loops). (In an undirected

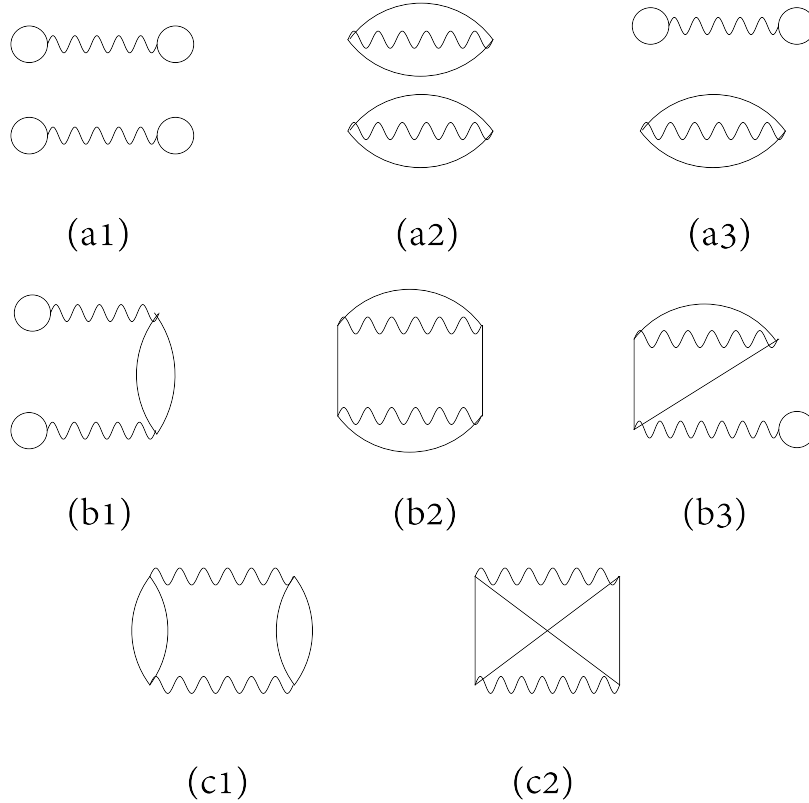


Figure 6: Unlabeled closed Feynman diagrams of order 2.

graph with self-edges, each self-edge contributes 2 to the degree of the vertex, so that the degree indicates the number of half-edges emanating from a vertex.)

In fact it is more natural to view closed Feynman diagrams as being specified via the linking of half-edges than it is to view them as undirected graphs specified by vertex and edge sets  $(V, E)$ . We now provide careful definitions.

**Definition 3.** An unlabeled closed Feynman diagram  $\Gamma$  of order  $n$  consists of a vertex set  $V$  with  $|V| = n$  and the following extra structure. To the vertices  $v \in V$  there are associated disjoint sets  $H_1(v)$  and  $H_2(v)$  each of cardinality 2. The union  $H(v) := H_1(v) \cup H_2(v)$  is the ‘half-edge set’ of the vertex (or ‘interaction’)  $v$ , and the partition  $\{H_1(v), H_2(v)\}$  reflects the separation of the half-edges into two pairs separated by a wiggled line. The (disjoint) union  $\bigcup_{v \in V} H(v)$  is equipped with a partition  $\Pi$  into  $2n$  pairs of half-edges.<sup>4</sup> In total we can view the unlabeled diagram  $\Gamma$  as the tuple  $\Gamma = (V, H_1, H_2, \Pi)$ . For any half-edge  $h \in \bigcup_{v \in V} H(v)$ , let the unique vertex  $v$  associated with this half-edge be denoted by  $v = v(h)$ .

<sup>4</sup>Intuitively speaking, these data specify a recipe for linking up half-edges to form a connected undirected graph of degree 4, but the previously specified data are a more natural representation of the diagram, especially once labels are introduced.

**Notation 4.** As a matter of notation going forward, we stress that we maintain a careful distinction in the notation between sets or pairs  $\{\cdot, \cdot\}$ , e.g., of half-edges, in which the order of the terms does not matter, and ordered pairs  $(\cdot, \cdot)$ , e.g., of half-edges, in which the order matters.

We will often refer to different flavors of Feynman diagrams simply as diagrams when the context is clear. However, if not otherwise specified, diagrams should be understood to be unlabeled.

The reader may notice that our depictions of unlabeled diagrams do not distinguish the sides of each interaction line from one another by the labels ‘1’ and ‘2,’ while the definition seems to do so. This labeling should indeed not be important when we decide whether or not two unlabeled diagrams are ‘the same.’ One could have instead defined an unlabeled diagram to have each vertex equipped merely with a partition of its four half-edges into two disjoint pairs, but such a definition would be a bit cumbersome to accommodate notationally without making use of the labels ‘1’ and ‘2’ anyway later on. What is really more important is to define an equivalence relation (a notion of isomorphism) between unlabeled diagrams that only cares about the partition of the half-edge set at each vertex, not the labeling of the pairs in the partition. Of course such a notion must be introduced regardless of our choice of definition:

**Definition 5.** Two unlabeled closed Feynman diagrams  $\Gamma = (V, H_1, H_2, \Pi)$  and  $\Gamma' = (V', H'_1, H'_2, \Pi')$  are isomorphic if there exists a bijection  $\varphi : V \rightarrow V'$  and bijections  $\psi_v : H(v) \rightarrow H'(\varphi(v))$  for all  $v \in V$ , such that

1.  $\psi_v(H_1(v)) = H'_1(\varphi(v))$  or  $\psi_v(H_1(v)) = H'_2(\varphi(v))$  for all  $v \in V$ .
2. for every  $v_1, v_2 \in V$ ,  $h_1 \in H(v_1)$ ,  $h_2 \in H(v_2)$ , we have  $\{h_1, h_2\} \subset \Pi$  if and only if  $\{\psi_{v_1}(h_1), \psi_{v_2}(h_2)\} \subset \Pi'$ .

We will often denote by  $\psi$  a bijection between the *entire* half-edge sets of two diagrams. Note that the  $\psi_v$  can be obtained directly from the map  $\psi$ .

Now we defined the *labeled* closed Feynman diagrams that were introduced informally earlier, as well as an appropriate notion of isomorphism for such diagrams.

**Definition 6.** A labeled closed Feynman diagram  $\Gamma$  is specified by an unlabeled closed Feynman diagram  $(V, H_1, H_2, \Pi)$ , together with a bijection  $\mathcal{V} : V \rightarrow \{1, \dots, n\}$ , viewed as a ‘labeling’ of the vertices, as well as a bijection  $\mathcal{H}_v : H(v) \rightarrow \{i, j, k, l\}$  for every  $v \in V$  which sends  $H_1(v)$  to either  $\{i, k\}$  or  $\{j, l\}$ , where  $i, j, k, l$  are understood as symbols or distinct letters, not numbers. We will denote the collection of these bijections, viewed as labelings of the half-edges associated to each vertex, by  $\mathcal{H}$ , so in total we can view the labeled diagram  $\Gamma$  as the tuple  $\Gamma = (V, H_1, H_2, \Pi, \mathcal{V}, \mathcal{H})$ . The data  $(\mathcal{V}, \mathcal{H})$  will be called a labeling of the unlabeled diagram  $(V, H_1, H_2, \Pi)$ .

**Definition 7.** Two closed labeled Feynman diagrams  $\Gamma = (V, H_1, H_2, \Pi, \mathcal{V}, \mathcal{H})$  and  $\Gamma' = (V', H'_1, H'_2, \Pi', \mathcal{V}', \mathcal{H}')$  are isomorphic if they are isomorphic as unlabeled Feynman diagrams via maps  $\varphi$  and  $\psi_v$  as in Definition 5, which additionally satisfy

1.  $\mathcal{V}(v) = \mathcal{V}'(\varphi(v))$  for all  $v \in V$ , and
2.  $\mathcal{H}_v(h) = \mathcal{H}'_{\varphi(v)}(\psi_v(h))$  for all  $v \in V, h \in H(v)$ .

*Remark 8.* We can think of two labeled closed Feynman diagrams are isomorphic when they represent the same pairing on the set  $\{i_1, j_1, k_1, l_1, \dots, i_n, j_n, k_n, l_n\}$  of labels. In other words, the new perspective on labeled diagrams as unlabeled diagrams with extra structure is compatible with the old perspective on labeled diagrams as pairings, represented graphically by drawing  $n$  interaction lines as in Fig. 3 (b) on the page and then linking their dangling half-edges. The definition ensures that the labels  $\{i, k\}$  and  $\{j, l\}$  appear on opposite sides of the  $p$ -th interaction line in order to ensure this correspondence.

*Remark 9.* Note that there is only one possible way for two labeled diagrams to be isomorphic, since an isomorphism must send each vertex in the one to its equivalently labeled vertex in the other, and it must send all half-edges associated to a given vertex in one to the equivalently labeled half-edges associated to the corresponding vertex in the other. This completely determines maps  $\varphi$  and  $\psi_v$ , so one need only to check whether or not these maps define an isomorphism of unlabeled diagrams.

Refer again to Fig. 4 for a depiction of labeled closed diagrams. Recall that one can assign a numerical value to a labeled diagram by taking a formal product of the factors for each edge and each vertex indicated by Fig. 2 and then summing over all half-edge labels. In fact, the value so obtained is independent of the choice of labeling, hence can be associated with the underlying unlabeled diagram as well.

**Definition 10.** The numerical value associated with a labeled or unlabeled diagram  $\Gamma$  as in the preceding discussion is called the Feynman amplitude of  $\Gamma$ , denoted  $F_\Gamma$ .

For instance, Fig. 4 (a) should be interpreted as

$$\sum_{i,j,k,l} (-v_{ij}) \delta_{ik} \delta_{jl} G_{ik}^0 G_{jl}^0 = - \sum_{i,j} v_{ij} G_{ii}^0 G_{jj}^0. \quad (3.7)$$

Comparing with the first term in Eq. (3.3), we see that we are missing only the pre-constant  $\frac{1}{8}$ . In fact the factor 8 in this denominator has a significance that can be understood in terms of the structure of Feynman diagrams. It is known as the *symmetry factor* for the Feynman diagram of Fig. 4 (a).

More generally the symmetry factor of any Feynman diagram, which we shall define shortly, allows us to likewise compute the pre-constants of the associated term in our series expansion for the partition function. Roughly speaking, the symmetry factor counts the number of different labelings of a given labeled Feynman diagram

that maintain its structure. In particular, after relabeling, two connected half-edges should remain connected.

To define the symmetry factor more precisely, we first describe more carefully what is meant by a ‘relabeling.’ Consider the permutation group  $S_4$  on the four letters  $\{i, j, k, l\}$ . Denote by  $R$  the subgroup of order 8 generated by  $(i, k)$ ,  $(j, l)$ , and  $(i, j)(k, l)$ . (In fact  $R$  is isomorphic to the dihedral group of order 8.) Observe that the group  $\mathbf{R}_n := S_n \times R^n$  acts in a natural way on the set of labelings of any fixed unlabeled diagram. Here  $S_n$  acts on the permutation of  $n$  vertices, while  $R^n$  permutes the associated half-edges. In other words,  $g = (\sigma, \tau_1, \dots, \tau_n) \in S_n \times R^n$  acts on labelings by permuting the vertex labelings according to  $\sigma$  and by permuting the half-edge labelings at the  $p$ -th vertex according to  $\tau_p$ . We may think of each such  $g$  as a ‘relabeling.’

**Definition 11.** An automorphism of a labeled closed Feynman diagram  $\Gamma$  of degree  $n$  is a relabeling  $g \in \mathbf{R}_n$  such that  $g \cdot \Gamma$  is isomorphic to  $\Gamma$  (as a labeled Feynman diagram). The set of all automorphisms of  $\Gamma$  forms a subgroup  $\text{Aut}(\Gamma)$  of  $\mathbf{R}_n$ , called the automorphism group of  $\Gamma$ . The size  $|\text{Aut}(\Gamma)|$  of the automorphism group is called the symmetry factor of  $\Gamma$  and denoted  $S_\Gamma$ . (Note that  $S_\Gamma$  is independent of the labeling of  $\Gamma$ , i.e., depends only on the structure of  $\Gamma$  as an unlabeled diagram.)

*Remark 12.* Any relabeling  $g \in \mathbf{R}_n$  of  $\Gamma$  determines maps  $\varphi$  and  $\psi$  from the vertex and half-edge sets of  $\Gamma$ , respectively, to themselves. The map  $\varphi$  is obtained by mapping the vertices of  $\Gamma$  to the equivalently labeled vertices of  $g \cdot \Gamma$ , and the map  $\psi$  is obtained by mapping the half-edges associated to each vertex in  $\Gamma$  to the equivalently labeled half-edges of the equivalently labeled vertex of  $\Gamma$ . Conversely, any such maps  $\varphi$  and  $\psi$  determine a relabeling  $g \in \mathbf{R}_n$  of  $\Gamma$ . For any  $g \in \mathbf{R}_n$ , we denote the associated maps by  $\varphi_g$  and  $\psi_g$ .

Recalling Remark 9, it follows that  $g \cdot \Gamma$  and  $\Gamma$  are isomorphic as labeled diagrams (i.e.,  $g \in \text{Aut}(\Gamma)$ ) if and only if the associated maps  $\varphi_g$  and  $\psi_g$  define an isomorphism from  $\Gamma$  to itself as an *unlabeled* diagram. In other words, automorphisms, which have been defined via actions on labelings, are really just equivalent to self-isomorphisms of unlabeled diagrams. However, the perspective of labeled diagrams is valuable to retain for the application of Wick’s theorem.

For example, Fig. 7 depicts all of the automorphisms of the diagram in Fig. 4 (b), so the symmetry factor of this diagram is 4. One may readily verify that  $S_\Gamma = 8$  for the diagram in Fig. 4 (a).

These symmetry factors recover the pre-factors from our first-order expansion of the partition function. This correspondence will be established in general in Theorem 14.

Before moving on, we comment that two non-isomorphic labeled diagrams can be isomorphic as unlabeled diagrams. In this case, the numerical values associated with

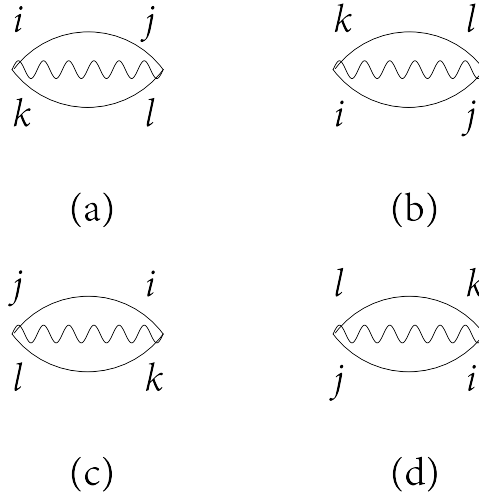


Figure 7: All automorphisms for Fig. 4 (b).

both are nonetheless the same. For instance, Fig. 4 (b) represents

$$\sum_{i,j,k,l} (-v_{ij}) \delta_{ik} \delta_{jl} G_{ij}^0 G_{kl}^0 = - \sum_{i,j} v_{ij} G_{ij}^0 G_{ij}^0,$$

while (b') represents

$$\sum_{i,j,k,l} (-v_{ij}) \delta_{ik} \delta_{jl} G_{il}^0 G_{kj}^0 = - \sum_{i,j} v_{ij} G_{ij}^0 G_{ij}^0,$$

i.e., the same term. When we ultimately sum over (isomorphism classes of) unlabeled diagrams in our series expansion for the partition function, (b) and (b') will *not* be counted as distinct diagrams. Therefore we record the following definition:

**Definition 13.** The set of (isomorphism classes of) unlabeled closed Feynman diagrams is denoted  $\mathfrak{F}_0$ .

In our new terminology, Fig. 5 and Fig. 6 depict all isomorphism classes of unlabeled closed diagrams of first and second order, respectively. Summation over the unlabeled diagrams, as opposed to the labeled diagrams, significantly simplifies the effort of bookkeeping, at the cost of computing symmetry factors for each diagram.

### 3.3 Feynman rules for $Z$

We are now ready to state and prove the so-called ‘Feynman rules’ for the diagrammatic expansion of the partition function, i.e., the recipe for producing the Taylor expansion via the enumeration of unlabeled diagrams.



**Theorem 14.** *The asymptotic series expansion for  $Z$  is given by*

$$Z = Z_0 \sum_{\Gamma \in \mathfrak{F}_0} \frac{F_\Gamma}{S_\Gamma}, \quad (3.8)$$

*i.e., the  $n$ -th term in the series of Eq. (1.5) is given by the sum of  $Z_0 \frac{F_\Gamma}{S_\Gamma}$  over isomorphism classes of unlabeled Feynman diagrams  $\Gamma$  of order  $n$ .*

*Remark 15.* We remind the reader that for a diagram  $\Gamma$  of order  $n$ , the Feynman amplitude  $F_\Gamma$  can be computed as follows:

1. Assign a dummy index to each of the  $4n$  half-edges.
2. Each edge with half-edge indices  $a, b$  yields a factor  $G_{ab}^0$ .
3. Each interaction line with half-edge indices  $a, b, c, d$  yields a factor  $-v_{ab}\delta_{ac}\delta_{bd}$ .
4. Multiply all factors obtained via steps 2 and 3, and sum over all dummy indices from 1 to  $N$ .

*Proof.* Recall Eq. (3.5), i.e., that we can write the  $n$ -th term in the series of Eq. (1.5) as

$$\frac{Z_0}{8^n n!} \sum_{i_1, j_1, k_1, l_1, \dots, i_n, j_n, k_n, l_n=1}^N \left( \prod_{m=1}^n -v_{i_m j_m} \delta_{i_m k_m} \delta_{j_m l_m} \right) \left\langle \prod_{m=1}^n x_{i_m} x_{j_m} x_{k_m} x_{l_m} \right\rangle_0.$$

By our preceding discussions (see Remark 8) this quantity can be written as

$$\frac{Z_0}{8^n n!} \sum_{\Gamma \text{ labeled, order } n} F_\Gamma.$$

We wish to replace the sum over (isomorphism classes of) labeled diagrams with a sum over unlabeled diagrams. The question is then: to any unlabeled diagram  $\Gamma$  of order  $n$ , how many distinct labeled diagrams can be obtained by labeling  $\Gamma$ ? To answer this question first assign an arbitrary labeling to obtain a labeled diagram which we shall also call  $\Gamma$ . Then the set of all labelings is the orbit of  $\Gamma$  under the group  $\mathbf{R}_n$ . By the orbit-stabilizer theorem, the size of this orbit is equal to  $|\mathbf{R}_n|/|(\mathbf{R}_n)_\Gamma|$ , where  $(\mathbf{R}_n)_\Gamma$  is the stabilizer subgroup of  $\mathbf{R}_n$  with respect to  $\Gamma$ . But this subgroup is precisely  $\text{Aut}(\Gamma)$  and  $|\mathbf{R}_n| = 8^n n!$ , so the number of distinct labeled diagrams associated with the underlying unlabeled diagram is  $\frac{8^n n!}{S_\Gamma}$ . Therefore the  $n$ -th term in the series of Eq. (1.5) is in fact

$$Z_0 \sum_{\Gamma \text{ unlabeled, order } n} \frac{F_\Gamma}{S_\Gamma},$$

as was to be shown. □

We now apply Theorem 14 to compute the second-order part of the expansion for  $Z$ . We can represent the 8 terms in the second-order part via the 8 (isomorphism classes of) unlabeled closed Feynman diagrams depicted in Fig. 6, applying Theorem 14 to compute the pre-factor of each term. The terms are organized into three groups according to the three groups of terms in Eq. (3.4). The diagrammatic approach facilitates the enumeration of these terms and allows us to classify the terms more intuitively. The first group of diagrams (a1)–(a3) in Fig. 6 are simply the diagrams obtained as ‘concatenations’ of two disconnected first-order diagrams. When computing the symmetry factor, we need to take into account the possible exchange of the two interaction lines as well as the symmetry factor of each disconnected piece as a first-order diagram. Unlike diagrams (a1) and (a2), diagram (a3) is not symmetric with respect to the exchange of the two interaction lines, so the former contribution is not included. One can readily verify the correspondence between the rest of diagrams and terms in Eq. (3.4). The distinction between the (b) and (c) diagrams will be made clear later on in our discussion of the so-called bold diagrams.

### 3.4 Comments on other interactions

We pause to make some brief comments on the development of Feynman diagrams for other interactions besides the generalized Coulomb interaction of Eq. (1.6).

First, consider an interaction of the form

$$U(x) = \frac{1}{4!} \sum_{i,j,k,l} u_{ikjl} x_i x_j x_k x_l, \quad (3.9)$$

where  $u_{ikjl}$  is a *symmetric* fourth-order tensor (i.e., invariant under any permutation of the indices). The inclusion of the factor of  $4!$  owes to the fact that the symmetry group of the interaction (i.e., the analog of  $R$ ) is now all of  $S_4$ , which is of order  $4!$ . Then the developments will be much the same, but with the role of the interaction line of Fig. 2 (b) assumed by the device shown in Fig. 8.

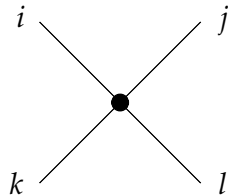


Figure 8: The interaction  $u_{ikjl}$ .

Since any fourth-order tensor can be symmetrized without changing the associated quartic form, why not just consider symmetric interactions? The reason is

that symmetrizing, e.g., the generalized Coulomb interaction throws away its lower-dimensional structure. While there are somewhat fewer diagram topologies to contend with, each vertex now involves a sum over four indices, not two. Moreover, the two-body interaction of quantum many-body physics has a natural asymmetry between creation and annihilation operators that is reflected in the structure of the Feynman diagrams for the generalized Coulomb interaction.

There is nonetheless a way to generalize the generalized Coulomb interaction without destroying its structure. Indeed, simply consider

$$U(x) = \frac{1}{8} \sum_{i,j,k,l} u_{ikjl} x_i x_j x_k x_l, \quad (3.10)$$

where  $u_{ikjl}$  is invariant under (1) the exchange  $i$  with  $k$ , (2) the exchange of  $k$  with  $l$ , and (3) the simultaneous exchange of  $i$  with  $j$  and  $k$  with  $l$ . In other words, the symmetry group of the interaction is  $R$  as for the generalized Coulomb interaction.

The developments for interactions of the form (3.10)—with the interaction line of Fig. 2 (b) now contributing the factor  $u_{ikjl}$  in the computation of Feynman amplitudes—are no different than for interactions of the form (1.6), with the exception of the GW approximation to be discussed in section 4.7.1. For the sake of writing down concrete expressions that correspond to various diagrams, we simply assume an interaction of the form (1.6).

### 3.5 Feynman rules for $\Omega$

The free energy  $\Omega$  is given by the negative logarithm of  $Z$  as in Eq. (2.8), which appears to be difficult to evaluate in terms of Feynman diagrams. It turns out that the logarithm in fact simplifies the diagrammatic expansion by removing the disconnected diagrams as in Fig. 6 (a). This is the content of Theorem 17 below, which is called the linked cluster expansion in physics literature.

Before stating the theorem, we establish some notation. Recall that a closed diagram induces an undirected graph of degree four. We say that a closed diagram is *connected* if the induced graph is connected.

**Definition 16.** The set of all connected closed diagrams is denoted  $\mathfrak{F}_0^c \subset \mathfrak{F}_0$ .

Similarly we can talk about connected components of a Feynman diagram in the obvious way. We can also consider the ‘union’  $\Gamma_1 \cup \Gamma_2$  of diagrams, i.e., the diagram constructed by viewing  $\Gamma_1$  and  $\Gamma_2$  as disconnected pieces of the same diagram. We leave more careful definitions of these notions to the reader. We establish a special notation for the union of several copies of the same diagram:

$$\Gamma^n := \bigcup_{j=1}^n \Gamma$$

A general diagram  $\Gamma \in \mathfrak{F}_0$  can be decomposed as

$$\Gamma = \bigcup_{i=1}^K \Gamma_i^{n_i}, \quad (3.11)$$

where  $\Gamma_1, \dots, \Gamma_K \in \mathfrak{F}_0^c$  are distinct.

For any diagram  $\Gamma$  expressed in the form of (3.11), the Feynman amplitude is

$$F_\Gamma = F_{\Gamma_1}^{n_1} \cdots F_{\Gamma_K}^{n_K}, \quad (3.12)$$

and since  $\Gamma_1, \dots, \Gamma_K$  are distinct diagrams, the symmetry factor is

$$S_\Gamma = (n_1! \cdots n_K!) S_{\Gamma_1}^{n_1} \cdots S_{\Gamma_K}^{n_K}. \quad (3.13)$$

It is convenient to define  $F_{\Gamma_\emptyset} = 1$  and  $S_{\Gamma_\emptyset} = 1$  for the ‘empty’ Feynman diagram  $\Gamma_\emptyset$  of order zero and moreover to let  $\Gamma^0 = \Gamma_\emptyset$  for any diagram  $\Gamma \in \mathfrak{F}_0$ .

Using this notation, we can then think of every diagram  $\Gamma \in \mathfrak{F}_0$  as being uniquely specified by a function  $n : \mathfrak{F}_0^c \rightarrow \mathbb{N}$  mapping  $\Gamma \mapsto n_\Gamma$ , where  $\mathbb{N}$  indicates the set of natural numbers including zero. We denote the set of such functions by  $\mathbb{N}^{\mathfrak{F}_0^c}$ . Indeed, any such function specifies a diagram  $\Gamma(n) := \bigcup_{\Gamma \in \mathfrak{F}_0^c} \Gamma^{n_\Gamma}$ . Moreover,  $F_{\Gamma(n)} = \prod_{\Gamma \in \mathfrak{F}_0^c} F_\Gamma^{n_\Gamma}$ , and  $S_{\Gamma(n)} = \prod_{\Gamma \in \mathfrak{F}_0^c} n_\Gamma! S_\Gamma^{n_\Gamma}$ .

Now we are ready to state and prove the diagrammatic expansion for the free energy.

**Theorem 17** (Linked cluster expansion for  $\Omega$ ). *The asymptotic series expansion for  $\Omega$  is*

$$\Omega = \Omega_0 - \sum_{\Gamma \in \mathfrak{F}_0^c} \frac{F_\Gamma}{S_\Gamma}, \quad (3.14)$$

where  $\Omega_0 = -\log Z_0$ .

*Proof.* Exponentiating both sides of Eq. (3.14) motivates the consideration of the following expression:

$$\exp \left( \sum_{\Gamma \in \mathfrak{F}_0^c} \frac{F_\Gamma}{S_\Gamma} \right) = \sum_{K=0}^{\infty} \frac{1}{K!} \left( \sum_{\Gamma \in \mathfrak{F}_0^c} \frac{F_\Gamma}{S_\Gamma} \right)^K. \quad (3.15)$$

We aim to relate this expansion to our expansion for the partition function from Theorem 14.

We will apply the multinomial theorem to compute the  $K$ -th power of the sum over  $\Gamma \in \mathfrak{F}_0^c$  appearing on the right-hand side of Eq. (3.15). This yields a sum over  $n \in \mathbb{N}^{\mathfrak{F}_0^c}$  such that  $\sum_{\Gamma \in \mathfrak{F}_0^c} n_\Gamma = K$  weighted by the multinomial coefficients  $\frac{K!}{\prod_{\Gamma \in \mathfrak{F}_0^c} (n_\Gamma)!}$ , as in

$$\exp \left( \sum_{\Gamma \in \mathfrak{F}_0^c} \frac{F_\Gamma}{S_\Gamma} \right) = \sum_{K=0}^{\infty} \frac{1}{K!} \sum_{n \in \mathbb{N}^{\mathfrak{F}_0^c} : \sum_{\Gamma} n_\Gamma = K} \frac{K!}{\prod_{\Gamma \in \mathfrak{F}_0^c} (n_\Gamma)!} \prod_{\Gamma \in \mathfrak{F}_0^c} \left( \frac{F_\Gamma}{S_\Gamma} \right)^{n_\Gamma}$$

$$\begin{aligned}
&= \sum_{K=0}^{\infty} \sum_{n \in \mathbb{N}^{\mathfrak{F}_0^c} : \sum_{\Gamma} n_{\Gamma} = K} \frac{F_{\Gamma(n)}}{S_{\Gamma(n)}} \\
&= \sum_{n \in \mathbb{N}^{\mathfrak{F}_0^c}} \frac{F_{\Gamma(n)}}{S_{\Gamma(n)}},
\end{aligned}$$

where in the penultimate step we have used our formulas for the Feynman amplitude and symmetry factor of the diagram  $\Gamma(n)$  associated to  $n \in \mathbb{N}^{\mathfrak{F}_0^c}$ . But since  $\mathbb{N}^{\mathfrak{F}_0^c}$  is in bijection with  $\mathfrak{F}_0$  via  $n \mapsto \Gamma(n)$ , we have proved:

$$\exp \left( \sum_{\Gamma \in \mathfrak{F}_0} \frac{F_{\Gamma}}{S_{\Gamma}} \right) = \sum_{\Gamma \in \mathfrak{F}_0} \frac{F_{\Gamma}}{S_{\Gamma}} = \frac{Z}{Z_0},$$

with the last equality following from Theorem 14. Taking logarithms yields the theorem.  $\square$

For example, the second-order contribution to  $\Omega$  is

$$\begin{aligned}
&\sum_{i_1, j_1, i_2, j_2} v_{i_1 j_1} v_{i_2 j_2} \left[ \left( \frac{1}{2! \cdot 8} G_{i_1 i_1}^0 G_{i_2 i_2}^0 G_{j_1 j_2}^0 G_{j_1 j_2}^0 \right. \right. \\
&\quad \left. \left. + \frac{1}{2 \cdot 2} G_{i_1 j_1}^0 G_{i_2 j_2}^0 G_{i_1 i_2}^0 G_{j_1 j_2}^0 + \frac{1}{4} G_{i_1 j_1}^0 G_{i_1 i_2}^0 G_{j_1 i_2}^0 G_{j_2 j_2}^0 \right) \right. \\
&\quad \left. + \left( \frac{1}{2! \cdot 8} G_{i_1 i_2}^0 G_{i_1 i_2}^0 G_{j_1 j_2}^0 G_{j_1 j_2}^0 + \frac{1}{2! \cdot 4} G_{i_1 i_2}^0 G_{j_1 i_2}^0 G_{i_1 j_2}^0 G_{j_1 j_2}^0 \right) \right], \quad (3.16)
\end{aligned}$$

and the terms are yielded by Fig. 6 (b), (c).

### 3.6 Feynman rules for $G$

Our next goal is to obtain a diagrammatic expansion for the Green's function  $G$ . First observe that the asymptotic series expansion for  $ZG$  can be written, similarly to that of  $Z$ , as

$$ZG_{ij} \sim \sum_{n=0}^{\infty} \frac{1}{n!} \int_{\mathbb{R}^N} x_i x_j (-U(x))^n e^{-\frac{1}{2} x^T A x} dx. \quad (3.17)$$

Again the interchange between the summation and integration is only formal. The right hand side of Eq. (3.17) can be evaluated using the Wick theorem and a new class of Feynman diagrams.

Similarly to Eq. (3.5), we see that the  $n$ -th term in the expansion of Eq. (3.17) is given by

$$\frac{Z_0}{8^n n!} \sum_{i_1, j_1, k_1, l_1, \dots, i_n, j_n, k_n, l_n=1}^N \left( \prod_{m=1}^n -v_{i_m j_m} \delta_{i_m k_m} \delta_{j_m l_m} \right) \left\langle x_i x_j \prod_{m=1}^n x_{i_m} x_{j_m} x_{k_m} x_{l_m} \right\rangle_0. \quad (3.18)$$

One can then use the Wick theorem to express this quantity as a sum over pairings of  $\mathcal{I}_{4n+2}$ , but once again it is easier to represent the pairings graphically. As before, for each  $m = 1, \dots, n$ , we draw one copy of Fig. 2 (b), i.e., an interaction line with four dangling *half-edges* labeled  $i, j, k, l$ . We can then number each interaction line as  $1, \dots, n$  and indicate this by adding an appropriate subscript to the labels  $i, j, k, l$  associated to this vertex. Now we also draw two additional freely floating half-edges with labels  $i$  and  $j$ . We can view the half-edges as terminating in a vertex indicated by a dot (which will distinguish these diagrams from the so-called ‘truncated’ diagrams that appear later on), while the other end of the half-edge is available for linking. The  $4n + 2$  half-edges  $\{i, j, i_1, \dots, l_n\}$ , each with a unique label, represent the set on which we consider pairings. We depict a pairing by linking the paired half-edges with a bare propagator. The resulting figure is a labeled Feynman diagram of order  $n$ . An example of order 2 is depicted in Fig. 9.

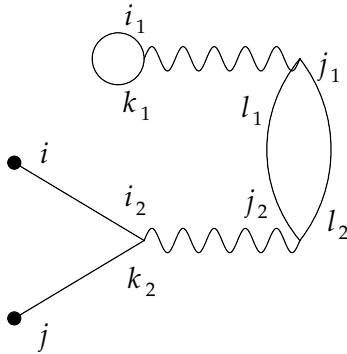


Figure 9: A labeled closed Feynman diagram of order 2.

The quantity associated via Wick’s theorem with the pairing represented by such a diagram can then be computed by taking a formal product over all propagators and interaction lines of the associated quantities indicated in Fig. 2 (a) and (b), respectively and then summing over the indices  $i_1, j_1, k_1, l_1, \dots, i_n, j_n, k_n, l_n$ . Importantly we *do not* sum over the indices  $i, j$ , as these specify the fixed entry of the Green’s function  $G_{ij}$  that we are computing via expansion. For the example depicted in Fig. 9, this procedure yields the sum

$$\sum_{i_1, j_1, k_1, l_1, i_2, j_2, k_2, l_2} v_{i_1 j_1} \delta_{i_1 k_1} \delta_{j_1 l_1} v_{i_2 j_2} \delta_{i_2 k_2} \delta_{j_2 l_2} G_{i_1 k_1}^0 G_{j_1 l_2}^0 G_{l_1 j_2}^0 G_{i i_2}^0 G_{j k_2}^0 = \sum_{i_1, j_1, i_2, j_2} v_{i_1 j_1} v_{i_2 j_2} G_{i i_2}^0 G_{j j_2}^0 G_{i_1 i_1}^0 G_{j_1 j_2}^0 G_{j_1 j_2}^0. \quad (3.19)$$

In summary, we can graphically represented the sum over pairings furnished by Wick’s theorem as a sum over such diagrams, which we call *labeled Feynman diagrams*

of order  $n$  with 2 external vertices. (Perhaps calling them diagrams with ‘external half-edges’ would be more appropriate, but ‘external vertices’ is the conventional terminology.)

One can similarly imagine the natural appearance of Feynman diagrams with  $2m$  external vertices in the expansion of the  $2m$ -point propagator  $\langle x_{p_1} \cdots x_{p_{2m}} \rangle$ .

We can define the (*partially labeled*) *Feynman diagrams of order  $n$  with  $2m$  external vertices* to be the  $\Gamma = (V, H_1, H_2, E, \Pi, \mathcal{E})$ , where  $V, H_1, H_2$  are as in the definition of closed diagrams,  $E$  is the set of  $2m$  external half-edges,  $\Pi$  is a partition of  $E \cup \bigcup_{v \in V} H(v)$  into  $2n + m$  pairs of half-edges, and  $\mathcal{E}$  is a labeling of the external half-edges only. More precisely,  $\mathcal{E}$  is a bijection from the external half-edge set  $E$  to the set of symbols  $\{p_1, \dots, p_{2m}\}$ . In the case  $m = 1$  we will instead adopt the convention  $\mathcal{E} : E \rightarrow \{i, j\}$ .

Two partially labeled diagrams of order 2 with 2 external vertices are depicted in Fig. 10. Notice that these diagrams are not isomorphic due to the distinction of the external half-edges  $i, j$ , though they would be isomorphic as ‘fully unlabeled’ diagrams.

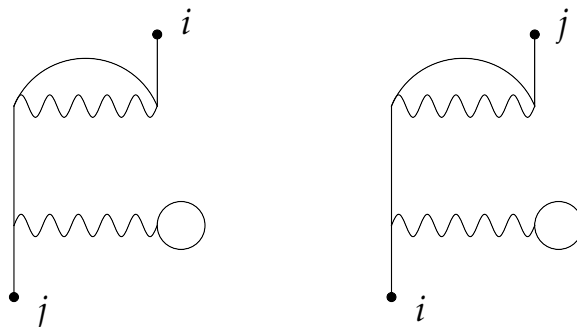


Figure 10: Non-isomorphic partially labeled diagrams of order 2 with 2 external vertices.

These diagrams can be additionally equipped with *internal* labelings  $(\mathcal{V}, \mathcal{H})$  to produce (*fully*) *labeled Feynman diagrams of order  $n$  with  $2m$  external vertices*. Here  $\mathcal{V}, \mathcal{H}$  are defined as before. More careful definitions of these classes of diagrams, as well as definitions of the notions of isomorphism for each, follow in the spirit of the analogous definitions for closed diagrams and are left to the reader.

Diagrams with external vertices will be understood to be partially labeled unless otherwise stated. The set of partially labeled diagrams (of any order) with  $2m$  external vertices is denoted  $\mathcal{F}_{2m}$ . In the case  $m = 1$  we often refer to these diagrams as *Green’s function diagrams*. Note that an unlabeled closed Feynman diagram can be viewed equivalently as a Feynman diagram with 0 external vertices.

The group  $\mathbf{R}_n$  acts naturally as before on internal labelings  $(\mathcal{V}, \mathcal{H})$  and induces a notion of automorphism for fully labeled diagrams with external vertices, as well as

a symmetry factor  $S_\Gamma$  defined to be the size of the automorphism group  $\text{Aut}(\Gamma)$  of a fully labeled diagram with external vertices (or, if  $\Gamma$  is only partially labeled, the size of the automorphism group of any full labeling of  $\Gamma$ ).

Moreover, each diagram with  $2m$  external vertices yields a Feynman amplitude which is no longer a scalar, but in fact a  $(2m)$ -tensor,  $F_\Gamma(p_1, \dots, p_{2m})$  which can be computed as follows

1. Assign a dummy index to each of the  $4n$  internal half-edges as well as indices  $p_1, \dots, p_{2m}$  to each of the external half-edges according to the labeling  $\mathcal{E}$
2. Each edge with half-edge indices  $a, b$  yields a factor  $G_{ab}^0$ .
3. Each interaction line with half-edge indices  $a, b, c, d$  yields a factor  $-v_{ab}\delta_{ac}\delta_{bd}$ .
4. Multiply all factors obtained via steps 2 and 3, and sum over all dummy indices from 1 to  $N$  to obtain a tensor in the indices  $p_1, \dots, p_{2m}$ .

For  $\Gamma \in \mathcal{F}_2$ , i.e., in the case  $m = 1$ , we usually indicate the tensor arguments by  $i, j$  as in  $F_\Gamma(i, j)$ .

Following the same discussion in section 3.3, we have

$$Z \langle x_{p_1} \cdots x_{p_{2m}} \rangle = Z_0 \sum_{\Gamma \in \mathfrak{F}_{2m}} \frac{F_\Gamma(p_1, \dots, p_{2m})}{S_\Gamma}, \quad (3.20)$$

so in particular

$$ZG_{ij} = Z_0 \sum_{\Gamma \in \mathfrak{F}_2} \frac{F_\Gamma(i, j)}{S_\Gamma}.$$

Denote by  $\mathfrak{F}_{2m}^c \subset \mathfrak{F}_{2m}$  the set of all diagrams with  $2m$  external vertices for which each connected component of the diagram contains at least one external half-edge. It is easy to see that  $\Gamma \in \mathfrak{F}_{2m}^c$  may have more than one connected component when  $m > 1$ . However, when  $m = 1$ , any diagram  $\Gamma \in \mathfrak{F}_2^c$  has only two external half-edges. Each internal vertex has 4 half-edges, and each connected component must contain an even number of half-edges. This implies that  $\Gamma$  must contain only one connected component, so  $\mathfrak{F}_2^c$  is in fact the subset of diagrams in  $\mathfrak{F}_2$  that are connected.

Theorem 18 below shows, perhaps surprisingly, that the expansion for the *correlator*  $\langle x_{p_1} \cdots x_{p_{2m}} \rangle$  removes many diagrams, and is therefore *simpler* than the expansion of  $Z \langle x_{p_1} \cdots x_{p_{2m}} \rangle$ . The combinatorial argument is similar in flavor to that of the proof of Theorem 17.

**Theorem 18** (Linked cluster expansion for correlators). *The asymptotic series expansion for  $\langle x_{p_1} \cdots x_{p_{2m}} \rangle$ , where  $1 \leq p_1, \dots, p_{2m} \leq N$ , is*

$$\langle x_{p_1} \cdots x_{p_{2m}} \rangle = \sum_{\Gamma \in \mathfrak{F}_{2m}^c} \frac{F_\Gamma(p_1, \dots, p_{2m})}{S_\Gamma}. \quad (3.21)$$



In particular, the series for  $G$  is

$$G_{ij} = \sum_{\Gamma \in \mathfrak{F}_2^c} \frac{F_\Gamma(i, j)}{S_\Gamma}. \quad (3.22)$$

*Proof.* Any diagram  $\Gamma \in \mathfrak{F}_{2m}$  can be decomposed uniquely as  $\Gamma = \Gamma' \cup \Gamma''$ , where  $\Gamma' \in \mathfrak{F}_{2m}^c$  and  $\Gamma'' \in \mathfrak{F}_0$ , and we allow  $\Gamma''$  to be the empty diagram. Hence according to Eq. (3.20),

$$Z \langle x_{p_1} \cdots x_{p_{2m}} \rangle = Z_0 \sum_{\Gamma' \in \mathfrak{F}_{2m}^c} \sum_{\Gamma'' \in \mathfrak{F}_0} \frac{F_{\Gamma' \cup \Gamma''}(p_1, \dots, p_{2m})}{S_{\Gamma' \cup \Gamma''}}.$$

Now for  $\Gamma' \in \mathfrak{F}_{2m}^c$  and  $\Gamma'' \in \mathfrak{F}_0$ ,

$$F_{\Gamma' \cup \Gamma''}(p_1, \dots, p_{2m}) = F_{\Gamma'}(p_1, \dots, p_{2m}) F_{\Gamma''}.$$

Also  $\Gamma'$  and  $\Gamma''$  have different numbers of external vertices, so  $\text{Aut}(\Gamma' \cup \Gamma'') = \text{Aut}(\Gamma') \times \text{Aut}(\Gamma'')$ , and consequently

$$S_\Gamma = S_{\Gamma'} S_{\Gamma''}.$$

Hence

$$\begin{aligned} Z \langle x_{p_1} \cdots x_{p_{2m}} \rangle &= Z_0 \sum_{\Gamma' \in \mathfrak{F}_{2m}^c} \sum_{\Gamma'' \in \mathfrak{F}_0} \frac{F_{\Gamma'}(p_1, \dots, p_{2m}) F_{\Gamma''}}{S_{\Gamma'} S_{\Gamma''}} \\ &= Z_0 \left( \sum_{\Gamma' \in \mathfrak{F}_{2m}^c} \frac{F_{\Gamma'}(p_1, \dots, p_{2m})}{S_{\Gamma'}} \right) \left( \sum_{\Gamma'' \in \mathfrak{F}_0} \frac{F_{\Gamma''}}{S_{\Gamma''}} \right) \\ &= Z \sum_{\Gamma' \in \mathfrak{F}_{2m}^c} \frac{F_{\Gamma'}(p_1, \dots, p_{2m})}{S_{\Gamma'}}, \end{aligned}$$

where the last equality follows from Theorem 14. Dividing by  $Z$  completes the proof.  $\square$

We now discuss the first few terms of the expansion for the Green's function  $G$ . The zeroth order expansion for  $G_{ij}$  is  $G_{ij}^0$ . Fig. 11 depicts the Feynman diagrams for the first-order contribution to  $G_{ij}$ , which amounts to the expression

$$-\frac{1}{2} \sum_{k,l} (v_{kl} G_{ik}^0 G_{jk}^0 G_{ll}^0) - \sum_{k,l} (v_{kl} G_{ik}^0 G_{jl}^0 G_{kl}^0). \quad (3.23)$$

Note how the symmetry factor of these diagrams is affected by the labeling of the external half-edges.

Fig. 12 depicts the Feynman diagrams for the second-order contribution to  $G_{ij}$ . These diagrams can be systematically obtained from the free energy diagrams of Fig. 6 (b) and (c) by cutting a propagator line to yield two external half-edges and then listing the non-isomorphic ways of labeling of these external half-edges. Note that all terms contain only one connected component due to Theorem 18. For simplicity we omit the resulting formula for the second-order contribution to  $G_{ij}$ .

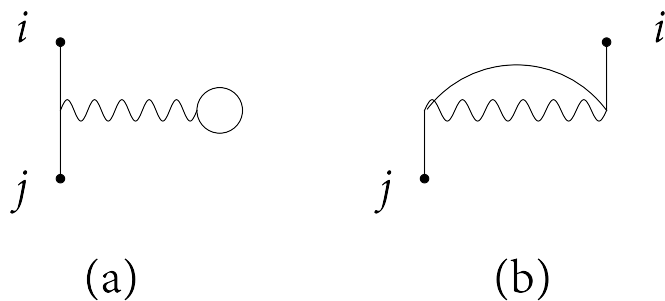


Figure 11: First-order expansion for  $G_{ij}$ .

### 3.7 Why we do not use fully unlabeled Green's function diagrams

Why not reduce the redundancy of diagrams by considering a notion of fully unlabeled Green's function diagrams? One reason is that the notion of symmetry factor would be different, yielding an unpleasant extra factor in Theorem 18. Moreover, in the development of the bold diagrammatic expansion of section 4, we will consider an operation in which propagator lines are replaced by Green's function diagrams. Since different orientations of such an 'insertion' might yield different topologies of the resulting diagram, it is good to keep track of non-isomorphic external labelings separately.

Finally, by retaining an external labeling, there is a clearer interpretation of each diagram as a matrix yielded by contracting out internal indices. Note carefully, however, that the Feynman amplitude of a non-symmetric diagram, i.e., a diagram whose isomorphism class is changed by a relabeling of the external vertices, is in general a non-symmetric matrix. By contrast,  $G$  is symmetric. Therefore any reasonable truncation of the expansion of Theorem 18 should not include any non-symmetric diagram without also including all diagrams obtained by different external labelings.

### 3.8 Feynman rules for $\Sigma$

The computation of  $G$  by diagrammatic methods can be further simplified via the introduction of the notion of the *self-energy*. This notion can be motivated diagrammatically as follows. Observe that diagrams such as (b1'), (b2'), and (b3'') in Fig. 12 are 'redundant' in that they can be constructed by 'stitching' first-order diagrams together at external vertices. Such diagrams will be removed in the diagrammatic expansion for the self-energy matrix  $\Sigma$ , defined as the difference between the inverse of  $G$  and that of  $G^0$  as

$$\Sigma = (G^0)^{-1} - G^{-1}. \quad (3.24)$$

Observe that once  $\Sigma$  is known,  $G$  can be computed simply via Eq. (3.24).

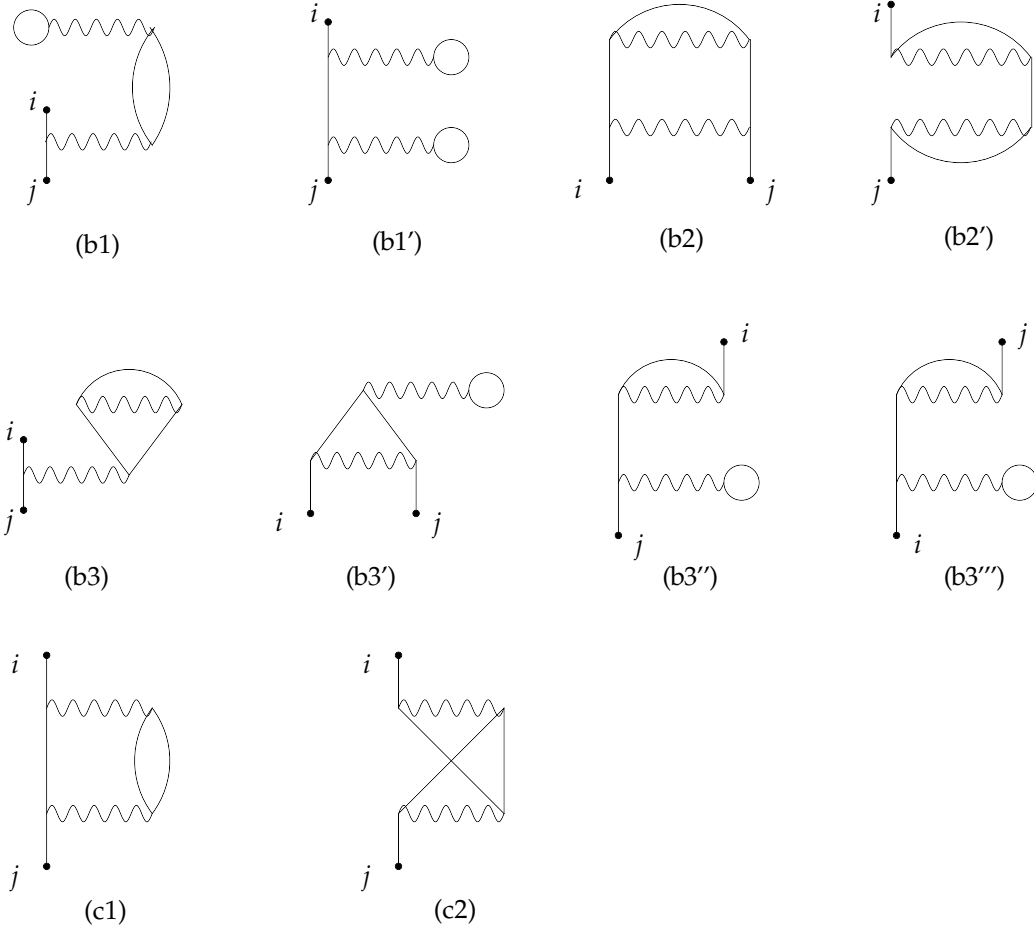


Figure 12: Second-order expansion for  $G_{ij}$ . The lettering is obtained from that of the free energy diagrams in Fig. 6 (b),(c), from which the Green's function diagrams may be obtained by cutting lines.

However, Eq. (3.24) does not clarify the diagrammatic motivation for the self-energy. Note that the definition of the self-energy matrix in Eq. (3.24) is equivalent to

$$G = G^0 + G^0 \Sigma G, \quad (3.25)$$

which is called the *Dyson equation*. By plugging the formula for  $G$  specified by the Dyson equation back into the right-hand side of Eq. (3.25) and then repeating this procedure *ad infinitum*, one obtains the formal equation

$$G = G^0 + G^0 \Sigma G^0 + G^0 \Sigma G^0 \Sigma G^0 + \dots, \quad (3.26)$$

which suggests a diagrammatic interpretation for  $\Sigma$ . To wit, in order to avoid counting the same Green's function diagram twice in the right-hand side of Eq. (3.26),  $\Sigma$  should only include those diagrams that cannot be separated into two disconnected

components when removing one bare propagator line. In physics terminology, these are called the *one-particle irreducible* (1PI) diagrams.

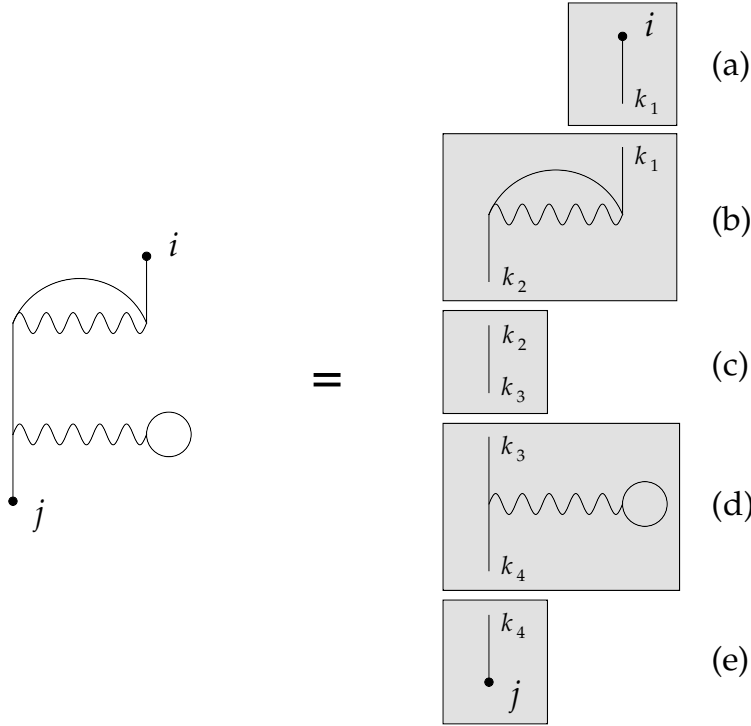


Figure 13: Decomposing a Green's function diagram into truncated 1PI diagrams and bare propagators.

We must be careful about what exactly is meant by such a diagram. We want to be able to produce Green's function diagrams by stitching together 1PI diagrams via a bare propagator line  $G^0$ , as depicted in Fig. 13. In order to avoid double-counting the propagators at each 'stitch,' our self-energy diagrams should *not* include a contribution from the propagator where the stitch is made. In the example shown in Fig. 13, we write the matrix represented by the diagram on the left-hand side as a product (a)(b)(c)(d)(e) of matrices represented by the diagrams on the right-hand side. Here (a), (c), and (e) simply represent the propagator  $G^0$ . Diagrams (b) and (d) are the self-energy diagrams representing the matrices with  $(k_1, k_2)$  entry given by  $v_{k_1 k_2} G^0_{k_1 k_2}$  and  $(k_3, k_4)$  entry given by  $\delta_{k_3 k_4} v_{k_3 k_3} G^0_{k_3 k_3}$ , respectively. Since these diagrams are like Green's function diagrams, except missing the external propagator contributions, we refer to them as *truncated Green's function diagrams*.

**Definition 19.** A truncated Green's function diagram  $\Gamma$  is obtained from a Green's function diagram  $\Gamma'$ . The internal half-edges of  $\Gamma'$  paired with the external half-edges of  $\Gamma'$  labeled  $i$  and  $j$  are referred to as the external half-edges of  $\Gamma$  and are labeled  $i$

and  $j$ , respectively. The 1PI diagrams are the truncated Green's function diagrams that cannot be disconnected by the removal of a single bare propagator line. The set of all truncated Green's function diagrams is denoted by  $\mathfrak{F}_2^{c,t}$ , and the set of all 1PI diagrams is denoted by  $\mathfrak{F}_2^{1PI}$ . The diagrams in  $\mathfrak{F}_2^{1PI}$  are alternatively referred to as self-energy diagrams.

Analogously one can define  $\mathfrak{F}_{2m}^{c,t}$  and  $\mathfrak{F}_{2m}^{1PI}$  for  $m > 1$ , but we will not make use of such notions.

As a data structure, a truncated Green's function diagram is really equivalent to its 'parent' Green's function diagram, but the interpretation is different, and we visually distinguish the truncated diagrams from their counterparts by removing the dot at the external vertex. In addition, a truncated Green's function diagram has a different notion of (matrix-valued) Feynman amplitude  $F_\Gamma(i, j)$ , computed as follows:

1. Assign a dummy index to each of the  $4n - 2$  internal half-edges as well as indices  $i, j$  to each of the external half-edges according to the labeling furnished by Definition 19.
2. Each *internal* edge with half-edge indices  $a, b$  yields a factor  $G_{ab}^0$ .
3. Each interaction line with half-edge indices  $a, b, c, d$  yields a factor  $-v_{ab}\delta_{ac}\delta_{bd}$ .
4. Multiply all factors obtained via steps 2 and 3, and sum over all dummy indices from 1 to  $N$  to obtain a matrix in the indices  $i, j$ .

However, the symmetry factor  $S_\Gamma$  of a truncated Green's function diagram is unchanged from that of the underlying Green's function diagram.

We may further introduce the concept of two-particle irreducible (2PI) Green's function diagrams as the subset of diagrams in  $\mathfrak{F}_2^{1PI}$  that cannot be disconnected by the removal of any two edges. The set of all such diagrams is denoted by  $\mathfrak{F}_2^{2PI}$ . The 2PI diagrams will be used to define the bold diagrams in section 4.

The first-order self-energy diagrams are depicted in Fig. 14. The only difference from Fig. 11 is that the external vertices are removed to produce truncated diagrams. The second-order self-energy diagrams are shown in Fig. 14. Note that the Green's function diagrams (b1'), (b2'), (b3''), and (b3''') in Fig. 12—after removing the external vertices to yield self-energy diagrams—are not 1PI, hence not self-energy diagrams.

**Theorem 20.** *The asymptotic series expansion for  $\Sigma_{ij}$ , where  $1 \leq i, j \leq N$ , is*

$$\Sigma_{ij} = \sum_{\Gamma \in \mathfrak{F}_2^{1PI}} \frac{F_\Gamma(i, j)}{S_\Gamma}. \quad (3.27)$$

*Proof.* One can think of the Dyson equation (3.25) as an equation of formal power series, where  $G$  and  $\Sigma$  indicate the asymptotic series expansions of  $G$  and  $\Sigma$ . (Recall

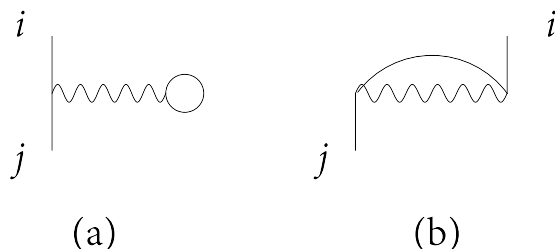


Figure 14: First-order diagrams for  $\Sigma_{ij}$ .

that we may think of our diagrammatic expansions as power series in a parameter  $\lambda$  that scales the interaction strength. It is not hard to see directly from the definition of  $\Sigma$ , as for all other quantities we consider, that an asymptotic series in  $\lambda$  exists in the first place.) Now the series for  $G$  is known from Theorem 18, and we claim that the series for  $\Sigma$  is the unique formal power series satisfying Eq. (3.25) as an equation of formal power series. Indeed, if some power series  $\Sigma$  satisfies Eq. (3.25), then  $\Sigma$  satisfies Eq. (3.24) as well (as an equation of formal power series), but inverses are unique in the ring of formal power series, so  $\Sigma$  satisfying Eq. (3.24) is uniquely determined.

Thus all we need to show is that Eq. (3.25) holds when we plug in the series for  $\Sigma$  from Eq. (3.27). To this end, write, via Theorem 18,

$$G = \sum_{\Gamma \in \mathfrak{F}_2^c} \frac{F_\Gamma}{S_\Gamma}, \quad (3.28)$$

where  $F_\Gamma$  appearing in the summand is a matrix. Now every  $\Gamma \in \mathfrak{F}_2^c$  that is of order greater than 1 can be decomposed uniquely into a bare propagator line at the external half-edge labeled  $i$ , a self-energy diagram  $\Gamma'$  connected to this propagator line at one external half-edge, and another Green's function diagram  $\Gamma''$  connected to  $\Gamma'$  at its other external half-edge. (This fact should be clear graphically, though a more careful proof is left to the reader.) For example, in Fig. 13,  $\Gamma'$  corresponds to (b) and  $\Gamma''$  corresponds to (c)(d)(e).

Moreover, we have the equality (of matrices)

$$F_\Gamma = G^0 F_{\Gamma'} F_{\Gamma''}.$$

Also, due to the fact that  $\Gamma$  distinguishes the labels  $i, j$ , we have that  $S_\Gamma = S_{\Gamma'} S_{\Gamma''}$ . Indeed, any automorphism of (a fully labeled version of)  $\Gamma$  must fix the label  $i$  and  $j$  of the external half-edges, as well as the labels of the internal half-edges connected directly to them. Then such an automorphism can only permute labels within the component  $\Gamma'$ ; otherwise, the automorphism would induce a graph automorphism of  $\Gamma'$  with another subgraph of  $\Gamma$  containing the external half-edge labeled by  $i$  as well

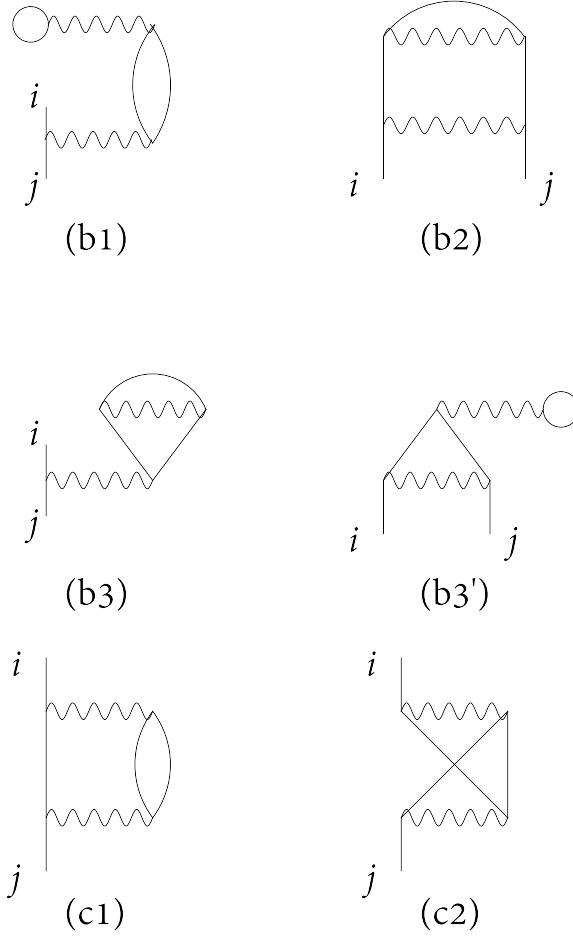


Figure 15: Second-order diagrams for  $\Sigma_{ij}$ . The labels correspond to those of the ‘parent’ Green’s function diagrams in Fig. 12.

as some half-edge in  $\Gamma''$ , which would consequently fail to be one-particle irreducible, contradicting the one-particle irreducibility of  $\Gamma'$ .

Thus from Eq. (3.28) we obtain the equality of power series

$$G = G^0 + G^0 \sum_{\Gamma' \in \mathfrak{F}_2^{1PI}} \sum_{\Gamma'' \in \mathfrak{F}_2^{\circ}} \frac{F_{\Gamma'} F_{\Gamma''}}{S_{\Gamma'} S_{\Gamma''}} = G^0 + G^0 \left[ \sum_{\Gamma' \in \mathfrak{F}_2^{1PI}} \frac{F_{\Gamma'}}{S_{\Gamma'}} \right] G,$$

as was to be shown. □

## 4 Bold diagrams

It turns out further redundancy can be removed from the diagrammatic series for the self-energy by consideration of the so-called bold diagrams. Note that so far all dia-

grammatic series are defined using the non-interacting Green's function  $G^0$  (alternatively the bare propagator), which can be viewed as the non-interacting counterpart to the *interacting* Green's function  $G$  (alternatively the 'dressed' or 'renormalized' propagator). What if we replace all of the  $G^0$  in our self-energy expansion by  $G$ ? Accordingly let us introduce the convention of a *doubled line* (also called a *bold line*) to denote  $G$ . After replacing all thin lines by bold lines in a diagram, the resulting diagram is called a *bold diagram*. (Topologically the diagram is not altered by this procedure, but the interpretation and Feynman amplitude, as well as our visual representation of the diagram, are changed.) A bold diagram can be understood as a shorthand for an infinite sum of bare diagrams by swapping each bold line out for the bare diagrammatic expansion of  $G$ . An example of a bold diagram and its representation as a sum of bare diagrams is provided in Fig. 16. Note that this representation is considered as an equality *only at the level of formal power series*.

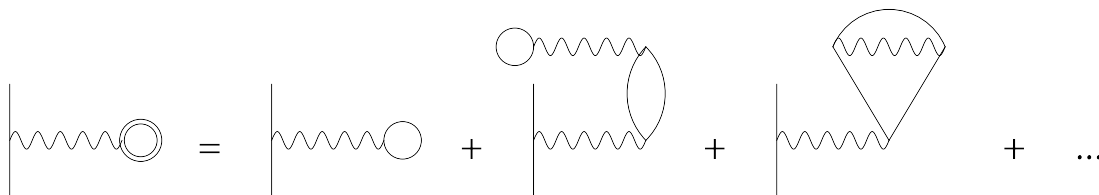


Figure 16: A bold self-energy diagram (the dumbbell, or Hartree, diagram), together with its expansion as a series of bare diagrams. Here we omit the labels  $i, j$  in the self-energy diagram, leaving a dangling half-edge (without a dot) to indicate their existence.

If one were to replace all self-energy diagrams by their bold counterparts, the resulting bold diagram expansion would overcount many of the original bare self-energy diagrams. Indeed, notice that the second and third terms on the right-hand side of Fig. 16 account for the bare self-energy diagrams (b1) and (b3) of Fig. 15. The bold versions of (b1) and (b3) would also count these terms, so as a result these contributions would be double-counted. Therefore if we can concoct a successful bold diagrammatic expansion for the self-energy, it should involve 'fewer' diagrams than the bare expansion. From a certain perspective, passage to the bold diagrams can then be thought of as a means to further economize on diagrammatic bookkeeping.

Which self-energy diagrams should be left out of the bold expansion? Notice that the disqualifying feature of diagrams (b1) and (b3) of Fig. 15 is that they contain Green's function diagram insertions—for short, simply *Green's function insertions* or even *insertions* when the context is clear. In other words, we can disconnect each of these diagrams into two separate diagrams by cutting two propagator lines. The resulting component *not* containing the external half-edges of the original diagram is itself a Green's function diagram with external half-edges at the cut locations.



(For now we are being a bit casual about the distinction between truncated and non-truncated diagrams because there is essentially no topological difference.) In the component that does contain the external half-edges of the original diagram, the two half-edges that have been left dangling due to the cuts can be sewn together with a bold line to yield the ‘parent’ bold diagram. The insertion procedure yielding diagram (b3) Fig. 15 is depicted in Fig. 17.

In general a bare self-energy diagram may contain many such Green’s function insertions, possibly viewed as being nested within one another. However, it will soon pay to introduce a notion of a maximal Green’s function insertion, or *maximal insertion* for short. This is a Green’s function insertion that is not contained within any other insertion. Then we will find that any bare self-energy diagram can be represented uniquely via its set of maximal insertions.

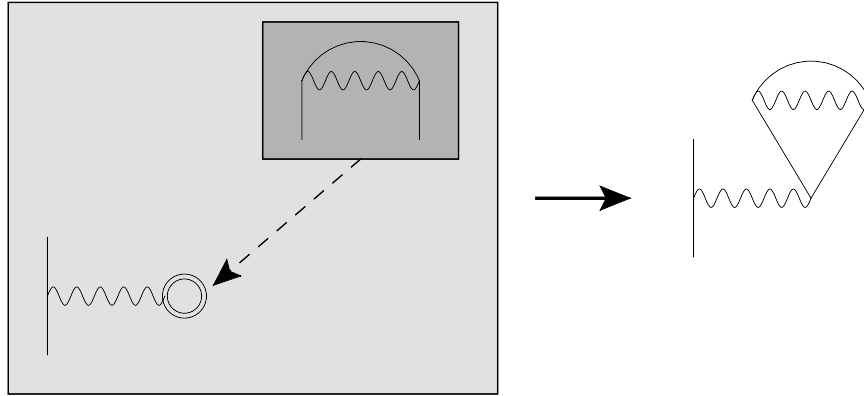


Figure 17: Green’s function insertion yielding diagram (b3) of Fig. 15.

Note that a diagram admits a Green’s function insertion if and only if it can be disconnected by removing two propagator lines. Then the candidates for the bold self-energy diagrams are the self-energy diagrams with Green’s function insertions; namely, the 2PI self-energy diagrams introduced earlier, though now considered with bold lines. This set will be denoted  $\mathfrak{F}_2^{2\text{PI}}$  as before. We distinguish diagrams a bold via the notation for the Feynman amplitude as  $\mathbf{F}_\Gamma$ , as opposed to  $F_\Gamma$ . We call diagrams in  $\mathfrak{F}_2^{2\text{PI}}$  *skeleton diagrams* and diagrams in  $\mathfrak{F}_2^{1\text{PI}} \setminus \mathfrak{F}_2^{2\text{PI}}$  *non-skeleton self-energy diagrams*, or simply *non-skeleton diagrams* for short.

The idea of the bold diagrammatic expansion is to write

$$\Sigma_{ij} = \sum_{\Gamma_s \in \mathfrak{F}_2^{2\text{PI}}} \frac{\mathbf{F}_{\Gamma_s}(i, j)}{S_{\Gamma_s}}. \quad (4.1)$$

This equation must be interpreted rather carefully to yield a rigorous statement. However, formally speaking for now, note that in order for the bold diagram expansion (4.1) to match the bare diagram expansion (3.27) for the self-energy,  $S_{\Gamma_s}^{-1}$  should

be the right guess of the pre-factor for the diagram  $\Gamma_s$  in (4.1). Indeed, if we formally substitute the bare expansion for the Green's function in for each bold propagator line of a bold diagram  $\Gamma_s$ , then the first term in the resulting expansion for the bold Feynman amplitude of  $\Gamma_s$  will be the Feynman amplitude of  $\Gamma_s$  interpreted as a *bare* diagram, which should indeed be counted with the pre-factor  $S_{\Gamma_s}^{-1}$  as in the bare expansion for the self-energy. But then we have to establish that the rest of the bare self-energy diagrams (i.e., those with non-skeleton topology) are counted with the appropriate pre-factors. This turns out to be non-trivial and constitutes the major task of this section. Our efforts culminate in Theorem 32 of section 4.6 below, in which we give precise meaning to and prove Eq. (4.1). We recommend readers to skip to section 4.6 for the applications of the bold diagrammatic expansion first and then to return to the intervening details later.

## 4.1 Skeleton decomposition

Our first goal is to show that every self-energy diagram can be decomposed (uniquely, in some sense) as a skeleton diagram with Green's function insertions. We now turn to defining the notion of insertion more carefully.

**Definition 21.** Given a truncated Green's function diagram  $\Gamma$ , together with a half-edge pair  $\{h_1, h_2\}$  in  $\Gamma^5$  and another truncated Green's function diagram  $\Gamma'$ , the insertion of  $\Gamma'$  into  $\Gamma$  at  $(h_1, h_2)$ , denoted  $\Gamma \oplus_{(h_1, h_2)} \Gamma'$  is defined to be the truncated Green's function diagram constructed by taking the collection of all vertices and half-edges (along with their pairings) from  $\Gamma$  and  $\Gamma'$ , then defining a new half-edge pairing by removing  $\{h_1, h_2\}$  and adding  $\{h_1, e_1\}$  and  $\{h_2, e_2\}$ , where  $e_1$  and  $e_2$  are the external half-edges of  $\Gamma'$  labeled  $i$  and  $j$ , respectively.

Notice that the ordering of  $(h_1, h_2)$  in  $\Gamma \oplus_{(h_1, h_2)} \Gamma'$  matters in this definition because it determines the orientation of the inserted diagram. Here the definition has also made use of the fact that truncated Green's function diagrams distinguish their external half-edges via the labels  $i$  and  $j$ .

We can define a simultaneous insertion of truncated Green's function diagrams  $\Gamma^{(1)}, \dots, \Gamma^{(K)}$  along several edges of a diagram, as follows:

**Definition 22.** Let  $\Gamma$  be a truncated Green's function diagram, and consider a collection of distinct half-edge pairs  $\{h_1^{(k)}, h_2^{(k)}\}$  for  $k = 1, \dots, K$ . Let  $\Gamma_0 = \Gamma$  and recursively define  $\Gamma_{k+1} := \Gamma_k \oplus_{(h_1^{(k+1)}, h_2^{(k+1)})} \Gamma^{(k+1)}$  for  $k = 0, \dots, K-1$ . Then the resulting  $\Gamma_K$  is the insertion of  $\Gamma^{(1)}, \dots, \Gamma^{(K)}$  into  $\Gamma$  along  $(h_1^{(1)}, h_2^{(1)}), \dots, (h_1^{(K)}, h_2^{(K)})$ , denoted

$$\Gamma \oplus_{(h_1^{(1)}, h_2^{(1)}), \dots, (h_1^{(K)}, h_2^{(K)})} [\Gamma^{(1)}, \dots, \Gamma^{(K)}].$$

---

<sup>5</sup>So  $\{h_1, h_2\}$  is contained in the pairing  $\Pi_\Gamma$  of half-edges associated with  $\Gamma$ .

Notice that the simultaneous insertion does not depend on the ordering of the  $k$  half-edge pairs, though it does depend in general on the ordering of the half-edges within each pair.

**Definition 23.** We say that a truncated Green's function diagram  $\Gamma$  admits an insertion  $\Gamma''$  at  $(h_1, h_2)$  if it can be written as  $\Gamma' \oplus_{(h_1, h_2)} \Gamma''$ , where  $\{h_1, h_2\}$  is a pair in  $\Gamma'$  and  $\Gamma''$  is a nonempty truncated Green's function diagram. (Note that  $\Gamma$  admits such an insertion if and only if  $\Gamma$  can be disconnected by removing the half-edges  $h_1$  and  $h_2$ .) We say that this insertion is maximal if  $\Gamma'$  does not in turn admit an insertion containing either of the half-edges  $h_1, h_2$ .

For example, consider in self-energy diagram of Fig. 18, which admits two maximal insertions, shown in blue and red, respectively. (Note that each of the maximal insertions admits insertions itself, i.e., the overall diagram admits several insertions that are not maximal.) The remaining half-edges and interaction lines in the diagram (shown in black) form the 'skeleton' of the diagram.

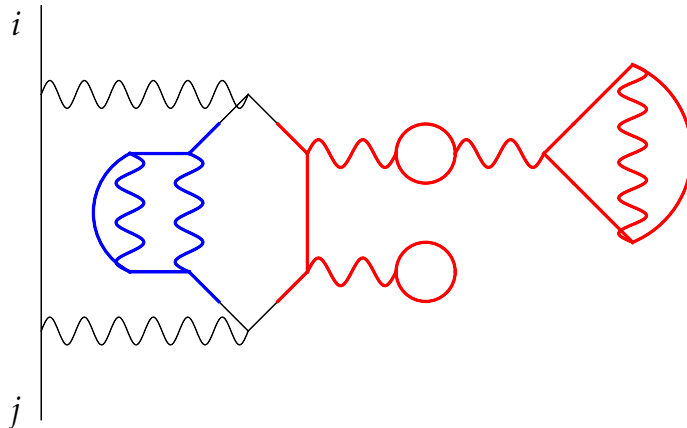


Figure 18: A self-energy diagram with two maximal insertions, depicted in blue and red, respectively.

The following result characterizes every self-energy diagram uniquely in terms of its maximal insertions and an underlying skeleton diagram (hence the name 'skeleton') obtained by collapsing each of these insertions into a single propagator line. The proof is given in Appendix B. It is somewhat technical and may be skipped on first reading to avoid interrupting the flow of the developments that follow.

**Proposition 24** (Skeleton decomposition). *Any diagram  $\Gamma \in \mathfrak{F}_2^{1\text{PI}}$  can be written as*

$$\Gamma = \Gamma_s \oplus_{(h_1^{(1)}, h_2^{(1)}, \dots, h_1^{(K)}, h_2^{(K)})} [\Gamma^{(1)}, \dots, \Gamma^{(K)}], \quad (4.2)$$

where  $\Gamma_s \in \mathfrak{F}_2^{2\text{PI}}$  and  $\Gamma^{(k)} \in \mathfrak{F}_2^{\text{c,t}}$  for  $k = 1, \dots, K$ . Moreover such a decomposition is unique up to the external labeling of the  $\Gamma^{(k)}$  and the ordering of the pair  $(h_1^{(k)}, h_2^{(k)})$  for

each fixed  $k$ , and the  $\Gamma^{(k)}$  are precisely the maximal insertions admitted by  $\Gamma$  (ignoring distinction of insertions based on external labelings).

*Remark 25.* Here we record some comments on the meaning of the uniqueness result of Proposition 24. It is purely an artifact of our ‘ $\oplus$ ’ notation for insertions, which privileges an ordering of each pair  $\{h_1^{(k)}, h_2^{(k)}\}$  of half-edges as in (4.2), that one could just as well write  $\Gamma$  in the form of (4.2) by exchanging the roles of  $h_1^{(k)}$  and  $h_2^{(k)}$  and permuting the external labels of the insertion  $\Gamma^{(k)}$ . The statement is then that the decomposition in Proposition 24 is unique up to this redundancy, which is resolved by fixing an external labeling for each of the maximal insertions admitted by  $\Gamma$ .

This sort of non-uniqueness (which is really just a notational artifact and reflects no interesting topological properties of a diagram), should be contrasted with a notion appearing later on, which is to be motivated in section 4.2 and fully sharpened in section 4.3. Indeed, we will be interested in the number of ‘ways’ (in a sense to be clarified) of producing a diagram *isomorphic* to  $\Gamma \in \mathfrak{F}_2^{1\text{PI}}$  from its skeleton  $\Gamma_s \in \mathfrak{F}_2^{2\text{PI}}$  via Green’s function insertions. By contrast, Proposition 24 above concerns the number of ways of producing the *actual* diagram  $\Gamma$  from its skeleton  $\Gamma_s$ , stating that that there is in fact only one (up to the notational ambiguity we have discussed).

Now we return to the task of developing a bold diagrammatic expansion for the self-energy. Proposition 24 tells us that each bare self-energy diagram can be constructed from a unique skeleton diagram via Green’s function insertions. It is not hard to see that, conversely, the result of making insertions into a skeleton diagram is a 1PI diagram, i.e., a self-energy diagram. If we view skeleton diagrams as *bold* diagrams, this implies that by summing over all (bold) skeleton diagrams (and then formally replacing each bold line with a sum over Green’s function diagrams), we recover all of the bare self-energy diagrams. However, there remains the question of whether these diagrams are counted appropriately. To answer this question we need to understand three items: (1) how many ways a given (isomorphism class of) self-energy diagram can be obtained via insertions from its underlying skeleton, (2) how to represent the automorphism groups (hence also symmetry factors) of self-energy diagrams in terms of the decomposition of Proposition 24, and (3) the relation between items (1) and (2). These items will be addressed in Sections 4.3, 4.4, and 4.5, respectively. First, however, to gain familiarity with what we are trying to prove, we discuss some motivating examples in section 4.2.

## 4.2 Motivating examples

Consider the non-skeleton diagram  $\Gamma$  in Fig. 19 (a), for which we have  $S_\Gamma = 2$ . It can be uniquely decomposed into a skeleton diagram  $\Gamma_s$  in Fig. 19 (b) and the single maximal insertion in  $S_{\Gamma_s}$  shown in (c). Evidently  $S_{\Gamma_s} = 2$  and  $S_{\Gamma_g} = 1$ . Roughly speaking (for now), there is only one ‘way’ in which (c) can be inserted into (b) to

produce a diagram isomorphic to (a), so we say that the *redundancy factor* of  $\Gamma$  is 1 and write  $r_\Gamma = 1$ . This notion will be defined more carefully below. For now we mention that we *do not* count separately the oppositely ‘oriented’ insertions of  $S_{\Gamma_g}$  into  $S_{\Gamma_s}$  because  $\Gamma_g$  is *symmetric*, i.e., its isomorphism class is unchanged by the exchange of its two external labels. Note, e.g., that all diagrams in Fig. 12 are symmetric except (b3'') and (b3''').

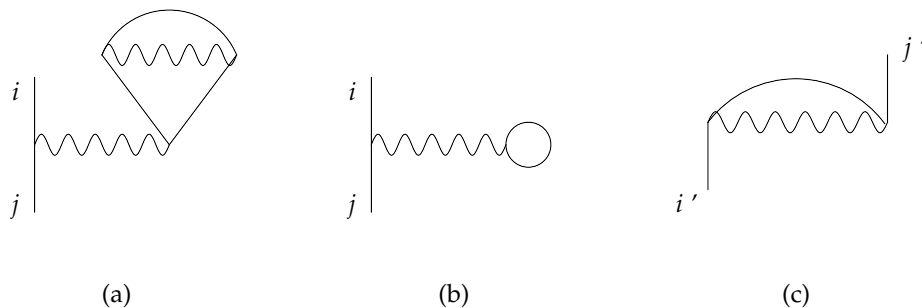


Figure 19: Decomposition of a non-skeleton diagram  $\Gamma$  in (a) into a skeleton diagram  $\Gamma_s$  in (b) and a truncated Green’s function diagram  $\Gamma_g$  in (c).

In our proposed bold diagram expansion Eq. (4.1) for the self-energy, when we substitute the bare diagrammatic expansion for the Green’s function in for each bold line, every bare self-energy diagram will be accounted for once for each of the  $r_\Gamma$  ‘ways’ that it can be produced from its skeleton via insertions, each time with a pre-factor equal to the reciprocal of the product of the symmetry factors of its skeleton and of all of its insertions. In other words, if  $\Gamma \in \mathfrak{F}_2^{\text{PI}}$  is decomposed as in Eq. (4.2), then  $\Gamma$  will be accounted for with a pre-factor of

$$\frac{r_\Gamma}{S_{\Gamma_s} \cdot \prod_{k=1}^K S_{\Gamma^{(k)}}}.$$

Then our hope is that

$$S_\Gamma = \frac{S_{\Gamma_s} \cdot \prod_{k=1}^K S_{\Gamma^{(k)}}}{r_\Gamma}. \quad (4.3)$$

This is a key reason for justifying bold diagrams, and is the content of Corollary 31 below. We can see from the above discussion that it holds in the case of Fig. 19.

For now we check Eq. (4.3) in a few more cases as we further develop the notion of the redundancy factor.

Consider Fig. 20. Here the insertion  $\Gamma_g$  is not symmetric, so we count  $r_\Gamma = 2$  different ways of inserting it into the skeleton  $\Gamma_s$  to make a diagram isomorphic to  $\Gamma$ . Moreover,  $S_{\Gamma_g} = 2$ ,  $S_{\Gamma_s} = 2$ , and  $S_\Gamma = 2$ , so Eq. (4.3) holds.

Next consider Fig. 21. The diagram  $\Gamma_g$  in (c) can be inserted into the skeleton  $\Gamma_s$  in (d) into two different locations, yielding the isomorphic diagrams in (a) and (b).

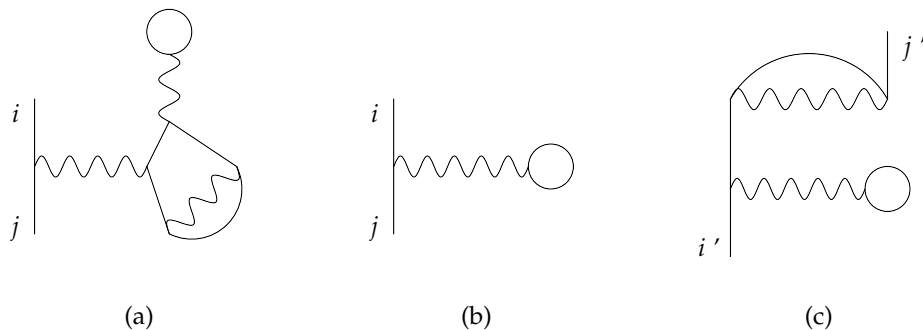


Figure 20: Decomposition of a non-skeleton diagram  $\Gamma$  in (a) into a skeleton diagram  $\Gamma_s$  in (b) and a truncated Green's function diagram  $\Gamma_g$  in (c).

Hence  $r_\Gamma = 2$ . Moreover, observe that  $S_{\Gamma_g} = 1$ ,  $S_{\Gamma_s} = 4$ , and  $S_\Gamma = 2$ , so Eq. (4.3) holds.

By contrast the diagram  $\Gamma$  in Fig. 22, which has the same skeleton  $\Gamma_s$  but admits two maximal insertions, has a redundancy factor of only  $r_\Gamma = 1$ . The left-right symmetry of the diamond does not yield additional redundancy because the maximal insertions exchanged by this symmetry are isomorphic to each other—and in fact to the insertion of Fig. 21 (c). Thus in the bold diagram expansion Eq. (4.1),  $\Gamma$  is only accounted for *once*. Meanwhile,  $S_\Gamma = 4$ ,  $S_{\Gamma_s} = 4$ , and the symmetry factors of the insertions are both one, so Eq. (4.3) holds.

In Fig. 23 (a), we show a non-skeleton diagram  $\Gamma$  which has the same skeleton  $\Gamma_s$  as in the last two examples.  $\Gamma$  admits two (non-symmetric) maximal insertions, each isomorphic to the diagram of Fig. 20 (c). There are two nonequivalent ways of inserting these diagrams into  $\Gamma_s$  to yield a diagram isomorphic to  $\Gamma$ , depicted separately in Fig. 23 (a), (b). We have  $S_\Gamma = 8$  (with a factor of 4 coming from the two half-dumbbells) and  $S_{\Gamma_s} = 4$ , and the symmetry factor of each insertion is 2 (due to the half-dumbbell), so Eq. (4.3) holds.

Finally, in Fig. 23, we show a diagram  $\Gamma$  which once again has the same skeleton  $\Gamma_s$  as in the last several examples.  $\Gamma$  admits one (non-symmetric) maximal insertion isomorphic to the diagram of Fig. 20 (c). This can be inserted into either propagator line of the 'bubble' in the center of  $\Gamma_s$  and with either orientation to yield  $\Gamma$  up to isomorphism, so  $r_\Gamma = 4$ . We have  $S_\Gamma = 2$  (due to the half-dumbbell) and  $S_{\Gamma_s} = 4$ , and the symmetry factor of the insertion is 2 (due to the half-dumbbell), so Eq. (4.3) holds.

We will refer back to these examples for concreteness in the developments that follow.

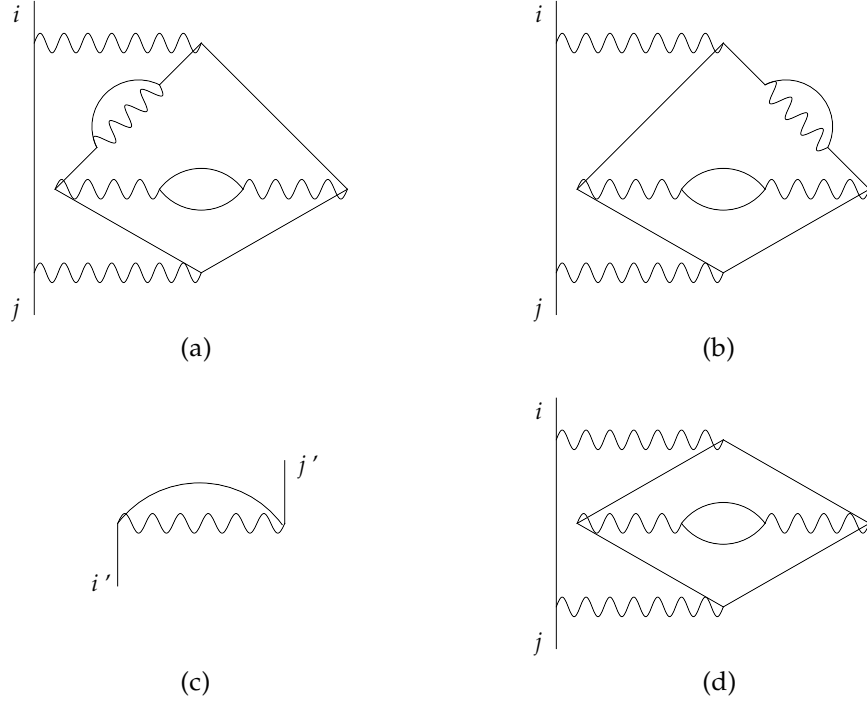


Figure 21: Decomposition of a more complex non-skeleton diagram  $\Gamma$  in (a) into a skeleton diagram  $\Gamma_s$  in (d) and a truncated Green's function diagram  $\Gamma_g$  in (c). The non-skeleton diagram in (b) is isomorphic to (a), but is obtained by the insertion of  $\Gamma_s$  into  $\Gamma_g$  at a different location.

### 4.3 Ways of producing a self-energy diagram from its skeleton

As promised we provide a rigorous definition of the redundancy factor, as well as the set of ways of producing a self-energy diagram from its skeleton. Consider a self-energy diagram  $\Gamma$ , and write

$$\Gamma = \Gamma_s \oplus_{(h_1^{(1)}, h_2^{(1)}, \dots, h_1^{(K)}, h_2^{(K)})} [\Gamma^{(1)}, \dots, \Gamma^{(K)}] \quad (4.4)$$

via Proposition 24.

*Remark 26.* We assume that the ordering within each pair  $(h_1^{(k)}, h_2^{(k)})$  is chosen so that if  $\Gamma^{(j)}, \Gamma^{(k)}$  are non-isomorphic for some  $j, k$ , then in fact  $\Gamma^{(j)}$  and  $\Gamma^{(k)}$  are non-isomorphic after any external relabeling. In other words, the insertions  $\Gamma^{(k)}$  are externally labeled such that if any two of them are isomorphic up to external labeling, then they are actually isomorphic (with external labeling taken into account). We follow this convention for all decompositions of the form of Eq. (4.4) in the sequel.

Implicitly  $\Gamma_s$  is a ‘subdiagram’ of  $\Gamma$  in that its vertex and half-edge sets are subsets of those of  $\Gamma$ . Roughly speaking, there is only one way to construct  $\Gamma$  from  $\Gamma_s$  via

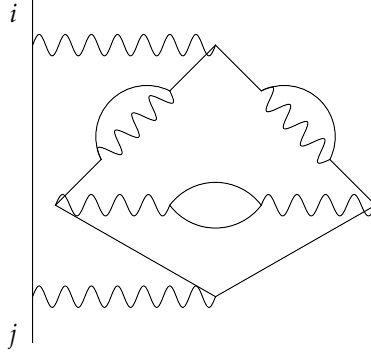


Figure 22: A non-skeleton diagram  $\Gamma$  with redundancy factor  $r_\Gamma = 1$  and the same skeleton  $\Gamma_s$  as in Fig. 21 (d).

Green's function insertions (namely, via the procedure represented in Eq. (4.4)), but there may be many ways to construct diagrams *isomorphic* to  $\Gamma$  from  $\Gamma_s$  via Green's function insertions. The uniqueness result of Proposition 24 guarantees that any such way must involve (up to isomorphism) the insertion of the same set  $\{\Gamma^{(1)}, \dots, \Gamma^{(K)}\}$  of truncated Green's function diagrams. Then let  $I(\Gamma, \Gamma_s)$  be the set of ways of replacing  $K$  propagator lines in  $\Gamma_s$  with  $\Gamma^{(1)}, \dots, \Gamma^{(K)}$  such that the resulting diagram is *isomorphic* to  $\Gamma$  as a truncated Green's function diagram. (There is some abuse of notation here because  $I(\Gamma, \Gamma_s)$  additionally depends on the decomposition of Eq. (4.4), but the meaning will be clear from context.)

More precisely, each such 'way' consists of the following data: a set of ordered pairs of half-edges  $(h_1^{(1)}, h_2^{(1)}), \dots, (h_1^{(K)}, h_2^{(K)})$  in  $\Gamma_s$  such that

$$\Gamma' := \Gamma_s \oplus_{(h_1^{(1)}, h_2^{(1)}), \dots, (h_1^{(K)}, h_2^{(K)})} [\Gamma^{(1)}, \dots, \Gamma^{(K)}]$$

is *isomorphic* to  $\Gamma$ , subject to an *equivalence relation*. Specifically, ways are not distinguished if they differ only by reordering the  $K$  half-edge pairs by a permutation  $\tau \in S_K$  such that  $\Gamma^{(\tau(k))}$  is isomorphic to  $\Gamma^{(k)}$  for all  $k$ . Moreover, ways are not distinguished if they differ only by the ordering *within* the  $k$ -th half-edge pair for  $k$  such that  $\Gamma^{(k)}$  is *symmetric*. We refer to the equivalence class of  $(h_1^{(1)}, h_2^{(1)}), \dots, (h_1^{(K)}, h_2^{(K)})$  as the element of  $I(\Gamma, \Gamma_s)$  *specified* by these half-edge pairs.

Observe that if we sum over the skeleton diagrams and then formally replace each bold line with a sum over Green's function diagrams, then in the resulting formal sum over self-energy diagrams, each self-energy diagram  $\Gamma$  will be counted precisely  $|I(\Gamma, \Gamma_s)|$  times, where  $\Gamma_s$  is the skeleton of  $\Gamma$ . This number  $r_\Gamma := |I(\Gamma, \Gamma_s)|$  depends only on  $\Gamma \in \mathfrak{F}_2^{1\text{PI}}$ , and as suggested earlier we call it the *redundancy factor* of  $\Gamma$ .

It is worthwhile to treat the distinction between symmetric and non-symmetric insertions a bit more elegantly (and, moreover, in a way that does not so clearly privilege the fact that our insertions have *two* external half-edges). For a truncated



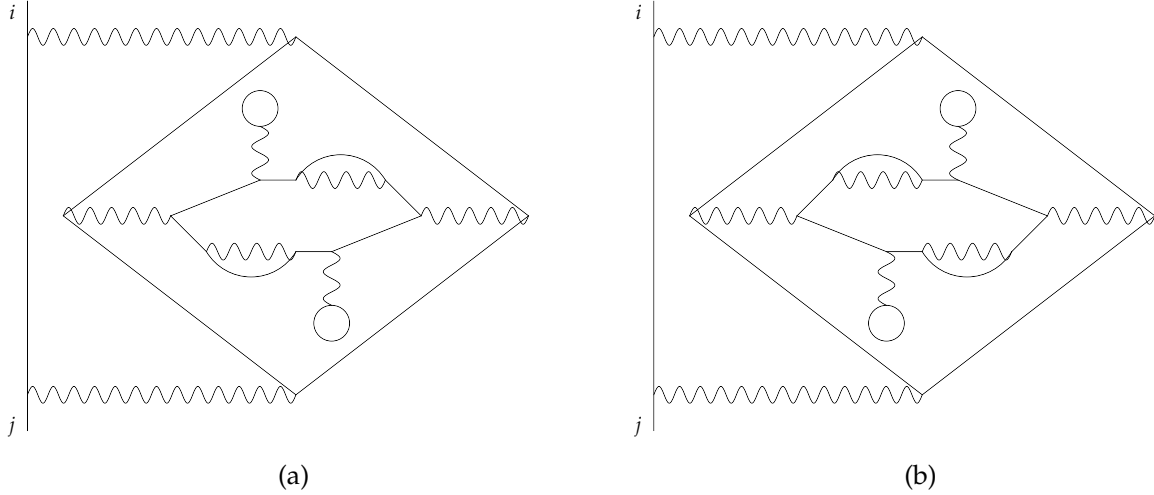


Figure 23: A non-skeleton diagram  $\Gamma$  in (a) with redundancy factor  $r_\Gamma = 2$  and the same skeleton  $\Gamma_s$  as in Fig. 21 (d). The diagram in (b) is a diagram isomorphic to  $\Gamma$  obtained from  $\Gamma_s$  by a nonequivalent set of insertions.

Green's function diagram, the external half-edges have labels ' $i$ ' and ' $j$ .' Let the *external symmetry group*, denoted  $S(\Gamma_g)$ , of a truncated Green's function diagram  $\Gamma_g$  be the subgroup of  $\text{Sym}\{i, j\} \simeq S_2$  consisting of permutations of the labels ' $i$ ' and ' $j$ ' that fix the isomorphism class of the diagram. Therefore for symmetric diagrams the external symmetry group is  $S_2$ , and for non-symmetric diagrams it is the trivial group. For future convenience, let the action of  $\sigma \in S_2$  on a truncated Green's function diagram  $\Gamma_g$  defined via permutation of the external labels be denoted  $\sigma \star \Gamma_g$ . (The ' $\star$ ' notation is meant to distinguish from the group action ' $\cdot$ ' defined earlier.)

Then using this language we can say that ways that for any  $\sigma \in S(\Gamma^{(k)})$  the modification of  $(h_1^{(k)}, h_2^{(k)})$  to  $(h_{\sigma(1)}^{(k)}, h_{\sigma(2)}^{(k)})$  does not yield a distinct element of  $I(\Gamma, \Gamma_s)$ .

#### 4.4 Understanding automorphisms in terms of the skeleton

Now we turn to item (2), i.e., characterizing the structure of automorphisms of  $\Gamma$  in terms of its decomposition furnished by Proposition 24.

With notation as in the section 4.3, let  $n$  be the order of  $\Gamma$ , and let  $p$  be the order of  $\Gamma_s$ . Then  $q = n - p$  is the order of  $\Gamma_g := \bigcup_{k=1}^K \Gamma^{(k)}$ .

We can view the skeleton diagram  $\Gamma_s$  as well as the insertions  $\Gamma^{(k)}$  as labeled truncated Green's function diagrams via the labeling of interaction lines and half-edges inherited from  $\Gamma$ . Let  $\text{Aut}(\Gamma, \Gamma_s)$  be the set of automorphisms of  $\Gamma_s$  that can be *extended* to automorphisms of  $\Gamma$  by relabeling the rest of the diagram, i.e., permuting the vertex and half-edge labels of  $\Gamma_g$ .

For example, the automorphism of  $\Gamma_s$  of Fig. 21 (d) corresponding to the left-right

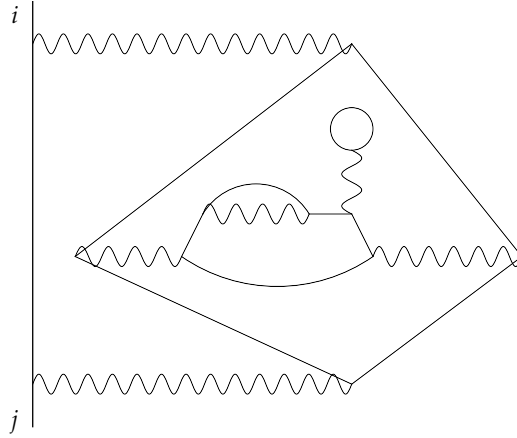


Figure 24: A non-skeleton diagram  $\Gamma$  with redundancy factor  $r_\Gamma = 4$  and the same skeleton  $\Gamma_s$  as in Fig. 21 (d).

reflection of the ‘diamond’ can be extended to an automorphism of the diagram in Fig. 22, but it *cannot* be extended to an automorphism of any of the diagrams of Fig. 21 (a), (b), nor of any of the diagrams of Figs. 23 and 24. Next, consider the automorphism of  $\Gamma_s$  obtained by composing a left-right reflection of the diamond with a swap of the two propagator lines in the ‘bubble’ at the center of the diamond. This extends to an automorphism of each of the diagrams of Fig. 21 (a), (b).

More precisely, viewing  $\mathbf{R}_q$  as acting on labelings of  $\Gamma_g$ , an element  $g \in \text{Aut}(\Gamma_s)$  is defined to be in  $\text{Aut}(\Gamma, \Gamma_s)$  if there exists  $h \in \mathbf{R}_q$  such that  $gh \in \text{Aut}(\Gamma)$ . (Since  $\Gamma_s$  and  $\Gamma_g$  are disjoint, elements  $g$  and  $h$  as in the preceding commute.) Note that  $\text{Aut}(\Gamma, \Gamma_s)$  is a subgroup of  $\text{Aut}(\Gamma_s)$ : indeed, if  $g_1, g_2 \in \text{Aut}(\Gamma, \Gamma_s)$ , then there exist  $h_1, h_2 \in \mathbf{R}_q$  such that  $g_1 h_1, g_2 h_2 \in \text{Aut}(\Gamma)$ , but then  $(g_1 g_2)(h_1 h_2) = (g_1 h_1)(g_2 h_2)$  is in  $\text{Aut}(\Gamma)$ , so  $g_1 g_2 \in \text{Aut}(\Gamma, \Gamma_s)$ .

We have the following characterization of  $\text{Aut}(\Gamma, \Gamma_s)$ :

**Lemma 27.** *Let  $\Gamma \in \mathfrak{F}_2^{\text{1PI}}$ , and write*

$$\Gamma = \Gamma_s \oplus_{(h_1^{(1)}, h_2^{(1)}, \dots, h_1^{(K)}, h_2^{(K)})} [\Gamma^{(1)}, \dots, \Gamma^{(K)}]$$

*via Proposition 24. An element  $g \in \text{Aut}(\Gamma_s)$  lies in  $\text{Aut}(\Gamma, \Gamma_s)$  if and only if for every  $k$ , there is some  $k'$  and some  $\sigma \in S_2$  such that  $\Gamma^{(k)}$  is isomorphic to  $\sigma \star \Gamma^{(k')}$  and  $\psi_g$  sends  $(h_1^{(k)}, h_2^{(k)})$  to  $(h_{\sigma(1)}^{(k')}, h_{\sigma(2)}^{(k')})$ .*

*Proof.* First we prove the forward direction, so let  $g \in \text{Aut}(\Gamma, \Gamma_s)$ . Then  $g$  extends to an automorphism of  $\Gamma$ , which we shall also call  $g$ . Let  $e_1^{(k)}, e_2^{(k)}$  be the external half-edges of the truncated Green’s function diagram  $\Gamma^{(k)}$  paired with  $h_1^{(k)}, h_2^{(k)}$ , respectively, in the overall diagram  $\Gamma$  (equivalently, the external half-edges labeled ‘ $i$ ’ and ‘ $j$ ,’ respectively). Then the maximal insertion  $\Gamma^{(k)}$  is disconnected from the

rest of  $\Gamma$  by unpairing  $\{e_1^{(k)}, h_1^{(k)}\}$  and  $\{e_2^{(k)}, h_2^{(k)}\}$  in  $\Gamma$ . Since  $g$  is an automorphism, removing the links  $\{\psi_g(e_1^{(k)}), \psi_g(h_1^{(k)})\}$  and  $\{\psi_g(e_2^{(k)}), \psi_g(h_2^{(k)})\}$  from  $\Gamma$  must also disconnect a maximal insertion (isomorphic to  $\Gamma^{(k)}$ ) at  $(\psi_g(h_1^{(k)}), \psi_g(h_2^{(k)}))$  with external half-edges  $\psi_g(e_1^{(k)}), \psi_g(e_2^{(k)})$  labeled ‘ $i$ ’ and ‘ $j$ ,’ respectively. Since this diagram is a maximal insertion, by Proposition 24 it must be  $\sigma \star \Gamma^{(k')}$  for some  $k'$ , where  $\sigma \in S_2$ , and moreover  $\psi_g(h_1^{(k)}) = h_{\sigma(1)}^{(k')}$  and  $\psi_g(h_2^{(k)}) = h_{\sigma(2)}^{(k')}$ . This concludes the proof of the forward direction.

Now assume that  $g \in \text{Aut}(\Gamma_s)$  and that for every  $k$ , there is some  $k' = k'(k)$  and some  $\sigma = \sigma(k) \in S_2$  such that  $\Gamma^{(k')}$  is isomorphic to  $\sigma \star \Gamma^{(k)}$  via some isomorphism  $(\varphi^{(k)}, \psi^{(k)})$  and  $\psi_g$  sends  $(h_1^{(k)}, h_2^{(k)})$  to  $(h_{\sigma(1)}^{(k')}, h_{\sigma(2)}^{(k')})$ . Then we aim to extend  $g$  to an automorphism of  $\Gamma$ , i.e., we aim to extend  $(\varphi_g, \psi_g)$  to an isomorphism from  $\Gamma$  to itself. This can be done simply by mapping vertices and half-edges lying in the  $\Gamma^{(k)}$  via  $(\varphi^{(k)}, \psi^{(k)})$ . It is straightforward to check that this indeed defines an automorphism.  $\square$

We also have the following result characterizing the structure of automorphisms of  $\Gamma$  in terms of  $\text{Aut}(\Gamma, \Gamma_s)$ :

**Lemma 28.** *Let  $\Gamma \in \mathfrak{F}_2^{\text{1PI}}$  be decomposed as in Eq. (4.2). Then any  $g \in \text{Aut}(\Gamma)$  restricts to an automorphism of  $\Gamma_s$  (in particular, only permutes vertex labels within the subdiagram  $\Gamma_s$ ). By definition, this induced automorphism of  $\Gamma_s$  then lies in  $\text{Aut}(\Gamma, \Gamma_s)$ . Moreover, if  $\Gamma$  admits a maximal insertion  $\Gamma'$  at  $(h'_1, h'_2)$ , then  $\Gamma$  also admits a maximal insertion  $\Gamma''$  isomorphic to  $\Gamma'$  at  $(\psi_g(h'_1), \psi_g(h'_2))$ . Furthermore,  $g$  sends all vertex labels from  $\Gamma'$  to  $\Gamma''$ , i.e.,  $\varphi_g$  sends each vertex of  $\Gamma'$  to a vertex of  $\Gamma''$ .*

*Proof.* Let  $g \in \text{Aut}(\Gamma)$ . First we prove that  $g$  only permutes vertex labels *within*  $\Gamma_s$ . Indeed, suppose not. Then  $\varphi_g$  sends a vertex that is not contained in any insertion to a vertex that is contained in some insertion. The property of whether or not a vertex is contained in an insertion is preserved under diagram isomorphism, so we have a contradiction.

Furthermore, an isomorphism of unlabeled diagrams sends maximal insertions to maximal insertions; i.e., if  $\Gamma$  admits a maximal insertion  $\Gamma'$  at  $(h'_1, h'_2)$ , then  $\Gamma$  also admits a maximal insertion  $\Gamma''$  isomorphic to  $\Gamma'$  at  $(\psi_g(h'_1), \psi_g(h'_2))$  via  $(\varphi_g, \psi_g)$ . (In particular  $\varphi_g$  sends each vertex of  $\Gamma'$  to a vertex of  $\Gamma''$ .)

Then it can be readily checked, by collapsing maximal insertions, that  $g$  descends to an automorphism of the skeleton  $\Gamma_s$ . Then by definition we can view  $g \in \text{Aut}(\Gamma, \Gamma_s)$ .  $\square$

The preceding two lemmas can be used to compute the symmetry factor of  $\Gamma \in \mathfrak{F}_2^{\text{1PI}}$  via its skeleton decomposition:

**Lemma 29.** *Let  $\Gamma \in \mathfrak{F}_2^{\text{1PI}}$ , decomposed as in Eq. (4.2). Then*

$$S_\Gamma = |\text{Aut}(\Gamma, \Gamma_s)| \cdot \prod_{k=1}^K S_{\Gamma^{(k)}} \quad (4.5)$$

*Proof.* Lemma 28 says that every  $g \in \text{Aut}(\Gamma)$  descends to  $g \in \text{Aut}(\Gamma, \Gamma_s)$  and moreover defines an isomorphism from each insertion  $\Gamma^{(k)}$  to its image under  $g$ .

Conversely, by Lemma 27, for any  $g \in \text{Aut}(\Gamma, \Gamma_s)$  and any  $k$ ,  $\psi_g$  sends  $(h_1^{(k)}, h_2^{(k)})$  to  $(h_{\sigma(1)}^{(k')}, h_{\sigma(2)}^{(k')})$  for some  $\sigma = \sigma(k) \in S_2$ , where  $k' = k'(k)$  is such that  $\Gamma^{(k)}$  is isomorphic to  $\sigma \star \Gamma^{(k')}$ . Any choice of isomorphisms from the  $\Gamma^{(k)}$  to the  $\sigma \star \Gamma^{(k')}$  defines an extension of  $g$  to an automorphism of  $\Gamma$ .

Thus any automorphism of  $\Gamma$  can be yielded constructively by starting with  $g \in \text{Aut}(\Gamma, \Gamma_s)$  and then choosing, for each  $k$ , an isomorphism from  $\Gamma^{(k)}$  to  $\sigma \star \Gamma^{(k')}$ . The number of such isomorphisms is the same as the number of automorphisms of  $\Gamma^{(k)}$ , so Eq. (4.5) follows.  $\square$

## 4.5 The action of $\text{Aut}(\Gamma_s)$ on $I(\Gamma, \Gamma_s)$

Finally, we turn to item (3). Again decompose  $\Gamma \in \mathfrak{F}_2^{\text{PI}}$  as in Eq. (4.2). Notice that Eq. (4.2) itself defines an element of  $I(\Gamma, \Gamma_s)$ . Call this element  $\iota^*$ .

The key observation here is that  $\text{Aut}(\Gamma_s)$  acts transitively on  $I(\Gamma, \Gamma_s)$  and that the stabilizer of any  $\iota \in I(\Gamma, \Gamma_s)$  is  $\text{Aut}(\Gamma, \Gamma_s)$ . We define the action as follows. Let  $g \in \text{Aut}(\Gamma_s)$ , and consider an element  $\iota$  of  $I(\Gamma, \Gamma_s)$  specified by a set of ordered pairs of half-edges  $(h_1^{(1)}, h_2^{(1)}), \dots, (h_1^{(K)}, h_2^{(K)})$  in  $\Gamma_s$ . Then  $g \cdot \iota$  is defined to be the element of  $I(\Gamma, \Gamma_s)$  specified by the ordered pairs  $(\psi_g(h_1^{(1)}), \psi_g(h_2^{(1)})), \dots, (\psi_g(h_1^{(K)}), \psi_g(h_2^{(K)}))$ .

For example, recall the automorphism of  $\Gamma_s$  of Fig. 21 (d) corresponding to the left-right reflection of the ‘diamond.’ The action of this automorphism fixes the only element in  $I(\Gamma, \Gamma_s)$  represented by Fig. 22. Meanwhile, it swaps the elements of  $I(\Gamma, \Gamma_s)$  represented in Fig. 21 (a) and (b). In a slightly more indirect way, it also swaps the elements of  $I(\Gamma, \Gamma_s)$  represented in Figs. 23 (a) and (b). Next, consider the automorphism of  $\Gamma_s$  obtained by a swap of the two propagator lines in the ‘bubble’ at the center of the diamond. The action of this automorphism also swaps the elements of  $I(\Gamma, \Gamma_s)$  represented in Figs. 23 (a) and (b).

**Lemma 30.** *With notation as in the preceding, the action of  $\text{Aut}(\Gamma_s)$  on  $I(\Gamma, \Gamma_s)$  is transitive, and the stabilizer of  $\iota^*$  is  $\text{Aut}(\Gamma, \Gamma_s)$ .*

*Proof.* First we establish that the action is transitive. To this end, consider arbitrary elements  $\iota_1, \iota_2 \in I(\Gamma, \Gamma_s)$ , i.e., two different ways of making insertions in  $\Gamma_s$  to yield diagrams  $\Gamma_1, \Gamma_2$ , respectively, that are each isomorphic to  $\Gamma$ . Our isomorphism from  $\Gamma_1$  to  $\Gamma_2$  descends (by collapsing the maximal insertions) to an isomorphism from  $\Gamma_s$  to itself, i.e., an automorphism  $g \in \text{Aut}(\Gamma_s)$ , and evidently this automorphism satisfies  $g \cdot \iota_1 = \iota_2$ . This establishes transitivity.

Now we turn to the claim about the stabilizer. Let  $g \in \text{Aut}(\Gamma, \Gamma_s)$ . We want to show that  $g \cdot \iota^* = \iota^*$ . By Lemma 27, there exists  $\tau \in S_K$  and  $\sigma_k \in S_2$  for  $k = 1, \dots, K$  such that  $\Gamma^{(k)}$  is isomorphic to  $\sigma_k \star \Gamma^{(\tau(k))}$  and  $\psi_g$  sends  $(h_1^{(k)}, h_2^{(k)})$  to  $(h_{\sigma_k(1)}^{(\tau(k))}, h_{\sigma_k(2)}^{(\tau(k))})$ . Hence  $g \cdot \iota^*$  is specified by the ordered pairs  $(h_{\sigma_k(1)}^{(\tau(k))}, h_{\sigma_k(2)}^{(\tau(k))})$ ,  $k = 1, \dots, K$ .

By Remark 26, since  $\Gamma^{(k)}$  and  $\Gamma^{(\tau(k))}$  are isomorphic up to external labeling, they are in fact isomorphic. But since  $\Gamma^{(k)}$  and  $\sigma_k \star \Gamma^{(\tau(k))}$  are isomorphic, this means that in turn  $\sigma_k \star \Gamma^{(\tau(k))}$  is isomorphic to  $\Gamma^{(\tau(k))}$ , i.e.,  $\sigma_k \in S(\Gamma^{(k)})$  for all  $k$ .

Then recalling the equivalence relation used to define  $I(\Gamma, \Gamma_s)$ , we see that  $g \cdot \iota^*$  is equivalently specified by the ordered pairs  $(h_1^{(k)}, h_2^{(k)})$ ,  $k = 1, \dots, K$ , i.e.,  $g \cdot \iota^* = \iota^*$ .

Conversely, suppose that  $g \cdot \iota^* = \iota^*$  for some  $g \in \text{Aut}(\Gamma_s)$ . Then there exist  $\tau \in S_K$  and  $\sigma_k \in S(\Gamma^{(k)})$  for  $k = 1, \dots, K$  such that  $\psi_g$  sends  $(h_1^{(k)}, h_2^{(k)})$  to  $(h_{\sigma_k(1)}^{(\tau(k))}, h_{\sigma_k(2)}^{(\tau(k))})$  for  $k = 1, \dots, K$ , and moreover  $\Gamma^{(k)}$  is isomorphic to  $\Gamma^{(\tau(k))}$  for all  $k$ . Since  $\sigma_k \in S(\Gamma^{(k)})$ , this means that  $\Gamma^{(k)}$  is isomorphic to  $\sigma_k \star \Gamma^{(\tau(k))}$  for all  $k$ . But then by Lemma 27 we have that  $g \in \text{Aut}(\Gamma, \Gamma_s)$ .  $\square$

Then the orbit-stabilizer theorem, together with Lemmas 29 and 30, yields the following corollary:

**Corollary 31.** *For  $\Gamma \in \mathfrak{F}_2^{\text{1PI}}$ , the redundancy factor of  $\Gamma$  is given by*

$$r_\Gamma = \frac{S_{\Gamma_s} \cdot \prod_{k=1}^K S_{\Gamma^{(k)}}}{S_\Gamma}.$$

*Proof.* Applying the orbit-stabilizer theorem via Lemma 30 we obtain

$$r_\Gamma = |I(\Gamma, \Gamma_s)| = \frac{|\text{Aut}(\Gamma_s)|}{|\text{Aut}(\Gamma, \Gamma_s)|} = \frac{S_{\Gamma_s}}{|\text{Aut}(\Gamma, \Gamma_s)|}.$$

The result then follows from Lemma 29.  $\square$

## 4.6 Bold diagrammatic expansion for the self-energy

At last we can prove the bold diagrammatic expansion for the self-energy, stated as follows:

**Theorem 32.** *For  $1 \leq i, j \leq N$ , we have the equality of formal power series (in the coupling constant)*

$$\Sigma_{ij} = \sum_{\Gamma_s \in \mathfrak{F}_2^{\text{2PI}}} \frac{\mathbf{F}_{\Gamma_s}(i, j)}{S_{\Gamma_s}}, \quad (4.6)$$

where  $\Sigma_{ij}$  is interpreted as a power series via Theorem 20 and where, for every  $\Gamma_s \in \mathfrak{F}_2^{\text{2PI}}$ , the expression  $\mathbf{F}_{\Gamma_s}(i, j)$  is interpreted as the power series obtained by applying the Feynman rules for  $\Gamma_s$  with propagator  $G$ , where  $G$  is in turn interpreted as a formal power series via Theorem 18.

*Remark 33.* The interpretation of the bold diagrammatic expansion of the self-energy is at this point somewhat cryptic. For the moment it can only be interpreted as a reorganization of the terms in the asymptotic series for the self-energy. However, since the terms on the right-hand side of Eq. (4.6) depend only on  $G$  and  $v$  (and not

on  $G^0$ ), one might conjecture based on the expansion that the self-energy depends only on  $G, v$ . This is indeed a major goal of Part III, where we will indeed construct the self-energy (non-perturbatively) as a (matrix-valued) functional  $\Sigma[G, v]$  of  $G$  and  $v$  only, and interpret the bold diagrammatic expansion as an asymptotic series in the coupling constant for the self-energy at fixed  $G$ , with terms given by the  $\Sigma^{(k)}[G, v]$  to be specified below. The non-perturbative perspective will guarantee the existence of such an asymptotic series, but Theorem 32 will be used to show that this series is in fact given by Eq. (4.6).

*Proof.* In the following all expressions should be suitably interpreted as in the statement of the theorem. For  $\Gamma_s \in \mathfrak{F}_2^{2\text{PI}}$ , the series  $\mathbf{F}_{\Gamma_s}(i, j)$  counts  $r_\Gamma$  times every self-energy diagram  $\Gamma \in \mathfrak{F}_2^{1\text{PI}}$  with skeleton isomorphic to  $\Gamma_s$ , each with factor

$$F_\Gamma(i, j) / \prod_{k=1}^K S_{\Gamma^{(k)}},$$

where the  $\Gamma^{(k)}$  are the maximal insertions of  $\Gamma$ . Then by Corollary 31,  $\mathbf{F}_{\Gamma_s}(i, j)/S_{\Gamma_s}$  equals the sum of  $F_\Gamma(i, j)/S_\Gamma$  over self-energy diagrams  $\Gamma$  with skeleton isomorphic to  $\Gamma_s$ . Therefore the right-hand side of Eq. (4.6) is the sum of  $F_\Gamma(i, j)/S_\Gamma$  over all self-energy diagrams  $\Gamma$ .  $\square$

Following Remark 33, each term in the bold diagrammatic expansion can be thought of as a functional of  $G$  and  $v$ . We indicate by  $\Sigma^{(k)}[G, v]$  the  $k$ -th order bold contribution to the self-energy, i.e., the contribution of the terms in the diagrammatic expansion that are of order  $k$  in the interaction  $v$ . In particular, the ‘first-order’ bold contribution to the self-energy is given by

$$(\Sigma^{(1)}[G, v])_{ij} = -\frac{1}{2} \left( \sum_k v_{ik} G_{kk} \right) \delta_{ij} - v_{ij} G_{ij}. \quad (4.7)$$

These two terms are represented in Fig. 25 (a), (b), and we denote them by  $\Sigma_{\text{H}}[G, v]$  and  $\Sigma_{\text{F}}[G, v]$  for ‘Hartree’ and ‘Fock,’ respectively. The associated diagrams happen to be the same as the first-order bare self-energy diagrams, but with thin lines replaced by bold lines.

The approximation  $\Sigma[G, v] \approx \Sigma^{(1)}[G, v]$  is known as the *Hartree-Fock approximation*. One can likewise approximate  $\Sigma[G, v] \approx \Sigma^{(1)}[G, v] + \Sigma^{(2)}[G, v]$ , where the second-order contribution can be written

$$(\Sigma^{(2)}[G])_{ij} = \frac{1}{2} G_{ij} \sum_{k,l} v_{ik} G_{kl}^2 v_{lj} + \sum_{k,l} v_{ik} G_{kj} G_{kl} G_{li} v_{jl}, \quad (4.8)$$

and the second-order bold diagrams are shown in Fig. 25 (c), (d). The latter is known as the ‘second-order exchange,’ while the former is an example of what is called a ‘ring diagram,’ for reasons to be made clear later.

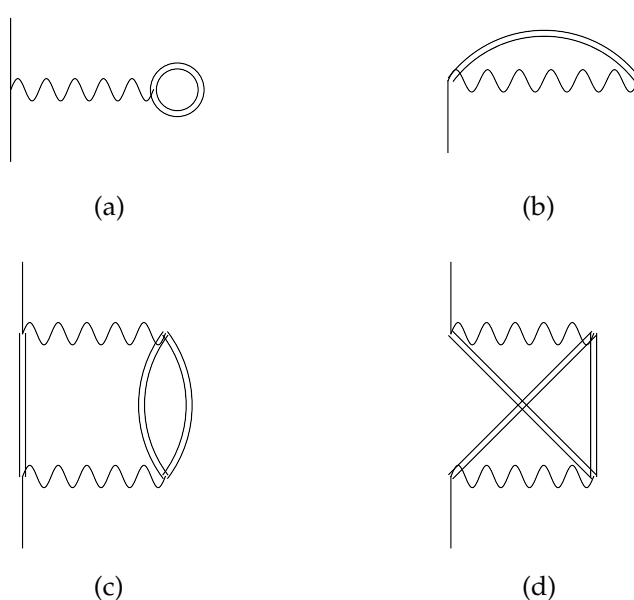


Figure 25: Bold diagrammatic expansion of the self-energy up to second order (external labelings omitted).

## 4.7 Green's function methods

A *Green's function method* using bold diagram expansion is a method for computing the Green's function via an ansatz for the self-energy  $\Sigma_{\text{ans}}[G, v] \approx \Sigma[G, v]$ . This ansatz should be viewed as some sort of approximation of the full self-energy, usually consisting of diagrammatic contributions meant to incorporate certain physical effects.

After choosing an ansatz, one substitutes  $\Sigma \leftarrow \Sigma_{\text{ans}}[G, v]$  in the Dyson equation (3.24) and attempts to solve it *self-consistently* for  $G$ . In other words, one solves

$$G = (A - \Sigma_{\text{ans}}[G, v])^{-1}$$

for  $G$ , where  $A$  and  $v$  are specified in advance.

The most immediate technique for solving this system is a fixed-point iteration that we refer to as the *Dyson iteration*, which is defined by the update

$$G^{(k+1)} = (A - \Sigma_{\text{ans}}[G^{(k)}, v])^{-1}.$$

This iteration can be combined with damping techniques to yield better convergence in practice, but in general the convergence behavior of the Dyson iteration (even with damping) may depend strongly on the ansatz for the self-energy.

The Green's function method obtained by the ansatz  $\Sigma_{\text{ans}}[G, v] = \Sigma^{(1)}[G, v]$  is known as the *Hartree-Fock method*. One is likewise free to consider higher-order approximations for the self-energy. However it should be emphasized that it is not

obvious *a priori* which order of approximation is optimal for a given problem specification. Some numerical comparison of methods will be undertaken in Section 5, but further comparisons are a subject of detailed study to be left for future work.

It should be noted that once the Green’s function is computed, it can be used to compute the internal energy of the system via the Galitskii-Migdal formula of Theorem 1. It can also be used to compute the Gibbs free energy via the Luttinger-Ward formalism in a way to be explained below. More remarkably, the framework of Green’s function methods can even be used, as in the dynamical mean field theory (DMFT) [37], to compute effective Hamiltonians on smaller subsystems (‘fragments’) that self-consistently account for their interaction with their environments. This will be studied in future work.

### 4.7.1 The GW approximation

In order to provide a broader perspective on both Green’s function methods and diagrammatic manipulations, we include here a diagrammatic derivation of the GW approximation [43] for the self-energy, which corresponds to an ansatz for the self-energy yielded by a further summation over an *infinite* subset of the bold diagrams. The GW method is the corresponding Green’s function method.

This summation, which is represented graphically in Fig. 26, includes the Hartree diagram, together with all of the so-called ring diagrams.

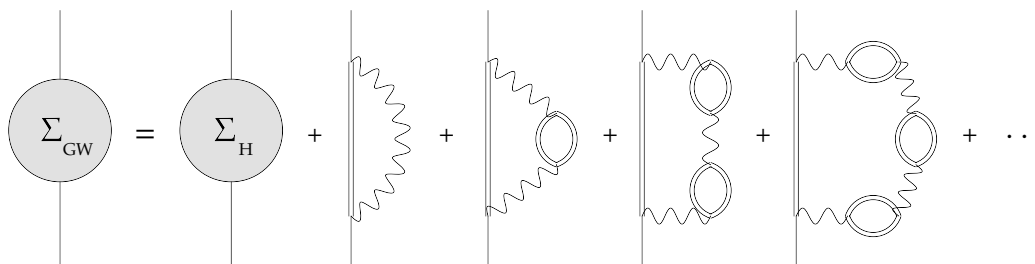


Figure 26: Diagrammatic depiction of the GW self-energy.

The Fock exchange diagram of Fig. 25 (b) can be thought of as the ‘zeroth’ ring diagram, and Fig. 25 (c) is the first ring diagram. Notice that the  $k$ -th ring diagram  $\Gamma_k$  has a symmetry factor of  $S_{\Gamma_k} = 2^k$ , with a factor of 2 deriving from the each symmetry that exchanges the two propagators in one of the  $k$  ‘bubbles.’

Furthermore, the  $k$ -th ring diagram has Feynman amplitude given by

$$\mathbf{F}_{\Gamma_k}(i, j) = - \left( [-v(G \odot G)]^k v \right)_{ij} G_{ij},$$

or equivalently,

$$\mathbf{F}_{\Gamma_k} = -G \odot \left( [-v(G \odot G)]^k v \right),$$

where ‘ $\odot$ ’ indicates the entrywise matrix product.



Then formally summing the geometric series we obtain

$$\sum_{k=0}^{\infty} \frac{\mathbf{F}_{\Gamma_k}(i, j)}{S_{\Gamma_k}} = -G \odot \left( \left[ I + \frac{1}{2}v(G \odot G) \right]^{-1} v \right),$$

where the factor  $\frac{1}{2}$  arises from the symmetry factor. Incorporating the Hartree term, we arrive at the GW ansatz for the self-energy:

$$\Sigma_{\text{GW}}[G, v] = \Sigma_{\text{H}}[G, v] - G \odot W[G, v],$$

where

$$W[G, v] := \left( \left[ I + \frac{1}{2}v(G \odot G) \right]^{-1} v \right) = \left[ v^{-1} + \frac{1}{2}(G \odot G) \right]^{-1},$$

is known as the *screened Coulomb interaction*. Thus the GW self-energy (whose name derives from the  $G \odot W$  term appearing therein) looks very much like the Hartree-Fock self energy, but with the Fock exchange replaced by a *screened exchange*, in which  $W$  assumes the role of  $v$ .

## 4.8 A preview of the Luttinger-Ward formalism

There is in fact a more fundamental formalism underlying the self-energy ansatzes and Green's function methods outlined above, which can be recovered from ansatzes for the so-called *Luttinger-Ward (LW) functional*.

It will turn out (as will be demonstrated in Part III) that the exact self-energy  $\Sigma[G, v]$ , viewed as a matrix-valued functional of  $G$  for fixed  $v$ , can be written as the *matrix derivative* of a scalar-valued functional of  $G$ , as in

$$\Sigma[G, v] = \frac{\partial \Phi}{\partial G}[G, v]. \quad (4.9)$$

Here  $\Phi[G, v]$  is the LW functional, which additionally satisfies  $\Phi[0, v] = 0$ .

We must include some technical commentary to make precise sense of Eq. (4.9). In fact,  $\Phi[\cdot, v]$  should be thought of as a function  $\mathcal{S}_{++}^n \rightarrow \mathbb{R}$ , where  $\mathcal{S}_{++}^n$  is the set of symmetric positive definite  $N \times N$  matrices. Then the derivative  $\frac{\partial}{\partial G}$  should be defined in terms of variations *within*  $\mathcal{S}_{++}^n$ . Letting  $E^{(ij)} \in \mathcal{S}^N$  be defined by  $E_{kl}^{(ij)} = \delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk}$ , we then define  $\frac{\partial}{\partial G} := \left( \frac{\partial}{\partial G_{ij}} \right) := \frac{1}{2}(D_{E^{(ij)}})$ , where  $D_{E^{(ij)}}$  indicates the directional derivative in the direction  $E^{(ij)}$ . If  $f : \mathcal{S}_{++}^n \rightarrow \mathbb{R}$  is obtained by the restriction of a function  $f : \mathbb{R}^{N \times N} \rightarrow \mathbb{R}$  that is specified by a formula  $X \mapsto f(X)$  in which the roles of  $X_{kl}$  and  $X_{lk}$  are the same for all  $l, k$ , then  $\frac{\partial}{\partial G}$  simply coincides with the usual matrix derivative. This is the basic scenario underlying various computations below.

The LW functional relates to the free energy in the following manner. Consider the free energy as a functional of  $A$  and  $v$  via  $\Omega = \Omega[A, v]$ . Then for  $A$  and  $v$  that yield Green's function  $G$ , we will derive in Part III the relation

$$\Omega[A, v] = \frac{1}{2} \text{Tr}[AG] - \frac{1}{2} \text{Tr} \log G - \frac{1}{2} (\Phi[G, v] + \Phi_0), \quad (4.10)$$

where  $\Phi_0 = N \log(2\pi e)$  is a constant.

Moreover, much like  $\Sigma[G, v]$ , for fixed  $G$  the LW functional  $\Phi[G, v]$  will admit an asymptotic series in the coupling constant with terms denoted  $\Phi^{(k)}[G, v]$ . It will follow, on the basis of Eq. (4.9), that

$$\Phi^{(k)}[G, v] = \frac{1}{2k} \text{Tr} (G \Sigma^{(k)}[G, v]). \quad (4.11)$$

Therefore  $\Phi$  admits a bold diagrammatic expansion obtained by linking the external half-edges of every skeleton diagram with a bold propagator to obtain a closed diagram, as in Fig. 25.

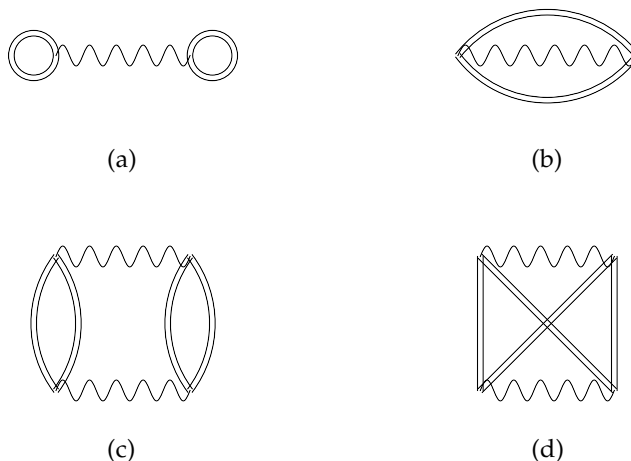


Figure 27: Bold diagrammatic expansion of the LW functional up to second order.

It is important to realize that the pre-factors for these diagrams are obtained rather differently than the diagrams we have already seen (though the Feynman amplitudes are computed as usual). Indeed, for each LW diagram, one must sum over the prefactors of all skeleton diagrams from which it can be produced, then divide by  $2k$ , where  $k$  is the number of interaction lines.

For example, we obtain

$$\Phi^{(1)}[G, v] = -\frac{1}{4} \sum_{i,j} v_{ij} G_{ii} G_{jj} - \frac{1}{2} \sum_{i,j} v_{ij} G_{ij} G_{ij} \quad (4.12)$$

and

$$\Phi^{(2)}[G] = \frac{1}{8} \sum_{i,j,k,l} G_{ij}^2 v_{ik} G_{kl}^2 v_{lj} + \frac{1}{4} \sum_{i,j,k,l} v_{ik} v_{jl} G_{ij} G_{kj} G_{kl} G_{li}, \quad (4.13)$$

for the first- and second-order contributions from Eqs. (4.7) and (4.8), together with Eq. (4.11). Moreover we denote by  $\Phi_{\text{H}}[G, v]$  and  $\Phi_{\text{F}}[G, v]$  the first and second terms of Eq. (4.12), for ‘Hartree’ and ‘Fock,’ respectively.

#### 4.8.1 $\Phi$ -derivability

An ansatz  $\Sigma_{\text{ans}}[G, v]$  for  $\Sigma[G, v]$  specifies a Green’s function method. Among these ansatzes, some can be written as matrix derivatives, i.e., can be viewed as being obtained from an ansatz  $\Phi_{\text{ans}}[G, v]$  for the Luttinger-Ward functional via

$$\Sigma_{\text{ans}}[G, v] := \frac{\partial \Phi_{\text{ans}}}{\partial G}[G, v].$$

These approximations are known as  $\Phi$ -*derivable* or *conserving* approximations. In the context of quantum many-body physics, the resulting  $\Phi$ -derivable Green’s function methods are physically motivated in that they respect certain conservation laws [100].

Notice that for fixed  $A$  and  $v$ , once an estimate  $G_{\text{ans}}$  for the Green’s function has been computed via such a method, Eq. (4.10) suggests a way to approximate the free energy, i.e.,

$$\Omega \approx \frac{1}{2} \text{Tr}[AG_{\text{ans}}] - \frac{1}{2} \text{Tr} \log(G_{\text{ans}}) - \frac{1}{2} (\Phi_{\text{ans}}[G, v] + \Phi_0). \quad (4.14)$$

In fact all of the self-energy approximations considered thus far, namely the first-order (Hartree-Fock) and second-order approximations and the GW approximation, are  $\Phi$ -derivable.

In fact the first- and second-order self-energy approximations of Eqs. (4.7) and Eqs. (4.8) can be obtained from the first- and second-order LW approximations of Eqs. (4.12) and Eqs. (4.13).

Meanwhile, the GW approximation can be obtained from

$$\Phi_{\text{GW}}[G, v] = \Phi_{\text{H}}[G, v] - \text{Tr} \log \left[ I + \frac{1}{2} v(G \odot G) \right].$$

Here the matrix derivative of the first term yields the Hartree contribution  $\Sigma_{\text{H}}[G, v]$ , and the matrix derivative of the second term yields  $G \odot W[G, v]$ .

In fact,  $\Phi_{\text{GW}}$  can be viewed as being obtained from an infinite summation of the (closed) ring diagrams from the bold diagrammatic expansion of the LW functional, together with the closed Hartree diagram. These are the obtained by closing up the diagrams of Fig. 26 and are themselves depicted in Fig. 28.

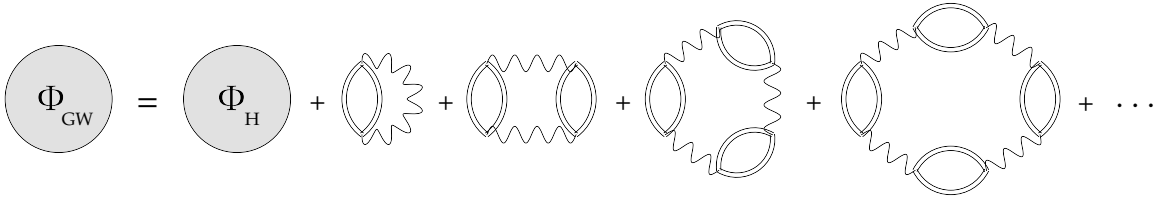


Figure 28: Diagrammatic depiction of the GW Luttinger-Ward functional.

What kind of (reasonable) self-energy approximation would *fail* to be  $\Phi$ -derivable? Roughly speaking, by the usual product rule for derivatives, taking the matrix derivative in  $G$  of a LW diagram yields a sum over all skeleton diagrams that can be obtained from the LW diagram by cutting a single propagator. This means that if one includes some skeleton diagram in the approximate self-energy but *does not* include another skeleton diagram that closes up to the same LW diagram as the first, then the approximation should not be  $\Phi$ -derivable.

Notice that for each of the diagrams in Fig. 27, any choice of the propagator line to be removed yields the same skeleton diagram, so the above scenario cannot apply to these terms. This is one way of seeing why each of the bold skeleton diagrams up to second order is *individually*  $\Phi$ -derivable. (The same is true for each of the ring diagrams.)

To find a non- $\Phi$ -derivable bold skeleton diagram, we have to go to the third order. In Fig. 29 (a), we depict a LW diagram that can be cut in different ways to obtain the distinct skeleton diagrams of Fig. 29 (b) and (c).

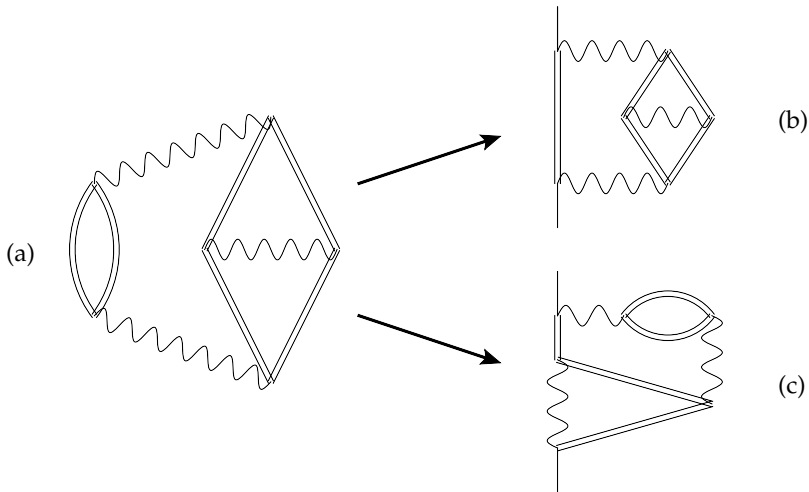


Figure 29: A LW diagram (a) that yields distinct classes of skeleton diagram (b), (c). (External labelings of the skeleton diagrams are ignored.)

Although some skeleton diagrams such as Fig. 29 (b) and (c) are not *individually*  $\Phi$ -derivable, the sum over all skeleton diagrams of a given order *is*  $\Phi$ -derivable.

## 4.9 Quantities that *do not* admit a bold diagrammatic expansion

We stress that here the bold diagrammatic expansion has only been established for the self-energy (and, by extension, the Luttinger-Ward functional). The same concept cannot be generalized to other quantities without verification.

For example, let us consider what goes wrong in the case of the free energy. Theorem 17 expresses the free energy as a sum over bare connected closed diagrams. Why can't we just replace the thin lines with bold lines and then remove redundant diagrams (i.e., the diagrams with Green's function insertions)?

First let us see what happens if we try this. Later we will discuss where the proof of the bold diagrammatic expansion of the self-energy breaks down when we attempt to adapt it to the free energy.

Recall that the first-order free energy diagrams are simply the Hartree and Fock diagrams (i.e. the dumbbell and the oyster). Neither of these admit Green's function insertions, so converting them bold diagrams and retain them.

Now recall that the second-order bare free energy diagrams are depicted in Fig. 6 (b), (c). In particular, consider diagram (b2), reproduced below in Fig. 30 (a). This diagram admits Green's function insertions, hence will not be retained as a bold diagram.

It can be obtained by inserting Green's function diagrams into the bold oyster diagram, shown in Fig. 30 (b), in *two* different ways; indeed, either of the bold propagators in diagram (b) can be replaced with the insertion depicted in Fig. 30 (c).

Since diagram (b) has a symmetry factor of 4 and (c) has a symmetry factor of 1, our attempted bold diagrammatic expansion for the free energy then counts diagram (a) with a pre-factor of  $\frac{1}{2}$ . However, diagram (a) has a symmetry factor of 4, hence has a pre-factor of  $\frac{1}{4}$  in the bare diagrammatic expansion for the free energy!

What went wrong? The problem is that there is no analog of the unique skeleton decomposition (Proposition 24) for *closed* diagrams.<sup>6</sup> Indeed, the diagram of Fig. 30 (a) can be built up from two different 'skeleton' subdiagrams, one containing the upper interaction line and the other containing the lower one.

---

<sup>6</sup>The proof of Proposition 24 fails for closed diagrams in that the case  $|E^{(j,k)}| = 0$  cannot be ruled out. (In the original proof, this case could be ruled out because it implied the existence of a closed subdiagram of a connected self-energy diagram, which is impossible.)

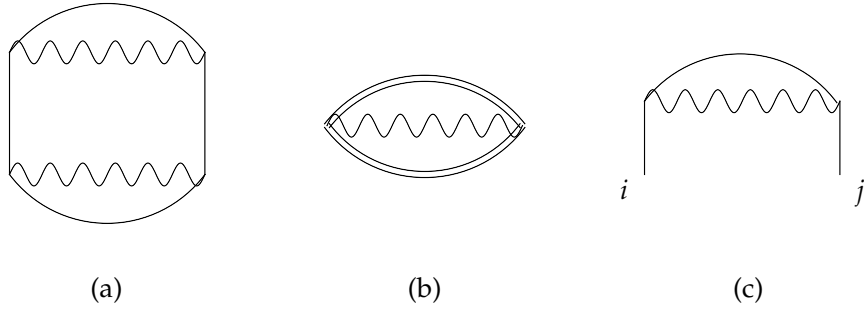


Figure 30: Decomposition of a closed diagram  $\Gamma$  in (a) into a closed diagram  $\Gamma_s$  in (b), and a truncated Green's function diagram  $\Gamma_g$  in (c).

## 5 Numerical experiments

For the Gibbs model, in contrast to the quantum many-body problem, Green's function methods as in section 4.7 can be implemented within a few lines of MATLAB code. In this section we provide a small snapshot of the application of the Gibbs model to the investigation of the numerical performance of MBPT.

In particular, we demonstrate that the use of the LW functional to compute the free energy via Eq. (4.14) can yield more accurate results than the use of the bare diagrammatic expansion of Theorem 17.

For simplicity we consider dimension  $N = 4$ , where all integrals can be evaluated directly via a quadrature scheme, such as Gauss-Hermite quadrature. The quadratic and the quartic terms of the Hamiltonian are specified by

$$A = \begin{pmatrix} 2 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 \\ 0 & -1 & 2 & -1 \\ 0 & 0 & -1 & 2 \end{pmatrix}, \quad v = \lambda \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad (5.1)$$

respectively.

Once the self-consistent Green's function  $G$  is obtained from some  $\Phi$ -derivable Green's function method, we evaluate the free energy using the LW functional via Eq. (4.14). Since the non-interacting free energy  $\Omega_0$  is not physically meaningful (as it corresponds to an additive constant in the Hamiltonian), we measure the relative error of  $\Omega - \Omega_0$ , i.e., we compute  $|\Omega_{\text{ans}} - \Omega_{\text{exact}}|/|\Omega_{\text{exact}} - \Omega_0|$ , where  $\Omega_{\text{ans}}$  is the free energy obtained via our approximation and  $\Omega_{\text{exact}}$  is the exact free energy. We carry out this procedure for the first-order (Hartree-Fock) and second-order approximations (GF2) of the LW functional, denoted 'Bold 1st' and 'Bold 2nd,' respectively, in Fig. 31, which plots the relative error against the coupling constant  $\lambda$ .

For comparison, we also consider the approximations of the free energy obtained directly from the first-order and second-order truncations of the bare diagrammatic

expansion for the free energy of Theorem 17, denoted ‘Bare 1st’ and ‘Bare 2nd,’ respectively, in Fig. 31.

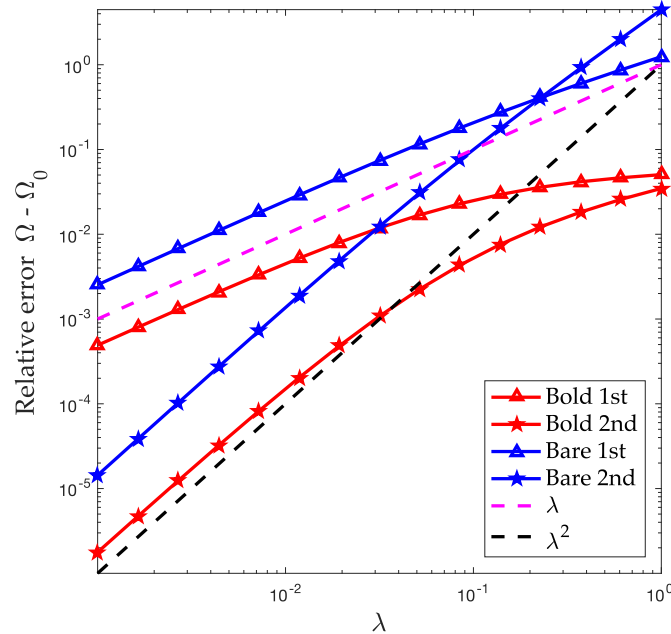


Figure 31: Comparison of different approximation schemes for the free energy.

Observe that the relative errors of the first- and second-order expansions scale as  $\lambda$  and  $\lambda^2$ , respectively, as  $\lambda \rightarrow 0$ . This makes sense because the ‘truncation’ error of these expansions should be of order  $\lambda^2$  and  $\lambda^3$ , respectively, but to obtain the relative error we are dividing by  $\Omega_{\text{exact}} - \Omega_0$ , which is of order  $\lambda$  as  $\lambda \rightarrow 0$ .

Note, however, that the pre-constant of the scaling is more favorable for the bold method in both cases. Furthermore, the bold methods scale more gracefully than their bare counterparts when  $\lambda$  is relatively large.

To demonstrate the simplicity of the implementation, we provide below a MATLAB code for computing the self-consistent Green’s function and the free energy using the first- and second-order bold diagrammatic expansions. The exact solution is evaluated directly using a quadrature code which is omitted here.

```
% luttingerward.m
d = 4;
Phi0 = d*(log(2*pi)+1);
A = [ 2 -1 0 0
      -1 2 -1 0;
```

```

        0 -1  2 -1;
        0  0 -1 2];
Umat = 0.1 * eye(d);

% First order bold diagram
maxiter = 100;
G = inv(A);
for iter = 1 : maxiter
    rho = diag(G);
    Sigma = -1/2*diag(Umat * rho) - (Umat.*G);
    GNew = inv(A-Sigma);
    err = norm(G-GNew)/norm(G);
    if( norm(G-GNew)/norm(G) < 1e-10 ) break; end
    G = GNew;
end
Phi = 1/2*trace(Sigma*G);
Omega = 0.5*(trace(A*G) - log(det(G)) - (Phi + Phi0));
fprintf('Free energy 1st order = %g\n', Omega);

% Second order bold diagram
G = inv(A);
for iter = 1 : maxiter
    rho = diag(G);
    % First order term
    Sigma1 = -1/2*diag(Umat * rho) - (Umat.*G);
    % Ring diagram
    Sigma2 = 1/2 * G.*(Umat * (G.*G) * Umat);
    % Second order exchange diagram
    for i = 1 : d
        for j = 1 : d
            Sigma2(i,j) = Sigma2(i,j) + ...
                (Umat(:,i).*G(:,j))'*G*(Umat(:,j).*G(:,i));
        end
    end
    Sigma = Sigma1 + Sigma2;
    GNew = inv(A-Sigma);
    nrmerr = norm(G-GNew)/norm(G);
    if( norm(G-GNew)/norm(G) < 1e-10 ) break; end
    G = GNew;
end
Phi = 1/2*trace(Sigma1*G) + 1/4*trace(Sigma2*G);

```



```
Omega = 0.5*(trace(A*G) - log(det(G)) - (Phi + Phi0));  
fprintf('Free energy 2nd order = %g\n', Omega);
```

```
OmegaExact = -2.7510737;  
fprintf('Free energy exact      = %g\n', OmegaExact);
```

```
>> luttingerward  
Free energy 1st order = -2.74745  
Free energy 2nd order = -2.75209  
Free energy exact    = -2.75107
```

## Part III

# The Luttinger-Ward formalism

## 1 Introduction

The bold Feynman diagrammatic expansion of many-body perturbation theory (abbreviated MBPT), along with the many practically used methods in quantum chemistry and condensed matter physics that derive from it, can be formally derived from the Luttinger-Ward (LW)<sup>7</sup> formalism [65]. Since its original proposal in 1960, the LW formalism has found widespread applicability [28, 45, 10, 90]. However, the LW formalism and the LW functional are defined only formally, and this shortcoming poses serious questions both in theory and in practice. Indeed, the very existence of the LW functional in the setting of fermionic systems is under debate, with numerical evidence to the contrary appearing in the past few years [54, 32, 103, 42] in the physics community.

In the preceding Part of this dissertation, we provided a self-contained explanation of MBPT in the setting of the *Gibbs model* (alternatively known as the ‘Euclidean lattice field theory’ in the physics literature). In this setting one is interested in the evaluation of the moments of certain Gibbs measures. While the exact computation of such possibly high-dimensional integrals is intractable in general, important exceptions are the Gaussian integrals, i.e., integrals for the moments of a Gaussian measure, which can be evaluated exactly. Perturbing about a reference system given by a Gaussian measure, one can evaluate quantities of interest by a series expansion of Feynman diagrams, which correspond to certain moments of Gaussian measures. For a specific form of quartic interaction that we refer to as the *generalized Coulomb interaction*, such a perturbation theory enjoys a correspondence with the Feynman diagrammatic expansion for the quantum many-body problem with a two-body interaction [77, 2, 1]. The generalized Coulomb interaction is also of interest in its own right and includes, e.g., the (lattice)  $\varphi^4$  interaction [2, 108], as a special case. The combinatorial study of its perturbation theory was the goal of Part II. Nonetheless, the techniques of Part II, and MBPT more broadly, are more generally applicable to various types of field theories and interactions.

The culmination of the developments of Part II is the bold diagrammatic expansion, which is obtained formally via a partial resummation technique which sums possibly divergent series of diagrams to infinite order. Indeed, the main technical contribution of Part II was to place the combinatorial side of this procedure on firm footing. One motivation for Part III is to interpret the bold diagrams analytically, which we accomplish by first constructing the LW formalism. In fact this construction

---

<sup>7</sup>The Luttinger-Ward formalism is also known as the Kadanoff-Baym formalism [7] depending on the context. In this work we always use the former.

is non-perturbative and valid for rather general forms of interaction.

## 1.1 Contributions

Please note that this Part is based on [62] (joint work with Lin Lin). The main contribution of this Part is to establish the LW formalism rigorously for the first time, in the context of Gibbs measures. In this setting, the role of the Green's function is assumed by the two-point correlator.

The construction of the LW functional proceeds via concave duality, in a spirit similar to that of the Levy-Lieb construction in density functional theory [56, 59] at zero temperature and the Mermin functional [76] at finite temperature, as well as the density matrix functional theory developed in [6, 98, 15]. With careful interpretation, this duality gives rise to a one-to-one correspondence between non-interacting and interacting Green's functions. The LW formalism yields a variational interpretation of the Dyson equation. To wit, the free energy can be expressed variationally as a minimum over all physical Green's functions, and the self-consistent solution of the Dyson equation yields its unique global minimizer. We also prove a number of useful properties of the LW functional, such as the transformation rule, the projection rule, and the continuous extension of the LW functional to the boundary of its domain, which can be interpreted as the domain of physical Green's functions. In particular, this last property suggests a novel interpretation of the LW functional as the non-divergent part of the concave dual of the free energy. These results allow us to interpret the appropriate analogs of quantum impurity problems in our simplified setting. In particular, we prove that the self-energy is always a sparse matrix for impurity problems, with nonzero entries appearing only in the block corresponding to the impurity sites. Such a result is at the foundation of numerical approaches such as the dynamical mean field theory (DMFT) [37, 53].

We prove that the bold diagrams for the generalized Coulomb interaction can be obtained as asymptotic series expansions of the LW and self-energy functionals, circumventing the formal strategy of performing resummation to infinite order. The proof of this fact proceeds by proving the existence of such series non-constructively and then employing the combinatorial results of Part II to ensure that the terms of these series are in fact given by the bold diagrams.

Although the bold diagrammatic expansion (evaluated in terms of the interacting Green's function, which is always defined) appears to be applicable in cases where the non-interacting Green's function is ill-defined, we demonstrate that caution should be exercised in practice in such cases. Using a one-dimensional example, we demonstrate that the approximate Dyson equation obtained via a *truncated* bold diagrammatic expansion may yield solutions with large error in the regime of vanishing interaction strength or fail to admit solutions at all.

## 1.2 Outline

In section 2 we review preliminary material and definitions needed to understand the results of this Part.

Section 3 concerns the construction of the LW formalism, beginning with a discussion of the variational formulation of the free energy and the relevant concave duality (section 3.1). This is followed by the introduction of the LW functional and the Dyson equation (section 3.2). Then we introduce several key properties of the LW functional: the transformation rule (section 3.3); the projection rule, accompanied by a discussion of impurity problems (section 3.4); and the continuous extension property (section 3.5). The proof of the continuous extension property, which is the most technically demanding part of the Part, is postponed to section 5, which has its own outline.

Section 4 concerns the bold diagrammatic expansion. In section 4.1 we prove the existence of asymptotic series for the LW functional and the self-energy, and in section 4.2 we relate the coefficients of the former to the latter. Then for the rigorous development of the bold diagrammatic expansion, it only remains at this point to prove that the asymptotic series for the self-energy matches the bold diagrammatic expansion of Part II. This is the most involved task of section 4. In section 4.3, we review the results that we need from Part II in a ‘diagram-free’ way that should be understandable to the reader who has not read Part II, and in section 4.4, we establish the claimed correspondence. Finally, in section 4.5 we illustrate the aforementioned warning about the truncation of the bold diagrammatic series in cases where the non-interacting Green’s function is ill-defined.

Relevant background material on convex analysis and the weak convergence of measures is collected in Appendices C and D, respectively. The proofs of many lemmas are provided in Appendix E, as noted in the text.

## 2 Preliminaries

In this section we discuss some preliminary definitions and notations.

### 2.1 Notation and quantities of interest

Throughout we shall let  $\mathcal{S}^N$ ,  $\mathcal{S}_+^N$ , and  $\mathcal{S}_{++}^N$  denote respectively the sets of symmetric, symmetric positive semidefinite, and symmetric positive definite  $N \times N$  real matrices. For simplicity we restrict our attention to real matrices, though analogous results can be obtained in the complex Hermitian case.

In this Part we will consider Gibbs measures defined by Hamiltonians  $h : \mathbb{R}^N \rightarrow \mathbb{R} \cup \{+\infty\}$  of the form

$$h(x) = \frac{1}{2}x^T Ax + U(x),$$

where  $A \in \mathcal{S}^N$ . The first term represents the quadratic or ‘non-interacting’ part of the Hamiltonian, while the second term,  $U$ , represents the interaction. We define the partition function accordingly as

$$Z[A, U] = \int_{\mathbb{R}^N} e^{-\frac{1}{2}x^T Ax - U(x)} dx. \quad (2.1)$$

For fixed interaction  $U$ , we may think of the partition function of  $A$  alone, i.e., as  $Z : \mathcal{S}^N \rightarrow \mathbb{R}$  sending  $A \mapsto Z[A]$ . In fact we adopt this perspective exclusively for the time being.

The free energy is then defined as a mapping  $\Omega : \mathcal{S}^N \rightarrow \mathbb{R} \cup \{-\infty\}$  via

$$\Omega[A] := -\log Z[A] = -\log \int_{\mathbb{R}^n} e^{-\frac{1}{2}x^T Ax - U(x)} dx, \quad (2.2)$$

We denote the *domain* of  $\Omega$  by

$$\text{dom } \Omega := \{A \in \mathcal{S}^N : \Omega[A] > -\infty\},$$

and the interior of the domain by  $\text{int dom } \Omega$ . As we will see,  $\Omega$  is concave in  $A$ , and this notion of domain is the usual notion from convex analysis (see Appendix C), and it is simply the set of  $A$  such that the integral in Eq. (2.2) is convergent.

For  $A \in \text{int dom } \Omega$ , in fact the integrand in Eq. (2.2) must decay exponentially, hence we can define the two-point correlator (which we call the *Green’s function* by analogy with the quantum many-body literature) in terms of  $A$  via

$$G_{ij}[A] := \frac{1}{Z[A]} \int_{\mathbb{R}^N} x_i x_j e^{-\frac{1}{2}x^T Ax - U(x)} dx,$$

and the integral on the right-hand side is convergent. More compactly, we have a mapping  $G : \text{int dom } \Omega \rightarrow \mathcal{S}_{++}^n$  defined by

$$G[A] := \frac{1}{Z[A]} \int_{\mathbb{R}^N} xx^T e^{-\frac{1}{2}x^T Ax - U(x)} dx. \quad (2.3)$$

It is important to note that  $G[A] \in \mathcal{S}_{++}^n$  for all  $A$ . As we shall see in section 3, this constraint defines the domain of ‘physical’ Green’s functions, in a certain sense. In the discussion below,  $G$  is also called the interacting Green’s function.

In the case of the ‘non-interacting’ Gibbs measure, where  $U \equiv 0$ , all quantities of interest can be computed exactly by straightforward multivariate integration. In particular, letting  $G^0[A] := G[A; 0]$ , we have for  $A \in \text{dom } \Omega = \mathcal{S}_{++}^n$  that

$$G^0[A] = A^{-1}. \quad (2.4)$$

The neatness of this relation is that it motivates the factor of one half included in the quadratic part of the Hamiltonian. We refer to  $G^0[A]$  as the *non-interacting Green’s*

function associated to  $A$ , whenever  $A \in \mathcal{S}_{++}^n$ . Note that for a general interaction  $U$ ,  $\text{int dom } \Omega$  may contain elements not in  $\mathcal{S}_{++}^n$ . For such  $A$  there is an associated (interacting) Green's function but not a non-interacting Green's function.

In general  $G$  can be viewed as the *gradient* of  $\Omega$ , for a suitably defined notion of gradient for functions of symmetric matrices, which we now define:

**Definition 1.** For  $i, j = 1, \dots, N$ , let  $E^{(ij)} \in \mathcal{S}^N$  be defined by  $E_{kl}^{(ij)} = \delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk}$ . For a differentiable function  $f : \mathcal{S}^N \rightarrow \mathbb{R}$ , define the gradient  $\nabla f : \mathcal{S}^N \rightarrow \mathcal{S}^N$  by

$$\nabla_{ij}f = (\nabla f)_{ij} := \lim_{\delta \rightarrow 0} \frac{f(A + \delta \cdot E^{(ij)}) - f(A)}{\delta}.$$

If  $f$  is obtained by restriction from a function  $f : \mathbb{R}^{N \times N} \rightarrow \mathbb{R}$ , then equivalently  $\nabla_{ij}f = \frac{\partial f}{\partial X_{ij}} + \frac{\partial f}{\partial X_{ji}}$ .

Then on  $\text{dom } \Omega$  the gradient map  $\nabla \Omega$  is given by

$$\nabla_{ij}\Omega[A] = \frac{1}{Z[A]} \int x_i x_j e^{-\frac{1}{2}x^T A x - U(x)} dx, \quad (2.5)$$

i.e.,  $G = \nabla \Omega$ , as claimed. The notion of gradient of Definition 1 is natural for our setting in that it yields this relation. However, it may seem a bit awkward when applied to specific computations. Indeed, consider a function  $X \mapsto f(X)$  on  $\mathcal{S}^N$  that is specified by a formula that can be applied to all  $N \times N$  matrices and in which the roles of  $X_{kl}$  and  $X_{lk}$  are the same for all  $l, k$ . For instance, such a formula is given by  $f(X) = \sum_{ij} X_{ij}^2$ . Then the usual matrix derivative of  $f$ , considered as a function on  $N \times N$  matrices, is given by  $\frac{\partial f}{\partial X_{ij}}(X) = 2X_{ij}$ , whereas, viewing  $f$  as a function on  $\mathcal{S}^N$  and with notation as specified in Definition 1, we have  $\nabla_{ij}f(X) = 4X_{ij}$ . More generally in this situation we have  $\nabla_{ij} = 2\frac{\partial}{\partial X_{ij}}$ . Since formulas like this arise from the bold diagrammatic expansion (as discussed in Part II), it is convenient then to further define:

**Definition 2.** For a differentiable function  $f : \mathcal{S}^N \rightarrow \mathbb{R}$ , define the matrix derivative  $\frac{\partial f}{\partial X} : \mathcal{S}^N \rightarrow \mathcal{S}^N$  by

$$\frac{\partial f}{\partial X_{ij}} = \frac{1}{2} \nabla_{ij}f.$$

Moreover, this notion of derivative will yield the relation

$$\Sigma[G] = \frac{\partial \Phi}{\partial G},$$

where  $\Sigma$  is the self-energy and  $\Phi$  is the LW functional, as was foreshadowed in Part II.

## 2.2 Interaction growth conditions

Note that  $\text{dom } \Omega$  depends on the shape of  $U(x)$ . For example, if  $U(x) = 0$ , then  $\text{dom } \Omega = \mathcal{S}_{++}^N$ . If  $U(x) = \sum_{i=1}^N x_i^4$ , then  $\text{dom } \Omega = \mathcal{S}^N$ . Our most basic condition on  $U$  is the following:

**Definition 3** (Weak growth condition). A measurable function  $U : \mathbb{R}^n \rightarrow \mathbb{R}$  satisfies the weak growth condition, if there exists a constant  $C_U$  such that  $U(x) + C_U(1 + \|x\|^2) \geq 0$  for all  $x \in \mathbb{R}^n$ , and  $\text{dom } \Omega$  is an open set.

The weak growth condition of Definition 3 specifies that  $U$  cannot decay to  $-\infty$  faster than quadratically, which ensures in particular that  $\text{dom } \Omega$  is non-empty. The assumption that  $\text{dom } \Omega$  is an open set (i.e.,  $\text{dom } \Omega = \text{int } \text{dom } \Omega$ ) will be used later to ensure that for fixed  $U$  there is a one-to-one correspondence between  $A$  and  $G$  (hence also between non-interacting and interacting Green's functions) over suitable domains.

Note that the condition of Definition 3 is weaker than the condition

$$\frac{1}{2}x^T Ax + U(x) \rightarrow +\infty, \quad \|x\| \rightarrow +\infty \quad (2.6)$$

For instance, if  $N = 2$  and  $U(x) = x_1^4$ , then the weak growth condition is satisfied with  $C_U = 0$ , but Eq. (2.6) is not satisfied for all  $A \in \mathcal{S}^N$ . In fact, when  $U(x)$  only depends on a subset of components of  $x \in \mathbb{R}^n$ , we call the Gibbs model an *impurity model* or *impurity problem*, in analogy with the impurity models of quantum many-body physics [68], and we call the subset of components on which  $U$  depends the *fragment*. The flexibility of the weak growth condition will allow us to rigorously establish the LW formalism for the impurity model. In the setting of the impurity model, the ‘projection rule’ of Proposition 25 then allows us to understand the LW formalism of the impurity model in terms of the lower-dimensional LW formalism of the fragment and to prove a special sparsity pattern of the self-energy.

One of our main results (Theorem 30) is that the LW functional, which is initially defined on the set  $\mathcal{S}_{++}^n$  of physical Green's functions, can in fact be extended continuously to the boundary of  $\mathcal{S}_{++}^n$ , a fact which will not be apparent from the definition of the LW functional. (In fact, this extension shall be specified by an explicit formula involving lower-dimensional LW functionals.) But in order for this result to hold, we need to strengthen the weak growth condition to the following:

**Definition 4** (Strong growth condition). A measurable function  $U : \mathbb{R}^n \rightarrow \mathbb{R}$  satisfies the strong growth condition if, for any  $\alpha \in \mathbb{R}$ , there exists a constant  $b \in \mathbb{R}$  such that  $U(x) + b \geq \alpha \|x\|^2$  for all  $x \in \mathbb{R}^n$ .

Note that the strong growth condition ensures that  $\text{dom } \Omega = \mathcal{S}^N$  and is hence an open set. If  $U$  is a polynomial function of  $x$  and satisfies the strong growth condition, then Eq. (2.6) will also be satisfied.

In Section 5 we will discuss the precise statement and proof of the aforementioned continuous extension property. In addition, a counterexample will be provided in the case where the weak growth condition holds but the strong growth condition does not. In fact, the continuous extension property is also valid for impurity models (which do not satisfy the strong growth condition) via the projection rule (Proposition 25), provided that the interaction satisfies the strong growth condition when restricted to the fragment.

For the generalized Coulomb interaction considered in Part II, i.e.,

$$U(x) = \frac{1}{8} \sum_{i,j=1}^N v_{ij} x_i^2 x_j^2, \quad (2.7)$$

there is a natural condition on the matrix  $v$  that ensures that  $U$  satisfies the strong growth condition, namely that the matrix  $v$  is positive definite. We will simply assume that this holds whenever we refer to the generalized Coulomb interaction. To see that this assumption implies the strong growth condition, first note that  $v \succ 0$  guarantees in particular that  $U$  is a nonnegative polynomial, strictly positive away from  $x = 0$ . Since  $U$  is homogeneous quartic, it follows that  $U \geq C^{-1}|x|^4$  for some constant  $C$  sufficiently large, which evidently implies the strong growth condition. Another sufficient assumption is that the entries of  $v$  are nonnegative and moreover that the diagonal entries are strictly positive.

Our interest in diagrammatic expansions leads us to adopt a further condition on the interaction. To see why this is necessary, recall from Part II that the perturbation about a non-interacting theory ( $U \equiv 0$ ) involves integrals such as

$$\int U(x) e^{-\frac{1}{2}x^T A x} dx,$$

which is clearly undefined if, e.g.,  $U(x) = e^{x^4}$ . In most applications of interest,  $U(x)$  is only of polynomial growth, but it is sufficient to assume growth that is at most exponential in the sense of Assumption 5, which is actually only needed in section 4 for our consideration of the bold diagrammatic expansion.

**Assumption 5** (At-most-exponential growth). *In this section, we assume that there exist constants  $B, C > 0$  such that  $|U(x)| \leq B e^{C\|x\|}$  for all  $x \in \mathbb{R}^n$ .*

Further technical reasons for this assumption will become clear in section 4.

### 2.3 Measures and entropy: notation and facts

Let  $\mathcal{M}$  be the space of probability measures on  $\mathbb{R}^n$  (equipped with the Borel  $\sigma$ -algebra), let  $\mathcal{M}_2 \subset \mathcal{M}$  be the subset of probability measures with moments up to second order, and let  $\lambda$  denote the Lebesgue measure on  $\mathbb{R}^n$ . For notational convenience we define a mapping that takes the second-order moments of a probability measure:



**Definition 6.** Define  $\mathcal{G} : \mathcal{M}_2 \rightarrow \mathcal{S}_+^N$  by  $\mathcal{G}(\mu) = \int xx^T d\mu$ . Writing  $\mathcal{G} = (\mathcal{G}_{ij})$ , we equivalently have  $\mathcal{G}_{ij}(\mu) = \int x_i x_j d\mu$ .

Therefore if  $\mu$  is defined via a density

$$d\mu = \rho(x) dx, \quad \text{where } \rho(x) = \frac{1}{Z[A]} e^{-\frac{1}{2}x^T A x - U(x)},$$

then  $\mathcal{G}(\mu) = G[A]$ .

We also denote by

$$\text{Cov}(\mu) = \int xx^T d\mu - \left( \int x d\mu \right) \left( \int x d\mu \right)^T$$

the covariance matrix of  $\mu$ .

For  $\mu \in \mathcal{M}$ , let  $H$  denote the (differential) entropy

$$H(\mu) = \begin{cases} - \int \log \frac{d\mu}{d\lambda} d\mu, & \mu \ll \lambda \\ -\infty, & \text{otherwise} \end{cases} \quad (2.8)$$

where  $\frac{d\mu}{d\lambda}$  denotes the Radon-Nikodym derivative (i.e., the probability density function of  $\mu$  with respect to the Lebesgue measure  $\lambda$ ) whenever  $\mu \ll \lambda$  (i.e., whenever  $\mu$  is absolutely continuous with respect to the Lebesgue measure). We will often refer to the differential entropy as the entropy for convenience.

For  $\mu, \nu \in \mathcal{M}$ , define the relative entropy  $H_\nu(\mu)$  via

$$H_\nu(\mu) = \begin{cases} - \int \log \frac{d\mu}{d\nu} d\mu, & \mu \ll \nu \\ -\infty, & \text{otherwise.} \end{cases} \quad (2.9)$$

Note carefully the sign convention.<sup>8</sup> The integral in (2.9) is well-defined with values in  $\mathbb{R} \cup \{-\infty\}$  for all  $\mu, \nu \in \mathcal{M}$ .

We now record some useful properties of the relative entropy.

**Fact 7.** For fixed  $\nu \in \mathcal{M}$ ,  $H_\nu$  is non-positive and strictly concave on  $\mathcal{M}$ , and  $H_\nu(\mu) = 0$  if and only if  $\mu = \nu$ . Moreover  $H_\nu$  is upper semi-continuous with respect to the topology of weak convergence; i.e., if the sequence  $\mu_k \in \mathcal{M}$  converges weakly to  $\mu \in \mathcal{M}$ , then  $\limsup_{k \rightarrow \infty} H_\nu(\mu_k) \leq H_\nu(\mu)$ .

*Proof.* For proofs see [88]. □

By contrast to the relative entropy, the differential entropy suffers from two analytical nuisances.

---

<sup>8</sup>Our relative entropy is then the negative of the Kullback-Leibler divergence, i.e.,  $H_\nu(\mu) = -D_{\text{KL}}(\mu \parallel \nu)$ .

First, in the definition of the entropy in (2.8), the entropy may actually fail to be defined for some measures (which simultaneously concentrate too much in some area and fail to decay fast enough at infinity, so the negative and positive parts of the integral are  $-\infty$  and  $+\infty$ , respectively, and the Lebesgue integral is ill-defined). However, Lemma 8 states that when we restrict to  $\mathcal{M}_2$ , the integral cannot have an infinite positive part and is well-defined.

**Lemma 8.** *For  $\mu \in \mathcal{M}_2$ , if  $\mu \ll \lambda$ , then the integral in (2.8) exists (in particular, the positive part of the integrand has finite integral) and moreover*

$$H(\mu) \leq \frac{1}{2} \log \left( (2\pi e)^N \det \text{Cov}(\mu) \right) \leq \frac{1}{2} \log \left( (2\pi e)^N \det \mathcal{G}(\mu) \right),$$

with possibly  $H(\mu) = -\infty$ . The first inequality is satisfied with equality if and only if  $\mu$  is a Gaussian measure with a positive definite covariance matrix. The second inequality is satisfied with equality if and only if  $\mu$  has mean zero.

Note that Lemma 8 also entails a useful bound on the entropy in terms of the second moments, as well as the classical fact that Gaussian measures are the measures of maximal entropy subject to second-order moment constraints.

The second analytical nuisance of the differential entropy is that we do not have the same semi-continuity guarantee as we have for the relative entropy in Fact 7. However, control on second moments allows a semi-continuity result that will suffice for our purposes.

**Lemma 9.** *Assume that  $\mu_j \in \mathcal{M}_2$  weakly converge to  $\mu \in \mathcal{M}$ , and that there exists a constant  $C$  such that  $\mathcal{G}(\mu_j) \preceq C \cdot I_N$  for all  $j$ . Then  $\limsup_{j \rightarrow \infty} H(\mu_j) \leq H(\mu)$ .*

*Remark 10.* In other words, the entropy is upper semi-continuous with respect to the topology of weak convergence on any subset of probability measures with uniformly bounded second moments. The subtle difference between the statements in Fact 7 and Lemma 9 is due to the fact that the Lebesgue measure  $\lambda \notin \mathcal{M}$ .

The proofs of Lemmas 8 and 9 are given in appendix E.

Finally we record the classical fact that subject to marginal constraints, the entropy is maximized by a product measure. In the statement and throughout this work, ‘ $\#$ ’ denotes the pushforward operation on measures.

**Fact 11.** *Suppose  $p < N$  and let  $\pi_1 : \mathbb{R}^n \rightarrow \mathbb{R}^p$  and  $\pi_2 : \mathbb{R}^n \rightarrow \mathbb{R}^{N-p}$  to be the projections onto the first  $p$  and last  $N-p$  components, respectively. Then for  $\mu \in \mathcal{M}_2$ ,  $H(\mu) \leq H(\pi_1\#\mu) + H(\pi_2\#\mu)$ .*

*Remark 12.* Note that  $\pi_1\#\mu$  and  $\pi_2\#\mu$  are the marginal distributions of  $\mu$  with respect to the product structure  $\mathbb{R}^n = \mathbb{R}^p \times \mathbb{R}^{N-p}$ .

See appendix E for a short proof.

### 3 Luttinger-Ward formalism

This section is organized as follows. In section 3.1, we provide a variational expression for the free energy via the classical Gibbs variational principle. For fixed  $U$ , this allows us to identify the Legendre dual of  $\Omega[A]$ , denoted by  $\mathcal{F}[G]$ , and to establish a bijection between  $A$  and the interacting Green's function  $G$ . In section 3.2, we define the Luttinger-Ward functional and show that the Dyson equation can be naturally derived by considering the first-order optimality condition associated to the minimization problem in the variational expression for the free energy. Then we prove that the LW functional satisfies a number of desirable properties. First, in section 3.3 we prove the transformation rule, which relates a change of the coordinates of the interaction with an appropriate transformation of the Green's function. The transformation rule leads to the projection rule in section 3.4, which implies the sparsity pattern of the self-energy for the impurity problem. Up until this point we assume only that  $U$  satisfy the weak growth condition. Then in section 3.5 we motivate and state our result that the LW functional is continuous up to the boundary of  $\mathcal{S}_{++}^n$ , for which we need the assumption that  $U$  satisfies the strong growth condition. The proof (as well as a counterexample demonstrating that weak growth is not sufficient) is deferred to section 5. Throughout we defer the proofs of some technical lemmas to Appendix E. Moreover we will invoke the language of convex analysis following Rockafellar [91] and Rockafellar and Wets [92]. See Appendix C for further background and details.

#### 3.1 Variational formulation of the free energy

The main result in this subsection is given by Theorem 13.

**Theorem 13** (Variational structure). *For  $U$  satisfying the weak growth condition, the free energy can be expressed variationally via the constrained minimization problem*

$$\Omega[A] = \inf_{G \in \mathcal{S}_+^N} \left( \frac{1}{2} \text{Tr}[AG] - \mathcal{F}[G] \right), \quad (3.1)$$

where

$$\mathcal{F}[G] := \sup_{\mu \in \mathcal{G}^{-1}(G)} \left[ H(\mu) - \int U \, d\mu \right] \quad (3.2)$$

is the concave conjugate of  $\Omega[A]$  with respect to the inner product  $\langle A, G \rangle = \frac{1}{2} \text{Tr}[AG]$ . (Note that by convention  $\mathcal{F}[G] = -\infty$  whenever  $\mathcal{G}^{-1}(G)$  is empty, i.e., whenever  $G \in \mathcal{S}^N \setminus \mathcal{S}_+^N$ .) Moreover  $\Omega$  and  $\mathcal{F}$  are smooth and strictly concave on their respective domains  $\text{dom } \Omega$  and  $\mathcal{S}_{++}^n$ . The mapping  $G[A] := \nabla \Omega[A]$  is a bijection  $\text{dom } \Omega \rightarrow \mathcal{S}_{++}^N$ , with inverse given by  $A[G] := \nabla \mathcal{F}[G]$ .

We first record some technical properties of  $\Omega$  in Lemma 2.

**Lemma 14.**  *$\Omega$  is an upper semi-continuous, proper (hence closed) concave function. Moreover,  $\Omega$  is strictly concave and  $C^\infty$ -smooth on  $\text{dom } \Omega$ .*

*Remark 15.* Recall that a function  $f$  on a metric space  $X$  is upper semi-continuous if for any sequence  $x_k \in X$  converging to  $x$ , we have  $\limsup_{k \rightarrow \infty} f(x_k) \leq f(x)$ .

We now turn to exploring the concave (or Legendre-Fenchel) duality associated to  $\Omega$ . The following lemma, a version of the classical Gibbs variational principle [88], is the first step toward identifying the dual of  $\Omega$ .

**Lemma 16.** *For any  $A \in \mathcal{S}^N$ ,*

$$\Omega[A] = \inf_{\mu \in \mathcal{M}_2} \left[ \int \left( \frac{1}{2} x^T A x + U(x) \right) d\mu(x) - H(\mu) \right]. \quad (3.3)$$

*If  $A \in \text{dom } \Omega$ , the infimum is uniquely attained at  $d\mu(x) = \frac{1}{Z[A]} e^{-\frac{1}{2} x^T A x - U(x)} dx$ .*

*Remark 17.* One might wonder whether the infimum in (3.3) can be taken over all of  $\mathcal{M}$ . Note that if  $\mu$  does not have a second moment, it is possible to have both  $H(\mu) = +\infty$  and  $\int \left( \frac{1}{2} x^T A x + U(x) \right) d\mu(x) = +\infty$ , so the expression in brackets is of the indeterminate form  $\infty - \infty$ . The restriction to  $\mu \in \mathcal{M}_2$  takes care of this problem because Lemma 8 guarantees that  $H(\mu) < +\infty$ , and by the weak growth condition, the other term in the infimum must be either finite or  $+\infty$ . Moreover,  $\mathcal{M}_2$  is still large enough to contain the minimizer, and restricting our attention to measures with finite second-order moments will be convenient in later developments.

From the previous lemma we can split up the infimum in (3.3) and obtain

$$\Omega[A] = \inf_{G \in \mathcal{S}_+^N} \inf_{\mu \in \mathcal{G}^{-1}(G)} \left[ \int \left( \frac{1}{2} x^T A x + U(x) \right) d\mu(x) - H(\mu) \right].$$

Since  $\int x^T A x d\mu = \text{Tr}[\mathcal{G}(\mu)A]$ , it follows that

$$\Omega[A] = \inf_{G \in \mathcal{S}_+^N} \left( \frac{1}{2} \text{Tr}[AG] + \inf_{\mu \in \mathcal{G}^{-1}(G)} \left[ \int U d\mu - H(\mu) \right] \right).$$

This proves Eq. (3.1) of Theorem 13 using the definition of  $\mathcal{F}[G]$  in Eq. (3.2).

*Remark 18.* For the perspective of the large deviations theory, we comment that the construction of  $\mathcal{F}$  from the entropy may be recognizable by analogy to the contraction principle [88]. Indeed, the expression  $\int U d\mu - H(\mu)$  is equal (modulo a constant offset) to  $-H_{\nu_U}(\mu)$ , where  $\nu_U$  is the measure with density proportional to  $e^{-U}$ . If one considers i.i.d. sampling from the probability measure  $\nu_U$ , by Sanov's theorem  $-H_{\nu_U}$  is the corresponding large deviations rate function for the empirical measure. The rate function for the second-order moment matrix (i.e.,  $-\mathcal{F}$ , modulo constant offset) is obtained via the contraction principle applied to the mapping  $\mu \mapsto \mathcal{G}(\mu)$ . This is analogous to the procedure by which one obtains Cramér's theorem from Sanov's theorem via application of the contraction principle to a map that maps  $\mu$  to its mean [88].

Now we record some technical facts about  $\mathcal{F}$  in Lemma 19, which demonstrates in particular that  $\mathcal{F}$  diverges (at least) logarithmically at the boundary  $\partial\mathcal{S}_+^N = \mathcal{S}_+^N \setminus \mathcal{S}_{++}^N$ .

**Lemma 19.**  $\mathcal{F}$  is finite on  $\mathcal{S}_{++}^n$  and  $-\infty$  elsewhere. Moreover,

$$\mathcal{F}[G] \leq \frac{1}{2} \log [(2\pi e)^N \det G] + C_U(1 + \text{Tr } G)$$

for all  $G \in \mathcal{S}_{++}^n$ .

Define

$$\Psi[\mu] := H(\mu) - \int U \, d\mu,$$

so  $\mathcal{F}[G] = \sup_{\mu \in \mathcal{G}^{-1}(G)} \Psi[\mu]$ . By the concavity of the entropy,  $\Psi$  is concave on  $\mathcal{M}_2$ . Thus, given  $G$ , we can in principle solve a concave maximization problem over  $\mu \in \mathcal{M}$  to find  $\mathcal{F}[G]$ , with the linear constraint  $\mu \in \mathcal{G}^{-1}(G)$ . Moreover, this variational representation of  $\mathcal{F}$  in terms of the concave function  $\Psi$  is enough to establish the concavity of  $\mathcal{F}$  by abstract considerations. This and other properties of  $\mathcal{F}$  are collected in the following.

**Lemma 20.**  $\mathcal{F}$  is an upper semi-continuous, proper (hence closed) concave function on  $\mathcal{S}^N$ .

Now Eq. (3.1) states precisely that  $\Omega$  is the *concave conjugate* of  $\mathcal{F}$  with respect to the inner product  $\langle A, G \rangle = \frac{1}{2} \text{Tr}[AG]$ , and accordingly we write  $\Omega = \mathcal{F}^*$ . Since  $\mathcal{F}$  is concave and closed, we have by Theorem 19 that  $\mathcal{F} = \mathcal{F}^{**} = \Omega^*$ , i.e.,  $\mathcal{F}$  and  $\Omega$  are concave duals of one another. Thus we expect that  $\nabla\mathcal{F}$  and  $\nabla\Omega$  are inverses of one another, but to make sense of this claim we need to establish the differentiability of  $\mathcal{F}$ . We collect this and other desirable properties of  $\mathcal{F}$  in the following:

**Lemma 21.**  $\mathcal{F}$  is  $C^\infty$ -smooth and strictly concave on  $\mathcal{S}_{++}^n$ .

Then Theorem 20 guarantees that  $\nabla\Omega$  is a bijection from  $\text{dom } \Omega \rightarrow \mathcal{S}_{++}^n$  with its inverse given by  $\nabla\mathcal{F}$ . This completes the proof of Theorem 13.

Finally, following Lemma 16, together with the splitting of (3.3) and the  $A \leftrightarrow G$  correspondence of Theorem 13, we observe that the supremum in (3.2) is attained uniquely at the measure  $d\mu := \frac{1}{Z[A[G]]} e^{-\frac{1}{2}x^T A[G]x - U(x)} dx$ .

### 3.2 The Luttinger-Ward functional and the Dyson equation

According to Lemma 19,  $\mathcal{F}$  should blow up at least logarithmically as  $G$  approaches the boundary of  $\mathcal{S}_{++}^N$ . Remarkably, we can explicitly separate the part that accounts for the blowup of  $\mathcal{F}$  at the boundary. In fact, subtracting away this part is how we define the Luttinger-Ward (LW) functional for the Gibbs model. We will see in this subsection that the definition of the Luttinger-Ward functional can also be motivated by the stipulation that its gradient (the self-energy) should satisfy the Dyson equation.

Consider for a moment the case in which  $U \equiv 0$ , so

$$\mathcal{F}[G] = \sup_{\mu \in \mathcal{G}^{-1}(G)} \left[ H(\mu) - \int U \, d\mu \right] = \sup_{\mu \in \mathcal{G}^{-1}(G)} H(\mu).$$

The random variable  $X$  achieving the maximum entropy subject to  $\mathbb{E}[X_i X_j] = G_{ij}$  follows a Gaussian distribution, i.e.,  $X \sim \mathcal{N}(0, G)$ . It follows that

$$\mathcal{F}[G] = \frac{1}{2} \log \left( (2\pi e)^N \det G \right) = \frac{1}{2} \text{Tr}[\log(G)] + \frac{N}{2} \log(2\pi e).$$

This motivates, for general  $U$ , the consideration of the *Luttinger-Ward functional*

$$\Phi[G] := 2\mathcal{F}[G] - \text{Tr}[\log(G)] - N \log(2\pi e). \quad (3.4)$$

For non-interacting systems,  $\Phi[G] \equiv 0$  by construction.

Now we turn to establishing the Dyson equation. Theorem 13 shows that for  $A \in \text{dom } \Omega$ , the minimizer  $G^*$  in (3.1) satisfies  $A = \nabla \mathcal{F}[G^*] = A[G^*]$ , so the minimizer is  $G^* = G[A]$ . Recall

$$\mathcal{F}[G] = \frac{1}{2} \text{Tr}[\log(G)] + \frac{1}{2} \Phi[G] + \frac{1}{2} N \log(2\pi e).$$

Taking gradients and plugging into  $A = \nabla \mathcal{F}[G^*]$  yields

$$0 = A - (G^*)^{-1} - \frac{1}{2} \nabla \Phi[G^*].$$

Define the self-energy  $\Sigma$  as a functional of  $G$  by  $\Sigma[G] := \frac{1}{2} \nabla \Phi[G] = \frac{\partial \Phi}{\partial G}[G]$ . Then we have established that for  $G = G[A]$ ,

$$G^{-1} = A - \Sigma[G]. \quad (3.5)$$

Moreover, by the strict concavity of  $\mathcal{F}$ ,  $G = G[A]$  is the *unique*  $G$  solving (3.5).

Eq. (3.5) is in fact the Dyson equation as in section 3.8 of Part II. To see this, recall from Eq. (2.4) that the non-interacting Green's function  $G^0$  is given by  $G^0 = A^{-1}$ , so we have

$$G^{-1} = (G^0)^{-1} - \Sigma[G].$$

Left- and right-multiplying by  $G^0$  and  $G$ , respectively, and then rearranging, we obtain

$$G = G^0 + G^0 \Sigma[G] G.$$

However, Eq. (2.4) requires  $G^0$  to be well defined, i.e.,  $A \in \mathcal{S}_{++}^N$ . On the other hand, the Dyson equation (3.5) derived from the LW functional does not rely on this assumption and makes sense for all  $A \in \text{dom } \Omega$ . Nonetheless, if for fixed  $A$  one seeks to approximately solve the Dyson equation for  $G$  by inserting an ansatz for the self-energy obtained from many-body perturbation theory, one must be wary in the case that  $A \notin \mathcal{S}_{++}^n$ ; see section 4.5.

### 3.3 Transformation rule for the LW functional

Though the dependence of the Luttinger-Ward functional on the interaction  $U$  was only implicit in the previous section, we now explicitly consider this dependence, including it in our notation as  $\Phi[G, U]$ . The same convention will be followed for other functionals without comment. Proposition 22 relates a transformation of the interaction with a corresponding transformation of the Green's function.

**Proposition 22** (Transformation rule). *Let  $G \in \mathcal{S}_{++}^N$ ,  $U$  be an interaction satisfying the weak growth condition. Let  $T$  denote an invertible matrix in  $\mathbb{R}^{N \times N}$ , as well as the corresponding linear transformation  $\mathbb{R}^n \rightarrow \mathbb{R}^n$ . Then*

$$\Phi[TGT^*, U] = \Phi[G, U \circ T].$$

*Proof.* For  $G \in \mathcal{S}_{++}^n$ , note that the supremum in (3.2) can be restricted to the set of  $\mu \in \mathcal{G}^{-1}(G)$  that have densities with respect to the Lebesgue measure. (Indeed, for any  $\mu \in \mathcal{M}_2$  that does not have a density,  $H(\mu) - \int U \, d\mu = -\infty$ .) Then observe

$$\begin{aligned} \Phi[G, U] &= -N \log(2\pi e) - \log \det G + 2 \sup_{\mu \in \mathcal{G}^{-1}(G)} \left[ H(\mu) - \int U \, d\mu \right] \\ &= -N \log(2\pi e) - \log \det G - 2 \inf_{\{\rho: \rho \, dx \in \mathcal{G}^{-1}(G)\}} \left[ \int (\log \rho + U) \, \rho \, dx \right] \\ &= -N \log(2\pi e) - 2 \inf_{\{\rho: \rho \, dx \in \mathcal{G}^{-1}(G)\}} \left[ \int (\log [(\det G)^{1/2} \rho] + U) \, \rho \, dx \right]. \end{aligned}$$

Going forward we will denote  $C := -N \log(2\pi e)$ .

Then for  $T$  invertible, we have

$$\Phi[TGT^*, U] = C - 2 \inf_{\rho \, dx \in \mathcal{G}^{-1}(TGT^*)} \left[ \int (\log [(\det G)^{1/2} \cdot |\det T| \cdot \rho] + U) \, \rho \, dx \right].$$

Now observe by changing variables that

$$\{\rho : \rho \, dx \in \mathcal{G}^{-1}(TGT^*)\} = \{|\det T|^{-1} \cdot \rho \circ T^{-1} : \rho \, dx \in \mathcal{G}^{-1}(G)\}.$$

Therefore

$$\begin{aligned} \Phi[TGT^*, U] &= C - 2 \inf_{\rho \, dx \in \mathcal{G}^{-1}(G)} \left[ |\det T|^{-1} \right. \\ &\quad \left. \int (\log [(\det G)^{1/2} \cdot \rho \circ T^{-1}] + U) \, \rho \circ T^{-1} \, dx \right] \\ &= C - 2 \inf_{\rho \, dx \in \mathcal{G}^{-1}(G)} \left[ \int (\log [(\det G)^{1/2} \cdot \rho] + U \circ T) \, \rho \, dx \right] \\ &= \Phi[G, U \circ T], \end{aligned}$$

as was to be shown. □

*Remark 23.* Since  $T$  is real, the Hermite conjugation  $T^*$  is the same as the matrix transpose, and this is used simply to avoid the notation  $T^T$ .

From the transformation rule we have the following corollary:

**Corollary 24.** *Let  $G \in \mathcal{S}_{++}^N$ , and consider an interaction  $U$  which is a homogeneous polynomial of degree 4 satisfying the weak growth condition. For  $\lambda > 0$ , we have*

$$\Phi[\lambda G, U] = \Phi[G, \lambda^2 U].$$

### 3.4 Impurity problems and the projection rule

For the impurity problem, the interaction only depends on a subset of the variables  $x_1, \dots, x_N$ , namely the fragment. In such a case, the Luttinger-Ward functional can be related to a lower-dimensional Luttinger-Ward functional corresponding to the fragment. This relation, called the projection rule, is given in Proposition 25 below. In the notation, we will now explicitly indicate the dimension  $d$  of the state space associated with the Luttinger-Ward functional via subscript as in  $\Phi_d[G, U]$ , since we will be considering functionals for state spaces of different dimensions. We will follow the same convention for other functionals without comment.

Before we state the projection rule, we record some remarks on the domain of  $\Omega$  and growth conditions in the context of impurity problems. Suppose that the interaction  $U$  depends only on  $x_1, \dots, x_p$ , where  $p \leq N$ , so  $U$  can alternatively be considered as a function on  $\mathbb{R}^p$ . Notice that even if  $U$  satisfies the strong growth condition as a function on  $\mathbb{R}^p$ , it is of course *not* true that  $\text{dom}(\Omega_N[\cdot, U]) = \mathcal{S}^N$ . As mentioned above, this provides a natural reason to consider interactions that do not grow fast in all directions and motivates the generality of our previous considerations.

In fact, for

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{12}^T & A_{22} \end{pmatrix},$$

one can show by Fubini's theorem, integrating out the last  $N - p$  variables in (2.2), that  $A \in \text{dom}(\Omega_N[\cdot, U])$  if and only if both

$$A_{22} \in \mathcal{S}_{++}^{N-p} \text{ and } A_{11} - A_{12}A_{22}^{-1}A_{12}^T \in \text{dom}(\Omega_p[\cdot, U]).$$

Moreover, one can show that for such  $A$ ,

$$\Omega_N[A, U] = \Omega_p[A_{11} - A_{12}A_{22}^{-1}A_{12}^T, U] + \frac{1}{2} \log((2\pi)^{p-N} \det A_{22}).$$

Therefore, if  $\text{dom}(\Omega_p[\cdot, U(\cdot, 0)])$  is open, then so is  $\text{dom}(\Omega_N[\cdot, U])$ . It follows that if  $U$  satisfies the weak growth condition as a function on  $\mathbb{R}^p$ , then  $U$  also satisfies the weak growth condition as a function on  $\mathbb{R}^N$ .



**Proposition 25** (Projection rule). *Let  $p \leq N$ . Suppose that  $U$  depends only on  $x_1, \dots, x_p$  and satisfies the weak growth condition. Hence we can think of  $U$  as a function on both  $\mathbb{R}^n$  and  $\mathbb{R}^p$ . Then for  $G \in \mathcal{S}_{++}^n$ ,*

$$\Phi_N[G, U] = \Phi_p[G_{11}, U],$$

where  $G_{11}$  is the upper-left  $p \times p$  block of  $G$ .

*Remark 26.* If  $U$  can be made to depend only on  $p \leq N$  variables by linearly changing variables, then we can use the projection rule in combination with the transformation rule (Proposition 22) to reveal the relationship with a lower-dimensional Luttinger-Ward functional, though we do not make this explicit here with a formula.

**Corollary 27.** *Let  $p \leq N$ , and  $P$  be the orthogonal projection onto the subspace  $\text{span}\{e_1^{(N)}, \dots, e_p^{(N)}\}$ . Suppose that  $U(\cdot, 0)$  satisfies the weak growth condition. Then for  $G \in \mathcal{S}_{++}^n$ ,*

$$\Phi_N[G, U \circ P] = \Phi_p[G_{11}, U(\cdot, 0)],$$

where  $G_{11}$  is the upper-left  $p \times p$  block of  $G$ .

*Proof.* (Of Proposition 25.) First we observe that we can assume that  $G$  is block-diagonal. To see this, let  $G \in \mathcal{S}_{++}^n$ , and write

$$G = \begin{pmatrix} G_{11} & G_{12} \\ G_{12}^T & G_{22} \end{pmatrix}.$$

Then block Gaussian elimination reveals that

$$G = \begin{pmatrix} I & 0 \\ G_{12}^T G_{11}^{-1} & I \end{pmatrix} \begin{pmatrix} G_{11} & 0 \\ 0 & G_{22} - G_{12}^T G_{11}^{-1} G_{12} \end{pmatrix} \begin{pmatrix} I & G_{11}^{-1} G_{12} \\ 0 & I \end{pmatrix}.$$

Define

$$T := \begin{pmatrix} I & 0 \\ G_{12}^T G_{11}^{-1} & I \end{pmatrix}, \quad \tilde{G} := \begin{pmatrix} G_{11} & 0 \\ 0 & G_{22} - G_{12}^T G_{11}^{-1} G_{12} \end{pmatrix},$$

so  $G = T\tilde{G}T^*$ . Then by the transformation rule, we have

$$\Phi_N[G, U] = \Phi_N[\tilde{G}, U \circ T] = \Phi_N[\tilde{G}, U],$$

where the last equality uses the fact that  $U$  depends only on the first  $p$  arguments, which are unchanged by the transformation  $T$ .

Since  $\tilde{G}$  is block-diagonal with the same upper-left block as  $G$ , we have reduced to the block-diagonal case, as claimed, so now assume that  $G \in \mathcal{S}_{++}^n$  with

$$G = \begin{pmatrix} G_{11} & 0 \\ 0 & G_{22} \end{pmatrix}.$$

Recall the expression for  $\mathcal{F}_N$ :

$$\mathcal{F}_N[G, U] = \sup_{\mu \in \mathcal{G}_N^{-1}(G)} \left[ H(\mu) - \int U \, d\mu \right].$$

Next define  $\pi_1 : \mathbb{R}^n \rightarrow \mathbb{R}^p$  and  $\pi_2 : \mathbb{R}^n \rightarrow \mathbb{R}^{N-p}$  to be the projections onto the first  $p$  and last  $N - p$  components, respectively. Then with ‘#’ denoting the pushforward operation on measures,  $\pi_1\#\mu$  and  $\pi_2\#\mu$  are the marginals of  $\mu$  with respect to the product structure  $\mathbb{R}^n = \mathbb{R}^p \times \mathbb{R}^{N-p}$ . Now recall Fact 11, in particular the inequality  $H(\mu) \leq H(\pi_1\#\mu) + H(\pi_2\#\mu)$ . Also note that if  $\mu \in \mathcal{G}_N^{-1}(G)$ , then  $\pi_1\#\mu \in \mathcal{G}_p^{-1}(G_{11})$  and  $\pi_2\#\mu \in \mathcal{G}_{N-p}^{-1}(G_{22})$ . Finally observe that since  $U$  depends only on the first  $p$  arguments,  $\int U \, d\mu = \int U \, d(\pi_1\#\mu)$  for any  $\mu$ . Therefore

$$\begin{aligned} \mathcal{F}_N[G, U] &\leq \sup_{\mu \in \mathcal{G}_N^{-1}(G)} \left[ H(\pi_1\#\mu) + H(\pi_2\#\mu) - \int U \, d(\pi_1\#\mu) \right] \\ &\leq \sup_{\mu_1 \in \mathcal{G}_p^{-1}(G_{11})} \left[ H(\mu_1) - \int U \, d\mu_1 \right] + \sup_{\mu_2 \in \mathcal{G}_{N-p}^{-1}(G_{22})} [H(\mu_2)] \\ &= \mathcal{F}_p[G_{11}, U] + \frac{1}{2} \log((2\pi e)^{N-p} \det G_{22}). \end{aligned}$$

Since  $\det G = \det G_{11} \det G_{22}$ , it follows that

$$\Phi_N[G, U] \leq \Phi_p[G_{11}, U].$$

For the reverse inequality, let  $\mu_1$  be arbitrary in  $\mathcal{G}_p^{-1}(G_{11})$ , and consider  $\mu := \mu_1 \times \mu_2$ , where  $\mu_2$  is given by the normal distribution with mean zero and covariance  $G_{22}$ . Then

$$\mathcal{F}_N[G, U] \geq H(\mu) - \int U \, d\mu = H(\mu_1) - \int U \, d\mu_1 + \frac{1}{2} \log((2\pi e)^{N-p} \det G_{22}).$$

Since  $\mu_1$  is arbitrary in  $\mathcal{G}_p^{-1}(G_{11})$ , it follows by taking the supremum over  $\mu_1$  that

$$\mathcal{F}_N[G, U] \geq \mathcal{F}_p[G_{11}, U] + \frac{1}{2} \log((2\pi e)^{N-p} \det G_{22}),$$

which implies

$$\Phi_N[G, U] \geq \Phi_p[G_{11}, U].$$

□

*Remark 28.* The proof suggests that for  $U$  depending only on the first  $p$  arguments and  $G$  block-diagonal, the supremum in the definition of  $\mathcal{F}$  is attained by a product measure, which is perhaps not surprising. The proof also suggests, however, that for such  $U$  and general  $G$ , the supremum is attained by taking a product measure and then ‘correlating’ it via the transformation  $T$ .

For the impurity problem, Proposition 25 immediately implies that the self-energy has a particular sparsity pattern:

**Corollary 29.** *Let  $p \leq N$  and suppose that  $U$  (satisfying the weak growth condition) depends only on  $x_1, \dots, x_p$ . Then*

$$\Sigma_N[G, U] = \begin{pmatrix} \Sigma_p[G_{11}, U] & 0 \\ 0 & 0 \end{pmatrix}.$$

For example, consider  $U(x) = \frac{1}{8} \sum_{ijkl} v_{ij} x_i^2 x_j^2$ . Here the stipulation that  $U$  depend only on the first  $p$  arguments corresponds to the stipulation that  $v_{ij} = 0$  unless  $i, j \leq p$ . For such an interaction, in the bold diagrammatic expansion for  $\Phi$  and  $\Sigma$ , any term in which  $G_{ij}$  appears will be zero unless  $i, j \leq p$ . This is a non-rigorous perturbative explanation of the fact that  $\Phi$  depends only on the upper-left block of  $G$ , which in turn explains the sparsity structure of  $\Sigma$ , as well as the fact that  $\Sigma$  also depends only on the upper-left block of  $G$ . However, the developments of this section apply to interactions  $U$  of far greater generality and which may indeed be non-polynomial, hence not admitting of a bold diagrammatic expansion.

### 3.5 Continuous extension of the LW functional to the boundary

The discussion in this subsection is only heuristic, and the proofs of the theorems stated here are deferred to section 5.

Now in section 3.1 we saw that the functional  $\mathcal{F}[G]$  diverges at the boundary  $\partial\mathcal{S}_+^N = \mathcal{S}_+^N \setminus \mathcal{S}_{++}^N$ . On the other hand, the projection rule together with the transformation rule, motivates the formula by which we can extend  $\Phi$  continuously up to the boundary  $\partial\mathcal{S}_+^N$ .

Indeed, suppose that  $T^{(j)} \rightarrow P$ , where  $T^{(j)}$  is invertible and  $P$  is the orthogonal projection onto the first  $p$  components, as in Corollary 27. Then for  $G \in \mathcal{S}_{++}^n$ ,

$$\Phi_N[T^{(j)}G(T^{(j)})^*, U] = \Phi_N[G, U \circ T^{(j)}].$$

By naively taking limits of both sides, we expect that

$$\Phi_N[PGP, U] = \Phi_N[G, U \circ P]$$

where  $G_{11}$  is the upper-left  $p \times p$  block of  $G$ . Then by the projection rule we expect

$$\Phi_N \left[ \begin{pmatrix} G_{11} & 0 \\ 0 & 0 \end{pmatrix}, U \right] = \Phi_p[G_{11}, U(\cdot, 0)],$$

where  $G_{11}$  is the upper-left  $p \times p$  block of  $G$ . After possibly changing coordinates via the transformation rule, this formula provides a general recipe for evaluating the LW functional on the boundary  $\partial\mathcal{S}_+^n$ , which is the content of Theorem 30 below.

Unfortunately, there are nontrivial analytic difficulties that are hidden by this heuristic derivation. In fact there exists an interaction  $U$  satisfying the weak growth condition for which the continuous extension property fails. Since the discussion of this counterexample is somewhat involved, it is postponed to section 5.5. However, the continuous extension property is true for  $U$  satisfying the strong growth condition of Definition 4.

Before stating the continuous extension property in Theorem 30, we provide a more careful discussion of the structure of the boundary  $\partial\mathcal{S}_+^n$ . Consider a  $q$ -dimensional subspace  $K$  of  $\mathbb{R}^n$ , and let  $p = N - q$ . Then the set

$$S_K := \{G \in \mathcal{S}_+^N : \ker G = K\}$$

forms a ‘stratum’ of the boundary of  $\mathcal{S}_+$ , which is itself isomorphic to the set of  $p \times p$  positive definite matrices. In turn, one can consider boundary strata (of smaller dimension) nested inside of  $S_K$ .

We will show that the restriction of the Luttinger-Ward function to such a stratum is precisely the Luttinger-Ward function for a lower-dimensional system. To this end, fix a subspace  $K$  and choose any orthonormal basis  $v_1, \dots, v_p$  for  $K^\perp$ . (The choice of basis is not canonical but can be made for the purpose of writing down results explicitly.) Define  $V_p := [v_1, \dots, v_p]$ . We use this notation to indicate both the matrix and the corresponding linear map.

**Theorem 30** (Continuous extension, I). *Suppose that  $U$  is continuous and satisfies the strong growth condition. With notation as in the preceding discussion,  $\Phi_N[\cdot, U]$  extends continuously to  $S_K$  via the rule*

$$\Phi_N[G, U] = \Phi_p[V_p^* G V_p, U \circ V_p]$$

for  $G \in S_K$ . Consequently,  $\Phi_N[\cdot, U]$  extends continuously to all of  $\mathcal{S}_+^N$ .

*Remark 31.* We interpret the extension rule as to set  $\Phi_N[0, U] = \Phi_0[U] := -2 \cdot U(0)$ . Moreover, it will become clear in the proof that even for continuous interactions  $U$  that do not satisfy the strong growth condition, the extension is still lower semi-continuous on  $\mathcal{S}_+^n$  and continuous on  $\mathcal{S}_{++}^n \cup \{0\}$ .

Changing coordinates via Proposition 22, we see that Theorem 30 is actually equivalent to the following:

**Theorem 32** (Continuous extension, II). *Suppose that  $U$  is continuous and satisfies the strong growth condition. For  $G \in \mathcal{S}_{++}^p$ ,  $\Phi[\cdot, U]$  extends continuously via the rule*

$$\Phi_N \left[ \begin{pmatrix} G & 0 \\ 0 & 0 \end{pmatrix}, U \right] = \Phi_p[G, U(\cdot, 0)].$$

Once again we comment that proof is deferred to section 5.

## 4 Bold diagrammatic expansion for the generalized Coulomb interaction

Using the Luttinger-Ward formalism, in this section we prove that the bold diagrammatic expansions from Part II of the self-energy and the LW functional (for the generalized Coulomb interaction (4.1)) can indeed be interpreted as asymptotic series expansions in the interaction strength at fixed  $G$ . This provides a rigorous interpretation of the bold expansions that is not merely combinatorial. Recall that when each  $G$  in the bold diagrammatic expansion of the self-energy is further expanded using  $G^0$  and  $U$ , the resulting expansion should be formally the same as the bare diagrammatic expansion of the self energy. The combinatorial argument in section 4 of Part II guaranteeing this fact does not need to be repeated in this setting, and we will be able to directly use Theorem 4.12 from Part II. The remaining hurdles are analytical, not combinatorial.

We summarize the results of this section as follows.

**Theorem 33.** *For any continuous interaction  $U : \mathbb{R}^n \rightarrow \mathbb{R}$  satisfying the weak growth condition and any  $G \in \mathcal{S}_{++}^n$ , the LW functional and the self-energy have asymptotic series expansions as*

$$\Phi[G, \varepsilon U] = \sum_{k=1}^{\infty} \Phi^{(k)}[G, U] \varepsilon^k, \quad \Sigma[G, \varepsilon U] = \sum_{k=1}^{\infty} \Sigma^{(k)}[G, U] \varepsilon^k. \quad (4.1)$$

Moreover, for  $U$  a homogeneous quartic polynomial, the coefficients of the asymptotic series satisfy

$$\Phi^{(k)}[G, U] = \frac{1}{2k} \text{Tr} [G \Sigma^{(k)}[G, U]]. \quad (4.2)$$

If  $U$  is moreover a generalized Coulomb interaction (1.6), we have (borrowing the language of Part II) that

$$\Sigma_{ij}^{(k)}[G, U] = \sum_{\Gamma_s \in \mathfrak{B}_2^{2\text{PI}}, \text{order } k} \frac{\mathbf{F}_{\Gamma_s}(i, j)}{S_{\Gamma_s}}, \quad (4.3)$$

i.e.,  $\Sigma^{(k)}$  is given the sum over bold skeleton diagrams of order  $k$  with bold propagator  $G$  and interaction  $v_{ij} \delta_{ik} \delta_{jl}$ .

*Remark 34.* For a series as in Eq. (4.1) to be asymptotic means that the error of the  $M$ -th partial sum is  $O(\varepsilon^{M+1})$  as  $\varepsilon \rightarrow 0$ .

Since  $U$  is fixed, for simplicity in the ensuing discussion we will omit the dependence on  $U$  from the notation via the definitions  $\Phi_G(\varepsilon) := \Phi[G, \varepsilon U]$ ,  $\Sigma_G(\varepsilon) = \Sigma[G, \varepsilon U]$ , and  $A_G(\varepsilon) := A[G, \varepsilon U]$ . We will also denote the series coefficients via

$\Phi_G^{(k)} := \Phi^{(k)}[G, U]$  and  $\Sigma_G^{(k)} := \Sigma^{(k)}[G, U]$ . In this notation, our asymptotic series take the form

$$\Phi_G(\varepsilon) = \sum_{k=1}^{\infty} \Phi_G^{(k)} \varepsilon^k, \quad \Sigma_G(\varepsilon) = \sum_{k=1}^{\infty} \Sigma_G^{(k)} \varepsilon^k. \quad (4.4)$$

**Notation 35.** *Note carefully that in this section the superscript  $(k)$  is merely a notation and does not indicate the  $k$ -th derivative. Such derivatives will be written out as  $\frac{d^k}{d\varepsilon^k}$ .*

Now we outline the remainder of this section. In section 4.1 we prove that the LW functional and the self-energy do indeed admit asymptotic series expansions. In section 4.2 we prove the relation between the LW and self-energy expansions for quartic interactions, namely Eq. (4.2). Interestingly, this relation—which is well-known formally based on diagrammatic observations—was originally assumed to be true to obtain a formal derivation of the LW functional [65, 68]. Our proof here does not rely on any diagrammatic manipulation, only making use of the transformation rule and the quartic nature of the interaction  $U$ . Similar relations for homogeneous polynomial interactions of different order could easily be obtained. Next, in section 4.3, we summarize and expand on the necessary results from Part II in diagram-free language; this both reduces the prerequisite knowledge needed for the remainder of the section and clarifies the arguments that follow. Finally, in section 4.4 we prove that when  $U$  is a generalized Coulomb interaction, the series for the self-energy is in fact the bold diagrammatic expansion of section 4 of Part II.

## 4.1 Existence of asymptotic series

In this section we assume that  $U$  is continuous and satisfies the weak growth condition. We first prove the following pair of lemmas.

**Lemma 36.** *For any  $G \in \mathcal{S}_{++}^n$ ,  $A_G(\varepsilon) \rightarrow G^{-1}$  as  $\varepsilon \rightarrow 0^+$ .*

**Lemma 37.** *For  $G \in \mathcal{S}_{++}^n$ , all derivatives of the functions  $\Phi_G : (0, \infty) \rightarrow \mathbb{R}$  and  $\Sigma_G : (0, \infty) \rightarrow \mathbb{R}^{N \times N}$  extend continuously to  $[0, \infty)$ .*

We will convey the continuous extension of the derivatives of  $\Phi_G$  to the origin by the notation  $\Phi_G^{(k)} := \Phi_G^{(k)}(0)$ , and similarly for the self-energy  $\Sigma_G^{(k)} := \Sigma_G^{(k)}(0)$ . From the preceding it will follow that the series (4.4) are indeed asymptotic series in the following sense:

**Proposition 38.** *For any nonnegative integer  $M$ ,  $\Phi_G(\varepsilon) - \sum_{k=1}^M \Phi_G^{(k)} \varepsilon^k = O(\varepsilon^{M+1})$  and  $\Sigma_G(\varepsilon) - \sum_{k=1}^M \Sigma_G^{(k)} \varepsilon^k = O(\varepsilon^{M+1})$  as  $\varepsilon \rightarrow 0^+$ .*

*Proof.* Consider any function  $f : [0, \infty) \rightarrow \mathbb{R}$  with all derivatives extending continuously up to the boundary (and so defined at 0). Let  $\delta > 0$ , so for  $\varepsilon \in (\delta, 1]$  we know

by the Lagrange error bound that

$$\left| f(\varepsilon) - \sum_{k=0}^M f^{(k)}(\delta)(\varepsilon - \delta)^k \right| \leq C(\varepsilon - \delta)^{M+1} \leq C\varepsilon^{M+1},$$

where  $C$  is a constant that depends only on a uniform bound on  $\left(\frac{d}{d\varepsilon}\right)^{k+1} f$  over  $[0, 1]$  (the existence of which is guaranteed by the continuous extension property). Simply taking the limit of our inequality as  $\delta \rightarrow 0^+$ , and again employing the continuous extension property, yields that  $\left| f(\varepsilon) - \sum_{k=0}^M f^{(k)}(0)\varepsilon^k \right| \leq C\varepsilon^{M+1}$ . This fact together with Lemma 37 proves the proposition.  $\square$

## 4.2 Relating the LW and self-energy expansions

The bold diagrams for the Luttinger-Ward functional are pinned down in terms of the bold diagrams for the self-energy via the following:

**Proposition 39.** *If  $U$  is a homogeneous quartic polynomial, then for all  $k$ ,*

$$\Phi_G^{(k)} = \frac{1}{2k} \text{Tr}[G\Sigma_G^{(k)}].$$

*Proof.* Observe that by the transformation rule that for any  $G \in \mathcal{S}_{++}^n$ ,  $\varepsilon, t > 0$ .

$$\Phi[tG, \varepsilon U] = \Phi[G, \varepsilon U \circ (t^{1/2}I)]$$

Taking the gradient in  $G$  of both sides, we have

$$t\Sigma[tG, \varepsilon U] = \Sigma[G, \varepsilon U \circ (t^{1/2}I)].$$

Since  $U$  is homogeneous quartic, in fact we have

$$\Sigma[tG, \varepsilon U] = \frac{1}{t} \Sigma[G, t^2\varepsilon U].$$

Then using this relation we compute:

$$\begin{aligned} \Phi[G, \varepsilon U] &= \int_0^1 \frac{d}{dt} \Phi[tG, \varepsilon U] dt \\ &= \int_0^1 \text{Tr}[G\Sigma[tG, \varepsilon U]] dt \\ &= \int_0^1 \frac{1}{t} \text{Tr}[G\Sigma[G, t^2\varepsilon U]] dt \\ &= \int_0^1 \frac{1}{t} \left[ \sum_{k=1}^M \text{Tr} \left[ G\Sigma_G^{(k)} \right] t^{2k} \varepsilon^k + O(t^{2(M+1)} \varepsilon^{M+1}) \right] dt \end{aligned}$$

$$= \int_0^1 \left[ \sum_{k=1}^M \text{Tr} \left[ G \Sigma_G^{(k)} \right] t^{2k-1} \varepsilon^k + O(t^{2M+1} \varepsilon^{M+1}) \right] dt.$$

Now since  $t$  ranges from 0 to 1 in the integrand, we have that  $t^{2N+1} \varepsilon^{N+1} \leq \varepsilon^{N+1}$ , and therefore

$$\begin{aligned} \Phi[G, \varepsilon U] &= \int_0^1 \left[ \sum_{k=1}^M \text{Tr} \left[ G \Sigma_G^{(k)} \right] t^{2k-1} \varepsilon^k \right] dt + O(\varepsilon^{M+1}) \\ &= \sum_{k=1}^M \frac{1}{2k} \text{Tr} \left[ G \Sigma_G^{(k)} \right] \varepsilon^k + O(\varepsilon^{M+1}). \end{aligned}$$

This establishes the proposition.  $\square$

### 4.3 Diagram-free discussion of results from Part II

For  $U$  satisfying the weak growth condition and  $A \in \text{dom } \Omega[\cdot, U]$ , define

$$\sigma[A, U] := A - (G[A, U])^{-1}.$$

Here we use the lowercase  $\sigma$  to emphasize that the self-energy here is being considered as a functional of  $A$  (not  $G$ ), together with the interaction.

Now we set the notation of  $U$  to indicated a fixed generalized Coulomb interaction (1.6). Further define

$$G_A(\varepsilon) := G[A, \varepsilon U], \quad \sigma_A(\varepsilon) := \sigma[A, \varepsilon U]. \quad (4.5)$$

The following lemma concerns the *bare* diagrammatic expansion of the Green's function and the self-energy, i.e., the asymptotic series for  $G_A$  and  $\sigma_A$ .

**Lemma 40.** *For fixed  $A \in \mathcal{S}_{++}^n$ , all derivatives  $\frac{d^n G_A}{d\varepsilon^n} : (0, \infty) \rightarrow \mathcal{S}_{++}^n$  and  $\frac{d^n \sigma_A}{d\varepsilon^n} : (0, \infty) \rightarrow \mathcal{S}^N$  extend continuously to  $[0, \infty)$ . In fact, interpreted as functions of both  $A$  and  $\varepsilon$ ,  $\frac{d^n G_A}{d\varepsilon^n}(\varepsilon)$  and  $\frac{d^n \sigma_A}{d\varepsilon^n}(\varepsilon)$  extend continuously to  $\mathcal{S}_{++}^n \times [0, \infty)$ . Moreover, we have asymptotic series expansions*

$$G_A(\varepsilon) = \sum_{k=0}^{\infty} g_A^{(k)} \varepsilon^k, \quad \sigma_A(\varepsilon) = \sum_{k=1}^{\infty} \sigma_A^{(k)} \varepsilon^k,$$

where the coefficient functions  $g_A^{(k)}$  and  $\sigma_A^{(k)}$  are polynomials in  $A^{-1}$ . More precisely,  $g_A^{(k)}$  and  $\sigma_A^{(k)}$  are homogeneous polynomials of degrees  $2k+1$  and  $2k-1$ , respectively. (Note that the zeroth-order term  $\sigma_A^{(0)}$  is implicitly zero.)

Finally, let  $G_A^{(\leq M)}(\varepsilon)$  and  $\sigma_A^{(\leq M)}(\varepsilon)$  denote the  $M$ -th partial sums of the above asymptotic series for  $G_A(\varepsilon)$  and  $\sigma_A(\varepsilon)$ , respectively. For every  $A \in \mathcal{S}_{++}^n$ , there exists a neighborhood  $\mathcal{N}$  of  $A$  in  $\mathcal{S}_{++}^n$  on which the truncation errors can actually be bounded

$$\left| G_A(\varepsilon) - G_A^{(\leq M)}(\varepsilon) \right| \leq C \varepsilon^{M+1}, \quad \left| \sigma_A(\varepsilon) - \sigma_A^{(\leq M)}(\varepsilon) \right| \leq C \varepsilon^{M+1}$$

for all  $\varepsilon \in [0, \tau]$ , with  $C, \tau$  independent of  $A \in \mathcal{N}$ .



*Proof.* The asymptotic series expansions for  $G_A$  and  $\Sigma_A$  are established in Theorems 3.15 and 3.17 of Part II. The continuous extension of the derivatives of  $G_A$  and  $\sigma_A$  to  $[0, \infty)$  follows from differentiation under the integral and simple dominated convergence arguments.

The uniform error bound follows from a Lagrange error bound argument as in Proposition 38, together with the continuity of  $\frac{d^n G_A}{d\varepsilon^n}(\varepsilon)$  and  $\frac{d^n \sigma_A}{d\varepsilon^n}(\varepsilon)$  on  $\mathcal{S}_{++}^n \times [0, \infty)$ .  $\square$

Inspired by Eq. (4.3), let

$$\mathbf{S}_G^{(k)} = \sum_{\Gamma_s \in \mathfrak{F}_2^{2\text{PI}}, \text{ order } k} \frac{\mathbf{F}_{\Gamma_s}}{S_{\Gamma_s}}.$$

In fact  $\mathbf{S}_G^{(k)}$  is polynomial in  $G$ , homogeneous of degree  $2k - 1$ . At this point we do not yet know that  $\mathbf{S}_G^{(k)}$  coincides with  $\Sigma_G^{(k)}$ , and indeed this is what we want to show. For any  $G$ , also define the partial sum

$$\mathbf{S}_G^{(\leq M)}(\varepsilon) := \sum_{k=1}^M \mathbf{S}_G^{(k)} \varepsilon^k.$$

Then the main result (Theorem 4.12) of Part II can be phrased as follows.

**Theorem 41.** *For any fixed  $A \in \mathcal{S}_{++}^n$ , the expressions*

$$\mathbf{S}_{G_A^{(\leq M)}(\varepsilon)}^{(\leq M)}(\varepsilon) = \sum_{k=1}^M \mathbf{S}_{G_A^{(\leq M)}(\varepsilon)}^{(k)} \varepsilon^k, \quad \sigma_A^{(\leq M)}(\varepsilon) = \sum_{k=1}^M \sigma_A^{(k)} \varepsilon^k$$

*agree as polynomials in  $\varepsilon$  up to order  $M$ , and hence they agree as joint polynomials in  $(A^{-1}, \varepsilon)$  after neglecting all terms in which  $\varepsilon$  appears degree at least  $M + 1$ .*

## 4.4 Derivation of self-energy bold diagrams

We have already shown that there *exist* asymptotic series for the LW functional and the self-energy. The remainder of Theorem 33 then consists of identifying that the self-energy coefficients  $\Sigma_G^{(k)}$  are indeed given by the bold diagrammatic expansion, i.e., that  $\Sigma_G^{(k)} = \mathbf{S}_G^{(k)}$ . Equivalently, we want to show that the partial sums  $\mathbf{S}_G^{(\leq M)}(\varepsilon)$  and  $\Sigma_G^{(\leq M)}(\varepsilon)$ , which are polynomials of degree  $M$  in  $\varepsilon$ , are equal. We will think of  $G \in \mathcal{S}_{++}^n$  as fixed throughout the following discussion, and we omit dependence on  $G$  from some of the notation below to avoid excess clutter. We will also think of  $M$  as a fixed positive integer and  $\varepsilon > 0$  as variable (and sufficiently small).

Since our series expansion is only valid in the asymptotic sense, for any finite  $M$  we consider the truncation

$$\Sigma_G^{(\leq M)}(\varepsilon) := \sum_{k=1}^M \Sigma_G^{(k)} \varepsilon^k.$$

Then we have  $\Sigma_G(\varepsilon) - \Sigma_G^{(\leq M)}(\varepsilon) = O(\varepsilon^{M+1})$ . For the purpose of this discussion,  $O(\varepsilon^{M+1})$  will be thought of as negligibly small, and ‘ $\approx$ ’ will be used to denote equality up to error  $O(\varepsilon^{M+1})$ . Meanwhile ‘ $\sim$ ’ will be used to denote error that is  $O(\varepsilon^{M+1-p})$  for all  $p \in (0, 1)$ , equivalently  $O(\varepsilon^{M+\delta})$  for all  $\delta \in (0, 1)$ . We remark that the difference between the relations ‘ $\approx$ ’ and ‘ $\sim$ ’ is due to technical reasons to be detailed later, and may be neglected on first reading.

Note that it actually suffices to show that  $\Sigma_G^{(\leq M)}(\varepsilon) \sim \mathbf{S}_G^{(\leq M)}(\varepsilon)$ . Indeed, both sides are polynomials of degree  $M$  in  $\varepsilon$ . Thus their difference is a polynomial of degree  $\leq M$ . If the degree- $n$  part of the difference is nonzero for some  $n = 1, \dots, M$ , then the difference is not  $O(\varepsilon^{n+\delta})$  for any  $\delta > 0$ . But if  $\Sigma_G^{(\leq M)}(\varepsilon) \sim \mathbf{S}_G^{(\leq M)}(\varepsilon)$ , then the difference is  $O(\varepsilon^{n+\delta})$  for all  $n = 1, \dots, M$ ,  $\delta \in (0, 1)$ . Thus in this case the difference is zero. With this reduction in mind, we now make a simple yet critical observation, namely that  $\Sigma_G^{(\leq M)}(\varepsilon)$  can be identified as the *exact* self-energy yielded by a modified interaction term. This will allow us to identify a quadratic form  $A^{(M)}(\varepsilon)$ , for which dependence on  $G$  has been suppressed from the notation, which generates (up to negligible error) the Green’s function  $G$  under the interaction  $\varepsilon U$ .

**Lemma 42.** *With notation as in the preceding discussion,  $\Sigma_G^{(\leq M)}(\varepsilon)$  is the self-energy induced by the interaction  $U_\varepsilon^{(M)}(x) := \varepsilon U(x) + \frac{1}{2}x^T \left[ \Sigma_G(\varepsilon) - \Sigma_G^{(\leq M)}(\varepsilon) \right] x$ , i.e.,*

$$\Sigma_G^{(\leq M)}(\varepsilon) = \Sigma[G, U_\varepsilon^{(M)}],$$

and moreover

$$A^{(M)}(\varepsilon) := A[G, U_\varepsilon^{(M)}] = G^{-1} + \Sigma_G^{(\leq M)}(\varepsilon).$$

Thus we may identify

$$G = G[A^{(M)}(\varepsilon), U_\varepsilon^{(M)}], \quad \Sigma_G^{(\leq M)}(\varepsilon) = \sigma[A^{(M)}(\varepsilon), U_\varepsilon^{(M)}].$$

*Proof.* Recalling that  $A_G(\varepsilon) = A[G, \varepsilon U]$  and  $\Sigma_G(\varepsilon) = \Sigma[G, \varepsilon U]$ , write

$$\begin{aligned} \frac{1}{2}x^T A_G(\varepsilon)x + U(x) &= \frac{1}{2}x^T \left( A_G(\varepsilon) - \Sigma_G(\varepsilon) + \Sigma_G^{(\leq M)}(\varepsilon) \right) x + U_\varepsilon^{(M)}(x) \\ &= \frac{1}{2}x^T \left( G^{-1} + \Sigma_G^{(\leq M)}(\varepsilon) \right) x + U_\varepsilon^{(M)}(x). \end{aligned}$$

It follows that under the interaction  $U_\varepsilon^{(M)}$ , the quadratic form  $G^{-1} + \Sigma_G^{(\leq M)}(\varepsilon)$  corresponds to the (interacting) Green’s function  $G$ . This establishes the second statement of the lemma, i.e., that

$$A[G, U_\varepsilon^{(M)}] = G^{-1} + \Sigma_G^{(\leq M)}(\varepsilon).$$

Moreover, by the Dyson equation we have that

$$\Sigma[G, U_\varepsilon^{(M)}] = A[G, U_\varepsilon^{(M)}] - G^{-1} = \Sigma_G^{(\leq M)}(\varepsilon),$$

which is the first statement of the lemma. The last statement then follows from the second, together with the definitions of  $G[\cdot, \cdot]$  and  $\sigma[\cdot, \cdot]$ .  $\square$

*Remark 43.* Note carefully that Lemma 42 is a non-perturbative fact and is valid for all  $\varepsilon > 0$ , though we shall apply it in a perturbative context.

At this point we have defined the terms needed to present a schematic diagram (Figure 32) of our proof that  $\Sigma_G^{(\leq M)}(\varepsilon) \sim \mathbf{S}_G^{(\leq M)}(\varepsilon)$ . Although the motivation for this schematic may not be fully clear at this point, the reader should refer back to it as needed for perspective.

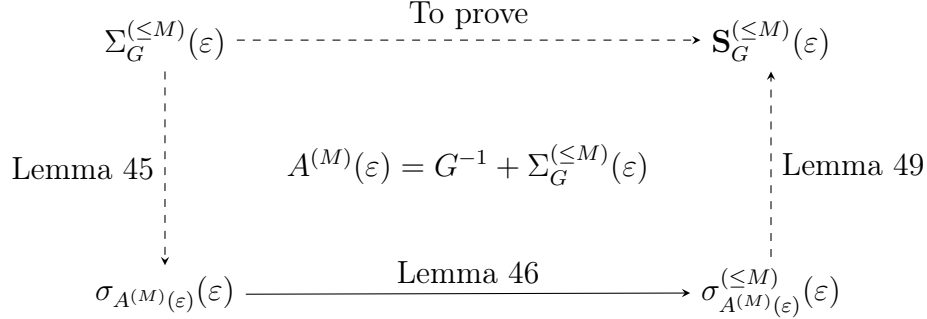


Figure 32: Schematic diagram for proving the bold diagrammatic expansion. Dashed lines indicate ‘ $\sim$ ’, and solid lines indicate ‘ $\approx$ ’.

Now recalling the definitions (4.5), we can write

$$G_{A^{(M)}(\varepsilon)}(\varepsilon) = G[A^{(M)}(\varepsilon), \varepsilon U], \quad \sigma_{A^{(M)}(\varepsilon)}(\varepsilon) := \sigma[A^{(M)}(\varepsilon), \varepsilon U]. \quad (4.6)$$

Meanwhile, following Lemma 42 we have the identities

$$G = G[A^{(M)}(\varepsilon), U_\varepsilon^{(M)}], \quad \Sigma_G^{(\leq M)}(\varepsilon) = \sigma[A^{(M)}(\varepsilon), U_\varepsilon^{(M)}]. \quad (4.7)$$

Note that pointwise,  $\varepsilon U$  and  $U_\varepsilon^{(M)}$  differ negligibly, but the form of  $\varepsilon U$  is simpler and easier to work with going forward.

Based on Eqs. (4.6) and (4.7), one then hopes that  $G_{A^{(M)}(\varepsilon)}(\varepsilon)$  is close to  $G$  and  $\sigma_{A^{(M)}(\varepsilon)}(\varepsilon)$  is close to  $\Sigma_G^{(\leq M)}(\varepsilon)$ . This is the content of the next two lemmas.

**Lemma 44.**  $G_{A^{(M)}(\varepsilon)}(\varepsilon) \sim G$ .

*Proof.* See appendix E.11. □

**Lemma 45.**  $\sigma_{A^{(M)}(\varepsilon)}(\varepsilon) \sim \Sigma_G^{(\leq M)}(\varepsilon)$ .

*Proof.* Based on Eqs. (4.6) and (4.7), we want to show that  $\sigma[A^{(M)}(\varepsilon), U_\varepsilon^{(M)}] \sim \sigma[A^{(M)}(\varepsilon), \varepsilon U]$ . We have already shown that  $G = G[A^{(M)}(\varepsilon), U_\varepsilon^{(M)}] \sim G[A^{(M)}(\varepsilon), \varepsilon U]$ , from which it follows that

$$A^{(M)}(\varepsilon) - (G[A^{(M)}(\varepsilon), U_\varepsilon^{(M)}])^{-1} \sim A^{(M)}(\varepsilon) - (G[A^{(M)}(\varepsilon), \varepsilon U])^{-1},$$

which is exactly what we want to show. □

Then we can use  $\sigma_{A^{(M)}(\varepsilon)}(\varepsilon)$  as a stepping stone to relate  $\Sigma_G^{(\leq M)}(\varepsilon)$  with the bare diagrammatic expansion for the self-energy via the following:

**Lemma 46.**  $\sigma_{A^{(M)}(\varepsilon)}(\varepsilon) \approx \sigma_{A^{(M)}(\varepsilon)}^{(\leq M)}(\varepsilon)$

*Proof.* Since  $A^{(M)}(\varepsilon) = G^{-1} + O(\varepsilon)$ , the result follows from Lemma 40 (in particular, the locally uniform bound on truncation error of the bare self-energy series).  $\square$

We can prove a similar fact (which will be useful later on) regarding the bare series for the interacting Green's function:

**Lemma 47.**  $G_{A^{(M)}(\varepsilon)}(\varepsilon) \approx G_{A^{(M)}(\varepsilon)}^{(\leq M)}(\varepsilon)$ .

*Proof.* Since  $A^{(M)}(\varepsilon) = G^{-1} + O(\varepsilon)$ , the result follows from Lemma 40 (in particular, the locally uniform bound on truncation error of the bare series for the interacting Green's function).  $\square$

From Lemmas 44 and 47 we immediately obtain:

**Lemma 48.**  $G_{A^{(M)}(\varepsilon)}^{(\leq M)}(\varepsilon) \sim G$ .

Finally, we are ready to state and prove the last leg of the schematic diagram (Figure 32):

**Lemma 49.**  $\mathbf{S}_G^{(\leq M)} \sim \sigma_{A^{(M)}(\varepsilon)}^{(\leq M)}(\varepsilon)$ .

*Proof.* Consider  $\mathbf{S}_{G_A^{(\leq M)}}^{(\leq M)}$  as a polynomial in  $(A^{-1}, \varepsilon)$ , and let  $P(A^{-1}, \varepsilon)$  be the contribution of terms in which  $\varepsilon$  appears with degree at least  $M + 1$ . By Theorem 41 we have the equality

$$\mathbf{S}_{G_A^{(\leq M)}}^{(\leq M)}(\varepsilon) - P(A^{-1}, \varepsilon) = \sigma_A^{(\leq M)}(\varepsilon)$$

of polynomials in  $(A^{-1}, \varepsilon)$ . Then substituting  $A \leftarrow A^{(M)}(\varepsilon)$  we obtain

$$\mathbf{S}_{G_{A^{(M)}(\varepsilon)}^{(\leq M)}}^{(\leq M)}(\varepsilon) - P([A^{(M)}(\varepsilon)]^{-1}, \varepsilon) = \sigma_{A^{(M)}(\varepsilon)}^{(\leq M)}(\varepsilon). \quad (4.8)$$

Although the first term on the left-hand side of Eq. (4.8) looks quite intimidating, we can recognize it as  $\mathbf{S}_{\mathbf{G}(\varepsilon)}^{(\leq M)}(\varepsilon)$ , where

$$\mathbf{G}(\varepsilon) := G_{A^{(M)}(\varepsilon)}^{(\leq M)}(\varepsilon) \sim G$$

is the expression from Lemma 48. Since  $\mathbf{S}_{[\cdot]}^{(\leq M)}(\varepsilon) = \sum_{k=1}^M \mathbf{S}_{[\cdot]}^{(k)} \varepsilon^k$ , where each  $\mathbf{S}_{[\cdot]}^{(k)}$  is a polynomial (homogeneous of positive degree) in the subscript slot, it follows that

$$\mathbf{S}_{\mathbf{G}(\varepsilon)}^{(\leq M)}(\varepsilon) \sim \mathbf{S}_G^{(\leq M)}(\varepsilon).$$

Then from Eq. (4.8) we obtain

$$\mathbf{S}_G^{(\leq M)}(\varepsilon) - P([A^{(M)}(\varepsilon)]^{-1}, \varepsilon) \sim \sigma_{A^{(M)}(\varepsilon)}^{(\leq M)}(\varepsilon).$$

But since  $[A^{(M)}(\varepsilon)]^{-1} = G + O(\varepsilon)$  and since  $P$  only includes terms of degree at least  $M + 1$  in the second slot, it follows that  $P([A^{(M)}(\varepsilon)]^{-1}, \varepsilon) \approx 0$ , and the desired result follows.  $\square$

Taken together (as indicated in Figure 32), Lemmas 45, 46, and 49 imply that  $\Sigma_G^{(\leq M)}(\varepsilon) \sim \mathbf{S}_G^{(\leq M)}(\varepsilon)$  as desired, and the proof of Theorem 33 is complete.

## 4.5 Caveat concerning truncation of the bold diagrammatic expansion

Although the LW and self-energy functionals are defined even for  $G$  such that the corresponding quadratic form  $A = A[G]$  is indefinite (and hence there is no physical bare non-interacting Green's function), Green's function methods (as discussed in section 4.7 of Part II) based on truncation of the bold diagrammatic expansion can fail dramatically in the case of indefinite  $A$ . One can encounter divergent behavior as the interaction becomes small, or the Green's function method may fail to admit a solution. Both failure modes can be demonstrated by simple one-dimensional examples. The relevance of these to the solution of the quantum many-body problem is at this point unclear.

Consider the one-dimensional example of

$$Z = \int_{\mathbb{R}} e^{\frac{1}{2}x^2 - \frac{1}{8}\lambda x^4} dx, \quad (4.9)$$

where  $a = -1$ . The corresponding non-interacting Green's function is  $G^0 = -1 < 0$  and hence is not even a physical Green's function.

Nonetheless with  $\lambda > 0$  the true Green's function is still well-defined via

$$G = \frac{1}{Z} \int_{\mathbb{R}} x^2 e^{\frac{1}{2}x^2 - \frac{1}{8}\lambda x^4} dx.$$

We now compute  $G$  via the Hartree-Fock method (cf. section 4.7 of Part II), i.e., we approximate the self-energy as

$$\Sigma^{(1)} = -\frac{1}{2}\lambda G - \lambda G = -\frac{3}{2}\lambda G.$$

Hence the self-consistent solution  $G^{(1)}$  of the Dyson equation solves

$$\frac{1}{G^{(1)}} = -1 + \frac{3}{2}\lambda G^{(1)}.$$

There is only one positive (physical) solution to this equation, namely

$$G^{(1)} = \frac{1 + \sqrt{1 + 6\lambda}}{3\lambda}.$$

In the spirit of perturbation theory, one might hope that  $G^{(1)}$  is a good approximation to  $G$  at least when  $\lambda \rightarrow 0$ . However we see just the opposite. This is perhaps not surprising because the exact Green's function  $G$  itself blows up in this limit.

The failure of the method as  $\lambda \rightarrow 0$  can be understood more precisely as follows. Rewrite the Hamiltonian from (4.9) as

$$\frac{1}{8}\lambda \left(x^2 - \frac{2}{\lambda}\right)^2 - \frac{1}{2\lambda}.$$

The corresponding Gibbs measure (which is unaffected by the additive constant) then concentrates about two peaks at  $x = \pm\sqrt{\frac{2}{\lambda}}$  as  $\lambda \rightarrow 0$ . Hence we expect

$$G \sim 2\lambda^{-1}.$$

We note that, in contrast with the statement of Lemma 40, the limit  $\lim_{\lambda \rightarrow 0+} G(\lambda)$  does not exist. According to Eq. (4.5)

$$G^{(1)} \sim \frac{2}{3}\lambda^{-1}.$$

We find that as  $\lambda \rightarrow 0+$ ,  $G$  and its first order approximation  $G^{(1)}$  do not agree.

If we include the second-order terms of the bold diagrammatic expansion

$$\Sigma^{(2)} = \frac{1}{2}\lambda^2 G^3 + \lambda^2 G^3 = \frac{3}{2}\lambda^2 G^3. \quad (4.10)$$

Then the self-consistent solution  $G^{(2)}$  of the Dyson equation solves

$$\frac{1}{G^{(2)}} = -1 + \frac{3}{2}\lambda G^{(2)} - \frac{3}{2}\lambda^2 (G^{(2)})^3.$$

This yields a quartic equation in the scalar  $G^{(2)}$ , which in fact has no solution for physical  $G^{(2)}$ , i.e.,  $G^{(2)} > 0$ .

To see this, first ease the notation by substituting  $x \leftarrow G^{(2)}$ , so we are interested in the solutions  $x > 0$  of

$$\frac{3}{2} [(\lambda^{1/2}x)^4 - (\lambda^{1/2}x)^2] + x + 1 = 0.$$

But  $y^4 - y^2 \geq -\frac{1}{4}$  for all  $y$ , so the first term is at least  $-\frac{3}{8}$ , which evidently implies that no solutions exist for  $x > 0$ .

## 5 Proof of the continuous extension of the LW functional

In section 3.5 we motivated the continuous extension of the LW functional to the boundary of  $\mathcal{S}_{++}^n$  and stated this result in two equivalent forms (Theorems 30 and 32). In this section we prove the continuous extension property (for interactions of strong growth). We also develop the counterexample promised earlier, an interaction of weak but not strong growth for which the continuous extension property fails.

The section is outlined as follows. In section 5.1, we describe some preliminary reductions in the proof of the continuous extension property, after which the proof can be divided into two parts: lower-bounding the limit inferior of the LW functional as the argument approaches the boundary and upper-bounding the limit supremum. In section 5.2, we prove the lower bound, and in section 5.3 we prove the upper bound. In section 5.4 we provide an alternate view on the continuous extension property from the Legendre dual side, and in section 5.5 we use this perspective to exhibit the aforementioned counterexample to the continuous extension property, which satisfies the weak growth condition but not the strong one.

### 5.1 Proof setup

We are going to prove Theorem 32, which as we have remarked suffices to prove Theorem 30 by changing coordinates via Proposition 22.

Suppose  $G \in \mathcal{S}_+^n$  is of the form

$$G = \begin{pmatrix} G_p & 0 \\ 0 & 0 \end{pmatrix},$$

where  $G_p \in \mathcal{S}_{++}^p$ , and suppose that  $G^{(j)} \in \mathcal{S}_{++}^n$  with  $G^{(j)} \rightarrow G$  as  $j \rightarrow \infty$ . For each  $j$ , diagonalize  $G^{(j)} = \sum_{i=1}^N \lambda_i^{(j)} v_i^{(j)} \left(v_i^{(j)}\right)^T$ , where the  $v_i^{(j)}$  are orthonormal,  $\lambda_i^{(j)} > 0$  for  $i = 1, \dots, N$ .

We want to show that

$$\Phi_n[G^{(j)}, U] \rightarrow \Phi_p[G_p, U(\cdot, 0)].$$

It suffices to show that every subsequence has a convergent subsequence with its limit being  $\Phi_p[G_p, U(\cdot, 0)]$ . The  $G^{(j)}$  are convergent, hence bounded (in the  $\|\cdot\|_2$  norm), so the  $\lambda_i^{(j)}$  are bounded. Moreover, the  $v_i^{(j)}$  are all of unit length, hence bounded, so by passing to a subsequence if necessary we can assume that, for each  $i$ , there exist  $\lambda_i, v_i$  such that  $\lambda_i^{(j)} \rightarrow \lambda_i$  and  $v_i^{(j)} \rightarrow v_i$  as  $j \rightarrow \infty$ . It follows that the  $v_i$  are orthonormal and that  $G$  can be diagonalized as  $G = \sum_{i=1}^N \lambda_i v_i v_i^T$ . Since  $G_p$  is positive definite, we must have  $\lambda_i > 0$  for  $i = 1, \dots, p$ , and moreover  $\lambda_i = 0$  for  $i = p + 1, \dots, N$ . Evidently, the eigenvectors of  $G$  with strictly positive eigenvalues must be precisely

the eigenvectors of  $G_p$ , concatenated with  $N - p$  zero entries, i.e., for  $i = 1, \dots, p$ ,  $v_i$  must be of the form  $(*, 0)$ . By orthogonality, for  $i = p + 1, \dots, n$ ,  $v_i$  must be of the form  $(0, *)$ .

For convenience we also establish the following notation:

$$V_G := \text{span}\{v_1, \dots, v_p\}, \quad V_{G^{(j)}} := \text{span}\{v_1^{(j)}, \dots, v_p^{(j)}\}.$$

Now the proof consists of proving two bounds: a lower bound

$$\liminf_{j \rightarrow \infty} \Phi_N[G^{(j)}, U] \geq \Phi_p[G_p, U(\cdot, 0)]$$

and an upper bound

$$\limsup_{j \rightarrow \infty} \Phi_N[G^{(j)}, U] \leq \Phi_p[G_p, U(\cdot, 0)].$$

These bounds will be proved in the next two sections, i.e., sections 5.2 and 5.3, respectively.

## 5.2 Lower bound

We want to establish a lower bound on  $\Phi_N[G_j, U]$  via our expression for  $\mathcal{F}_N$  as a supremum:

$$\mathcal{F}_N[G^{(j)}, U] = \sup_{\mu \in \mathcal{G}_N^{-1}(G^{(j)})} \left[ H(\mu) - \int U \, d\mu \right]. \quad (5.1)$$

This strategy requires us to construct measures  $\mu^{(j)} \in \mathcal{G}_N^{-1}(G^{(j)})$ . Intuitively, what one hopes to do (though this strategy will require some modification) is the following: consider the measure  $\alpha$  on  $\mathbb{R}^p$  that attains the supremum in the analogous expression for  $\mathcal{F}_p[G_p, U(\cdot, 0)]$ , identify this measure with a measure on  $V_G \simeq \mathbb{R}^p$ , rotate and scale appropriately to obtain a measure  $\alpha^{(j)}$  supported on  $V_{G^{(j)}}$  with the correct second-order moments with respect to this subspace, and finally take the direct sum with an appropriate Gaussian measure  $\beta^{(j)}$  on  $V_{G^{(j)}}^\perp$ . Unfortunately, due to difficulties of analysis, it is not clear how to then prove the desired limit as  $j \rightarrow \infty$ .

However, the analysis of this limit would be feasible if the  $\mu^{(j)}$  had compact support (which they evidently do not). Then our approach is to carry out a construction that preserves the spirit of the ‘ideal’ construction just described but instead works with  $\mu^{(j)}$  of (uniform) compact support.

For convenience we let  $\mathcal{M}_c \subset \mathcal{M}_2$  denote the subset of measures of compact support. The acceptability of working with measures of compact support can be motivated by the following lemma, which will be used below. (In the statement we temporarily suppress dependence on the interaction and the dimension from the notation.)



**Lemma 50.** For all  $G \in \mathcal{S}^N$ ,

$$\mathcal{F}[G] = \sup_{\mu \in \mathcal{G}^{-1}(G) \cap \mathcal{M}_c} \left[ H(\mu) - \int U \, d\mu \right].$$

Now we outline our *actual* construction of the  $\mu^{(j)}$ . Consider an *arbitrary* measure  $\alpha \in \mathcal{G}_p^{-1}(G_p)$  with compact support on  $\mathbb{R}^p \simeq V_G$ . (We abuse notation slightly by considering  $\alpha$  as a measure on both  $\mathbb{R}^p$  and  $V_G$ .) The idea now is to construct a measure in  $\mu^{(j)} \in \mathcal{G}_N^{-1}(G^{(j)})$  by rotating  $\alpha$  and scaling appropriately to obtain a measure  $\alpha^{(j)}$  supported on  $V_{G^{(j)}}$  and then taking the direct sum with a compactly supported measure  $\beta^{(j)}$  on  $V_{G^{(j)}}^\perp$  (the details of which will be discussed later). In fact the supremum in (5.1) will be approximately attained by a measure of this form as  $j \rightarrow \infty$ , i.e., our lower bound will be tight as  $j \rightarrow \infty$ .

Accordingly, for the construction of  $\alpha^{(j)}$ , let  $O^{(j)}$  be the orthogonal linear transformation sending  $v_i \mapsto v_i^{(j)}$ , and let  $D^{(j)}$  be the linear transformation with matrix (in the  $v_i^{(j)}$  basis) given by

$$\text{diag} \left( \sqrt{\lambda_1^{(j)}/\lambda_1}, \dots, \sqrt{\lambda_p^{(j)}/\lambda_p}, 1, \dots, 1 \right).$$

Then define  $T^{(j)} := D^{(j)}O^{(j)}$  and  $\alpha^{(j)} := T^{(j)}\#\alpha$ . Note that  $T^{(j)} \rightarrow I_n$  as  $j \rightarrow \infty$ . Moreover, observe that  $\alpha^{(j)}$  is a measure supported on  $V_{G^{(j)}}$  with second-order moment matrix given by  $\text{diag}(\lambda_1^{(j)}, \dots, \lambda_p^{(j)})$  with respect to the coordinates on  $V_{G^{(j)}}$  induced by the orthonormal basis  $v_1^{(j)}, \dots, v_p^{(j)}$ .

Now we turn to the construction of  $\beta^{(j)}$ . Let  $R > 1$  and let  $\gamma$  be a measure supported on  $[-R, R]$  with  $\int x^2 \, d\gamma = 1$ . The parameter  $R$  will control the size of the support of  $\beta^{(j)}$  and will be sent to  $+\infty$  at the very end of the proof of the lower bound (after the limit in  $j$  has been taken). Then define

$$\Lambda^{(j)} := \text{diag} \left( \sqrt{\lambda_{p+1}^{(j)}}, \dots, \sqrt{\lambda_N^{(j)}} \right),$$

and define a measure  $\beta^{(j)}$  on  $\mathbb{R}^{N-p}$  by  $\beta^{(j)} := \Lambda^{(j)}\#(\gamma \times \dots \times \gamma)$ . Note that  $\Lambda^{(j)} \rightarrow 0$  as  $j \rightarrow \infty$ . Abusing notation slightly, we will also identify  $\beta^{(j)}$  with a measure supported on  $V_{G^{(j)}}^\perp \simeq \mathbb{R}^{N-p}$  via the identification of the orthonormal basis  $v_{p+1}^{(j)}, \dots, v_N^{(j)}$  for  $V_{G^{(j)}}^\perp$  with the standard basis of  $\mathbb{R}^{N-p}$ .

Finally, define the product measure  $\mu^{(j)} := \alpha^{(j)} \times \beta^{(j)}$  with respect to the product structure  $\mathbb{R}^n = V_{G^{(j)}} \times V_{G^{(j)}}^\perp$ , and note that  $\mu^{(j)} \in \mathcal{G}_N^{-1}(G^{(j)})$ , so by (5.1),

$$\begin{aligned} \mathcal{F}_N[G^{(j)}, U] &\geq H(\alpha^{(j)} \times \beta^{(j)}) - \int U \, d\mu^{(j)} \\ &= H(\alpha^{(j)}) + H(\beta^{(j)}) - \int U \, d\mu^{(j)} \end{aligned}$$

$$= H(\alpha) - \int U \, d\mu^{(j)} + \frac{1}{2} \sum_{i=p+1}^N \log \lambda_i^{(j)} + (N-p)H(\gamma),$$

where  $H(\alpha^{(j)})$  and  $H(\beta^{(j)})$  are the entropies of  $\alpha^{(j)}$  and  $\beta^{(j)}$  on the probability spaces  $V_{G^{(j)}}$  and  $V_{G^{(j)}}^\perp$ , respectively.

Notice that there is a compact set on which *all* of the measures  $\mu^{(j)}$  are supported. It is then not difficult to see that  $\mu^{(j)}$  converges weakly to the measure  $\alpha \times \delta_0$ , where the product is with respect to the product structure  $\mathbb{R}^n = V_G \times V_G^\perp$  and  $\delta_0$  is the Dirac delta measure localized at the origin. By the continuity of  $U$  and the uniform boundedness of the supports of  $\mu^{(j)}$ , this is enough to guarantee that

$$\int U \, d\mu^{(j)} \rightarrow \int U \, d(\alpha \times \delta_0) = \int U(\cdot, 0) \, d\alpha$$

as  $j \rightarrow \infty$ .

Next we write the Luttinger-Ward functional in terms of  $\mathcal{F}_N$ :

$$\begin{aligned} \frac{1}{2} \Phi_N[G^{(j)}, U] &= \mathcal{F}_N[G^{(j)}, U] - \frac{1}{2} \text{Tr}[\log(G^{(j)})] - \frac{N}{2} \log(2\pi e) \\ &= \mathcal{F}_N[G^{(j)}, U] - \frac{1}{2} \sum_{i=1}^N \log \lambda_i^{(j)} - \frac{N}{2} \log(2\pi e). \end{aligned}$$

Then combining the preceding observations yields

$$\begin{aligned} &\liminf_{j \rightarrow \infty} \frac{1}{2} \Phi_N[G^{(j)}, U] \\ &\geq \liminf_{j \rightarrow \infty} \left[ H(\alpha) - \int U \, d\mu^{(j)} - \frac{1}{2} \sum_{i=1}^p \log \lambda_i^{(j)} - \frac{N}{2} \log(2\pi e) + (N-p)H(\gamma) \right] \\ &= H(\alpha) - \int U(\cdot, 0) \, d\alpha - \frac{1}{2} \sum_{i=1}^p \log \lambda_i - \frac{N}{2} \log(2\pi e) + (N-p)H(\gamma) \\ &= H(\alpha) - \int U(\cdot, 0) \, d\alpha - \frac{1}{2} \text{Tr}[\log(G_p)] - \frac{N}{2} \log(2\pi e) + (N-p)H(\gamma). \end{aligned}$$

Now for any  $\varepsilon > 0$ , we can choose  $R$  sufficiently large and  $\gamma$  supported on  $[-R, R]$  such that  $H(\gamma) \geq \frac{1}{2} \log(2\pi e) - \varepsilon$ . Indeed, note that  $\frac{1}{2} \log(2\pi e)$  is the entropy of the standard normal distribution, i.e., the maximal entropy over measures of unit variance. By restricting the normal distribution to  $[-R, R]$  for  $R$  sufficiently large, we can become arbitrarily close to saturating this bound. Therefore we have that

$$\liminf_{j \rightarrow \infty} \frac{1}{2} \Phi_N[G^{(j)}, U] \geq H(\alpha) - \int U(\cdot, 0) \, d\alpha - \frac{1}{2} \text{Tr}[\log(G_p)] - \frac{p}{2} \log(2\pi e).$$

Since  $\alpha$  was arbitrary in  $\mathcal{G}_p^{-1}(G_p) \cap \mathcal{M}_c$ , this establishes the desired upper bound

$$\begin{aligned} \frac{1}{2} \liminf_{j \rightarrow \infty} \Phi_N[G^{(j)}, U] &\geq \sup_{\alpha \in \mathcal{G}_p^{-1}(G_p) \cap \mathcal{M}_c} \left[ H(\alpha) - \int U(\cdot, 0) d\alpha \right] \\ &\quad - \frac{1}{2} \text{Tr} [\log(G_p)] - \frac{p}{2} \log(2\pi e) \\ &= \frac{1}{2} \Phi_p[G_p, U(\cdot, 0)], \end{aligned}$$

where we have used Lemma 50, which allows us to look at the supremum over compactly supported measures.

Observe that the proof of the lower bound did not require the strong growth assumption, hence the semi-continuity claim of Remark 31.

### 5.3 Upper bound

Next we turn to establishing an upper bound. The basic strategy is to select measures  $\mu^{(j)}$  that (approximately) attain the supremum in (5.1) and take a limit as  $j \rightarrow \infty$ .

Before proceeding, let  $\varepsilon > 0$ . Moreover, define  $\pi_1$  to be the orthogonal projection onto  $V_G \simeq \mathbb{R}^p$ , and define  $\pi_2$  to be the orthogonal projection onto  $V_G^\perp \simeq \mathbb{R}^{N-p}$ .

Now for every  $j$ , as suggested above choose  $\mu^{(j)} \in \mathcal{G}_N^{-1}(G^{(j)})$  such that

$$\mathcal{F}_N[G^{(j)}, U] \leq H(\mu^{(j)}) - \int U d\mu^{(j)} + \varepsilon.$$

Therefore

$$\Phi_N[G^{(j)}, U] \leq \underbrace{H(\mu^{(j)}) - \int U d\mu^{(j)}}_{=: a_j} - \frac{1}{2} \sum_{i=1}^N \log(2\pi e \lambda_i^{(j)}) + \varepsilon. \quad (5.2)$$

Then choose a subsequence  $j_k$  such that  $\lim_{k \rightarrow \infty} a_{j_k} = \limsup_{j \rightarrow \infty} a_j$ .

Now the  $\mu^{(j)}$  have uniformly bounded second moments, so by Markov's inequality, the sequence  $\mu^{(j)}$  is tight. Then by Prokhorov's theorem (Theorem 31), we can assume, by extracting a further subsequence if necessary, that  $\mu^{(j_k)}$  converges weakly to some measure  $\mu$ .

We claim that  $\mathcal{G}_N(\mu) \preceq G$  (so in particular,  $\mu \in \mathcal{M}_2$ ). Indeed, for any  $z \in \mathbb{R}^n$ , by the Portmanteau theorem for weak convergence of measures (Theorem 28) we have

$$\begin{aligned} \int (z^T x)^2 d\mu &\leq \liminf_{k \rightarrow \infty} \int (z^T x)^2 d\mu^{(j_k)} \\ &= \liminf_{k \rightarrow \infty} \int z^T x x^T z d\mu^{(j_k)} = \liminf_{k \rightarrow \infty} z^T G^{(j_k)} z = z^T G z. \end{aligned}$$

It follows that  $\mu \in \mathcal{M}_2$  and moreover  $z^T \mathcal{G}_n(\mu) z \leq z^T G z$  for all  $z$ , i.e.,  $\mathcal{G}_n(\mu) \preceq G$ . In particular,  $\mu$  is supported on  $V_G$ .

Define  $T^{(j)}$  to be the *orthogonal* transformation that sends  $v_i^{(j)} \mapsto v_i$ , so  $T^{(j)} \rightarrow I_n$  as  $j \rightarrow \infty$ . Define  $\nu^{(j)} := T^{(j)} \# \mu^{(j)}$ . Again by Prokhorov's theorem, we can assume that  $\nu^{(j_k)}$  converges weakly to some measure  $\nu$ . In fact, we must have  $\nu = \mu$ . To see this, note that for any continuous compactly supported function  $\phi$  on  $\mathbb{R}^n$ , we have that  $\phi \circ T^{(j)} \rightarrow \phi$  uniformly as  $j \rightarrow \infty$ . Therefore

$$\lim_{j \rightarrow \infty} \int |\phi - \phi \circ T^{(j)}| d\mu^{(j)} \rightarrow 0.$$

Consequently

$$\int \phi d\mu = \lim_{k \rightarrow \infty} \int \phi d\mu^{(j_k)} = \lim_{k \rightarrow \infty} \int \phi \circ T^{(j_k)} d\mu^{(j_k)} = \lim_{k \rightarrow \infty} \int \phi d\nu^{(j_k)} = \int \phi d\nu.$$

Since  $\mu$  and  $\nu$  agree on all continuous compactly supported functions, they must be equal (Riesz representation theorem), and  $\nu^{(j_k)} \rightarrow \mu$  weakly.

Define  $\mu_i^{(j)} := \pi_i \# \nu^{(j)} = (\pi_i \circ T^{(j)}) \# \mu^{(j)}$  and  $\mu_i := \pi_i \# \mu$  for  $i = 1, 2$ . It follows that  $\mu_i^{(j_k)} \rightarrow \mu_i$  weakly. Notice (using Fact 11) that

$$H(\mu^{(j)}) = H(\nu^{(j)}) \leq H(\mu_1^{(j)}) + H(\mu_2^{(j)}) \leq H(\mu_1^{(j)}) + \frac{1}{2} \sum_{i=p+1}^N \log(2\pi e \lambda_i^{(j)}).$$

Therefore, using Lemma 9 with the weak convergence  $\mu_1^{(j_k)} \rightarrow \mu_1$ , we obtain

$$\begin{aligned} \lim_{k \rightarrow \infty} a_{j_k} &= \lim_{k \rightarrow \infty} \left[ H(\mu^{(j_k)}) - \int U d\mu^{(j_k)} - \frac{1}{2} \sum_{i=1}^N \log(2\pi e \lambda_i^{(j_k)}) \right] \\ &\leq \limsup_{k \rightarrow \infty} \left[ H(\mu_1^{(j_k)}) - \frac{1}{2} \sum_{i=1}^p \log(2\pi e \lambda_i^{(j)}) \right] - \liminf_{k \rightarrow \infty} \left[ \int U d\mu^{(j_k)} \right] \\ &\leq H(\mu_1) - \liminf_{k \rightarrow \infty} \left[ \int U d\mu^{(j_k)} \right] - \frac{1}{2} \log((2\pi e)^p \det G_p). \end{aligned}$$

Now for any  $\alpha \in \mathbb{R}$ , define  $U_\alpha(x) = U(x) - \alpha \|x\|^2$ . Then

$$\int U d\mu^{(j)} = \int U_\alpha d\mu^{(j)} + \alpha \text{Tr}[G^{(j)}].$$

The utility of this manipulation will be made clear later. By the strong growth condition,  $U_\alpha$  is bounded below. Therefore, by the Portmanteau theorem for weak convergence of measures,

$$\liminf_{k \rightarrow \infty} \left[ \int U d\mu^{(j_k)} \right] = \alpha \text{Tr}[G] + \liminf_{k \rightarrow \infty} \left[ \int U_\alpha d\mu^{(j_k)} \right] \geq \alpha \text{Tr}[G_p] + \int U_\alpha d\mu.$$

Since  $\mu$  is supported on  $V_G$ , in fact

$$\int U_\alpha \, d\mu = \int U_\alpha(\cdot, 0) \, d\mu_1 = \int U(\cdot, 0) \, d\mu_1 - \alpha \text{Tr}[\mathcal{G}_p(\mu_1)],$$

and therefore

$$\begin{aligned} \lim_{k \rightarrow \infty} a_{j_k} &\leq H(\mu_1) - \int U(\cdot, 0) \, d\mu_1 - \frac{1}{2} \log((2\pi e)^p \det G_p) + \alpha \text{Tr}[\mathcal{G}_p(\mu_1) - G_p] \\ &\leq \mathcal{F}_p[\mathcal{G}_p(\mu_1), U(\cdot, 0)] - \frac{1}{2} \log((2\pi e)^p \det G_p) + \alpha \text{Tr}[\mathcal{G}_p(\mu_1) - G_p]. \end{aligned}$$

Recall from (5.2) that

$$\limsup_{j \rightarrow \infty} \Phi[G^{(j)}, U] \leq \lim_{k \rightarrow \infty} a_{j_k} + \varepsilon.$$

Since  $\varepsilon > 0$  was arbitrary, this means that

$$\limsup_{j \rightarrow \infty} \Phi[G^{(j)}, U] \leq \mathcal{F}_p[\mathcal{G}_p(\mu_1), U(\cdot, 0)] - \frac{1}{2} \log((2\pi e)^p \det G_p) + \alpha \text{Tr}[\mathcal{G}_p(\mu_1) - G_p].$$

If we had  $\mathcal{G}_N(\mu) = G$ , i.e.,  $\mathcal{G}_p(\mu_1) = G_p$ , then we would be done. We have  $\mathcal{G}_p(\mu_1) \preceq G_p$ , so it will suffice to show that  $\text{Tr}[\mathcal{G}_p(\mu_1) - G_p] = 0$ . Suppose for contradiction that  $\text{Tr}[\mathcal{G}_p(\mu_1) - G_p] < 0$ . But then, by taking  $\alpha$  arbitrarily large we see that  $\limsup_{j \rightarrow \infty} \Phi[G^{(j)}, U] = -\infty$ , which is impossible because we already have a lower bound on  $\liminf_{j \rightarrow \infty} \Phi[G^{(j)}, U]$ . Therefore  $\mathcal{G}_p(\mu_1) = G_p$ , as desired, and we have

$$\limsup_{j \rightarrow \infty} \Phi[G^{(j)}, U] \leq \Phi_p[G_p, U(\cdot, 0)],$$

which completes the proof.

Notice the strong growth assumption was only used in this part of the proof (i.e., the proof of the upper bound). In particular, it was only used to ensure that the measure  $\mu^{(j)}$  of maximum entropy relative to  $\nu_U$  (as in Remark 18) subject to the moment constraint  $\mathcal{G}(\mu^{(j)}) = G^{(j)}$  cannot weakly converge to a measure  $\mu$  with  $\mathcal{G}(\mu) \neq G = \lim_{j \rightarrow \infty} G^{(j)}$ .

## 5.4 Dual perspective on continuous extension

We now outline how Theorem 30 can be reinterpreted via the transformation rule. This perspective provides another way of understanding Theorem 30 and allows us to present a counterexample that illustrates the necessity of the strong growth condition of Definition 4.

Suppose that  $T_j$  are linear transformations such that  $T_j \rightarrow P$ , where  $P = I_p \oplus 0_{N-p}$  is the orthogonal projection onto  $\text{span}\{e_1^{(n)}, \dots, e_p^{(n)}\}$ . Let  $G \in \mathcal{S}_{++}^n$  with upper-left block given by  $G_p$ . Then, using the transformation rule, Theorem 30, and the projection rule, we obtain

$$\Phi_N[G, U \circ T_j] = \Phi_N[T_j G T_j^*, U] \rightarrow \Phi_p[G_p, U(\cdot, 0)] = \Phi_N[G, U \circ P].$$

This manipulation suggests that Theorem 30 is equivalent to the pointwise convergence

$$\Phi_N[\cdot, U \circ T_j] \rightarrow \Phi_N[\cdot, U \circ P] \quad (5.3)$$

for all  $T_j \rightarrow P$ . To see the equivalence, consider an arbitrary sequence  $G^{(j)} \in \mathcal{S}_{++}^n$  converging, as before, to the block-diagonal matrix  $G = G_p \oplus 0_{N-p} \in \mathcal{S}_+^n$ , where  $G_p \in \mathcal{S}_{++}^p$ . Then we want to show, using Eq. (5.3), that  $\Phi_N[G^{(j)}, U] \rightarrow \Phi_p[G_p, U(\cdot, 0)]$ .

To this end, let  $T_j = [G^{(j)}]^{1/2}[G_p \oplus I_{N-p}]^{-1/2}$ , so  $G^{(j)} = T_j(G_p \oplus I_{N-p})T_j^*$ , and  $T_j \rightarrow P$ . Then (5.3) implies that  $\Phi_N[G_p \oplus I_{N-p}, U \circ T_j] \rightarrow \Phi_N[G_p \oplus I_{N-p}, U \circ P]$ , and combining with the transformation and projection rules yields Theorem 30.

Note that (5.3) is equivalent to the pointwise convergence of concave functions  $\mathcal{F}_N[\cdot, U \circ T_j] \rightarrow \mathcal{F}_N[\cdot, U \circ P]$  as  $T_j \rightarrow P$ . Since the domains of these concave functions are open (namely,  $\mathcal{S}_{++}^n$ ), by Theorem 27 this is actually equivalent to uniform convergence on all compact subsets of  $\mathcal{S}_{++}^n$ . Furthermore, since  $\mathcal{F}_N[\cdot, U \circ T_j]$  and  $\mathcal{F}_N[\cdot, U \circ P]$  are both uniformly  $-\infty$  on  $\mathcal{S}^N \setminus \mathcal{S}_{++}^n$ , this is in turn equivalent to uniform convergence on all compact subsets of  $\mathcal{S}^N$  that do not contain a boundary point of  $\mathcal{S}_{++}^n$ , which by Theorem 25 is equivalent to the hypo-convergence (see Definition 24)  $\mathcal{F}_N[\cdot, U \circ T_j] \xrightarrow{h} \mathcal{F}_N[\cdot, U \circ P]$ . (Note that the role of epi-convergence for convex functions is assumed by hypo-convergence for concave functions.) But then hypo-convergence is equivalent to hypo-convergence of the concave conjugates (Theorem 26), i.e., of  $\Omega[\cdot, U \circ T_j]$  to  $\Omega[\cdot, U \circ P]$  as  $j \rightarrow \infty$ .

In summary, the continuous extension property is equivalent to the hypo-convergence  $\Omega[\cdot, U \circ T_j] \xrightarrow{c} \Omega[\cdot, U \circ P]$ .

## 5.5 Counterexample of weak but not strong growth

Here we give a counter example to show that the weak growth condition is insufficient for guaranteeing the continuous extension property. By the discussion of section 5.4, we need only find  $U$  satisfying the weak growth condition for which  $\Omega[\cdot, U \circ T_j]$  fails to hypo-converge to  $\Omega[\cdot, U \circ P]$ .

For example, consider  $N = 2$  and

$$U(x_1, x_2) = \begin{cases} |x_1|^4 & |x_1| \leq |x_2|^{-1} \\ |x_2|^{-4} & |x_1| \geq |x_2|^{-1}. \end{cases}$$

If  $x_2 = 0$ , then the first case holds for all  $x_1$ . This interaction is nonnegative, and hence satisfies the first part of the weak growth condition of Definition 3 with  $C_U = 0$ . To see that  $U$  satisfies the weak growth condition, we need only show that  $\text{dom } \Omega$  is open. Clearly  $\text{dom } \Omega \supset \mathcal{S}_{++}^n$ . Moreover, the restriction of  $U$  to any line except the  $x_1$ -axis is bounded, and it follows that in fact  $\text{dom } \Omega = \mathcal{S}_{++}^n$ , hence  $\text{dom } \Omega$  is open, as desired.

Now let

$$T_j := \begin{pmatrix} 1 & 0 \\ 0 & j^{-1} \end{pmatrix} \rightarrow P := \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}.$$

Since  $\Omega[\cdot, U \circ P]$  has an open domain, namely,

$$\text{dom} (\Omega[\cdot, U \circ P]) = \{A = (a_{ij}) \in \mathcal{S}^2 : a_{22} > 0\},$$

the hypo-convergence of  $\Omega[\cdot, U \circ T_j]$  to  $\Omega[\cdot, U \circ P]$  is equivalent to pointwise convergence (by Theorems 25 and 27), which is the same as the pointwise convergence  $Z[\cdot, U \circ T_j] \rightarrow Z[\cdot, U \circ P]$ .

Set  $A = (a_{ij})$  via  $a_{11} = a_{12} = 0$ ,  $a_{22} = 1$ , so  $A$  is in the domain of  $\Omega[\cdot, U \circ P]$ , i.e.,  $Z[A, U \circ P] < +\infty$ . However,

$$Z[A, U \circ T_j] = \int e^{-\frac{1}{2}|x_2|^2 - U(x_1, j^{-1}x_2)} dx_1 dx_2 = j \cdot \int e^{-j^2 \frac{1}{2}|x_2|^2 - U(x_1, x_2)} dx_1 dx_2.$$

Now the restriction of the last integrand to any line of constant  $x_2 \neq 0$  is asymptotically equal to  $e^{-j^2|x_2|^2 - |x_2|^{-4}} > 0$ , so the integral along any such line is  $+\infty$ , and by Fubini's theorem,  $Z[A, U \circ T_j] = +\infty$ . Thus convergence fails at  $A$ , and we have a counterexample as claimed.

## Part IV

# Large-interaction Luttinger-Ward asymptotics

## 1 Introduction

Motivated by the preceding Parts II and III, we are interested in Gibbs measures proportional to

$$e^{-\frac{1}{2}x^T Ax - U_\lambda(x)},$$

where the interaction  $U_\lambda = \lambda U$  is specified by

$$U(x) := \frac{1}{8} \sum_{ij} v_{ij} x_i^2 x_j^2,$$

with  $v = (v_{ij})$  symmetric positive definite and  $\lambda > 0$ . Here  $i$  should be understood as an index on a set of discrete sites, while  $x_i$  measures the ‘amount of charge’ present on the  $i$ -th site.

Parts II and III we have analyzed the perturbative regime  $\lambda \approx 0$ . Hence it is natural to ask whether the  $\lambda \rightarrow \infty$  limit can be characterized as well. More specifically, in order to examine the behavior of the Luttinger-Ward functional in this limit, we will consider the limit where  $A = A_\lambda$  also depends on  $\lambda$  and is chosen such that the second moments of the corresponding Gibbs measure are independent of  $\lambda$ . We will find that the Hartree contribution predominates in this limit, and leading-order corrections are determined in terms of the statistical mechanics of a classical spin system. From the point of view of perturbation theory, the dominance of the Hartree contribution may be thought of as somewhat surprising, since it is in this sense derived as a term in the first-order expansion about  $\lambda = 0$  (see Part II).

## 2 Preliminaries

We recapitulate the results of Part III relevant to this setting. Let  $\mathcal{S}^n \supset \mathcal{S}_+^n \supset \mathcal{S}_{++}^n$  denote the sets of real symmetric, positive semidefinite, and positive definite matrices, respectively, and define  $Z_\lambda, \Omega_\lambda : \mathcal{S}^n \rightarrow \mathbb{R}$  via

$$Z_\lambda[A] := \int e^{-\frac{1}{2}x^T Ax - U_\lambda(x)} dx, \quad \Omega_\lambda[A] := -\log Z_\lambda[A].$$

Then define  $G_\lambda : \mathcal{S}^n \rightarrow \mathcal{S}_{++}^n$  by

$$G_\lambda[A] := \nabla \Omega_\lambda[A] = \int xx^T e^{-\frac{1}{2}x^T Ax - U_\lambda(x)} dx.$$



The naturality of the definiteness assumption on  $v$  will become clear below, but for now, note that this assumption ensures in particular that  $U$  is a nonnegative polynomial, strictly positive away from  $x = 0$ . Since  $U$  is homogeneous quartic, it follows that  $U \geq C^{-1}|x|^4$  for some constant  $C$  sufficiently large. In particular, for any  $\lambda$ ,  $U_\lambda$  satisfies the ‘strong growth assumption’ of the Part III.

Let  $\mathcal{M}^n$  be the set of probability measures on  $\mathbb{R}^n$  with finite second-order moments, and define  $\mathcal{G} : \mathcal{M}^n \rightarrow \mathcal{S}_+^n$  by  $\mathcal{G}[\mu] = \int xx^T d\mu$ . Then define  $\mathcal{F}_\lambda : \mathcal{S}_{++}^n \rightarrow \mathbb{R}$  by

$$\mathcal{F}_\lambda[G] := \sup_{\mu \in \mathcal{G}^{-1}(G)} \left[ S(\mu) - \lambda \int U d\mu \right],$$

where  $S$  is the differential entropy:

$$S(\mu) = \begin{cases} - \int \log \frac{d\mu}{dx} d\mu, & \text{if } d\mu \ll dx \\ -\infty, & \text{otherwise.} \end{cases}$$

By the results of Part III,  $\mathcal{F}_\lambda = \Omega_\lambda^*$ , i.e.,  $\mathcal{F}_\lambda$  is the concave conjugate of  $\Omega_\lambda$ , and vice-versa, with respect to the inner product  $\langle A, G \rangle = \frac{1}{2} \text{Tr}[AG]$ , and the mapping  $G_\lambda$  defined above is a bijection  $\mathcal{S}^n \rightarrow \mathcal{S}_{++}^n$  with inverse given by  $A_\lambda[G] := \nabla \mathcal{F}_\lambda[G]$ . In fact, the supremum in the definition of  $\mathcal{F}_\lambda$  is attained uniquely by the Gibbs measure with density proportional to

$$e^{-\frac{1}{2}x^T A_\lambda x - \lambda U(x)}.$$

Moreover, the *Luttinger-Ward (LW) functional*, defined

$$\Phi_\lambda[G] := 2\mathcal{F}_\lambda[G] - \text{Tr} \log(2\pi eG)$$

is continuous up to the boundary of  $\mathcal{S}_{++}^n$ , and  $\Phi_0 \equiv 0$ . Finally, defining the *self-energy* by  $\Sigma_\lambda[G] := \nabla \Phi_\lambda[G]$ , the following Dyson equation holds for any  $G$ :

$$G^{-1} = A_\lambda[G] - \Sigma_\lambda[G],$$

and likewise, for any  $A$ :

$$(G_\lambda[A])^{-1} = A - \Sigma_\lambda[G_\lambda[A]].$$

We are interested in studying the asymptotic behavior (as  $\lambda \rightarrow \infty$ ) of  $\Phi_\lambda[G]$  for fixed  $G \in \mathcal{S}_{++}^n$ . The *transformation rule* implies that

$$\Phi_\lambda[G] = \Phi_1[\lambda^{1/2}G],$$

so this is the same as studying the behavior of the LW functional for fixed interaction and large  $G$ . Therefore, understanding these asymptotics will complement our existing understanding of the asymptotics of the LW functional near the boundary of  $\mathcal{S}_{++}^n$ . To understand these asymptotics it suffices to understand the asymptotics of  $\mathcal{F}_\lambda[G]$  for fixed  $G$ .

### 3 Heuristic asymptotics

Via the aforementioned Legendre duality, we can write

$$\mathcal{F}_\lambda[G] = \text{Tr}[A_\lambda[G]G] - 2\Omega_\lambda[A_\lambda[G]].$$

Then to study the asymptotic behavior of interest, it is of interest to study the asymptotic behavior of  $A_\lambda[G]$ , and indeed this is the perspective that we will take for now. In fact, our rigorous developments will run in the opposite direction: first we establish the asymptotics of  $\mathcal{F}_\lambda[G]$  and then use these to derive asymptotics for  $A_\lambda$ . Nonetheless, before proving anything rigorous about the asymptotics of  $A_\lambda[G]$ , we use heuristic reasoning to figure out what we want to prove.

Recall  $A_\lambda[G]$  is the unique  $A$  such that  $G_\lambda[A] = G$ . Fixing  $G$  and omitting it from the notation, we write  $A_\lambda = A_\lambda[G]$ . We will try to identify  $\tilde{A}_\lambda$  such that  $G_\lambda[\tilde{A}_\lambda] \approx G$ , where the approximation becomes better as  $\lambda \rightarrow \infty$ . Then one hopes that, in the limit as  $\lambda \rightarrow \infty$ , the asymptotics of  $A_\lambda$  are described by those of  $\tilde{A}_\lambda$ .

Recall that  $G_\lambda[\tilde{A}_\lambda]$  is the second-moment matrix of the Gibbs measure  $\tilde{\mu}_\lambda$  with density  $\tilde{\rho}_\lambda$  proportional to

$$\exp\left(-\frac{1}{2}x^T \tilde{A}_\lambda x - \frac{1}{8} \sum_{ij} \lambda v_{ij} x_i^2 x_j^2\right).$$

Now observe the following identity:

$$\frac{1}{8} \sum_{ij} \lambda v_{ij} (x_i^2 - G_{ii})(x_j^2 - G_{jj}) = \frac{1}{8} \sum_{ij} \lambda v_{ij} x_i^2 x_j^2 - \frac{1}{4} \sum_{ij} \lambda v_{ij} G_{jj} x_i^2 + \frac{1}{8} \sum_{ij} \lambda v_{ij} G_{ii} G_{jj}.$$

Then choose

$$\tilde{A}_\lambda = -\frac{\lambda}{2} \text{diag} \left[ \sum_j v_{1j} G_{jj}, \dots, \sum_j v_{nj} G_{jj} \right] + B,$$

where  $B$  is some symmetric matrix, independent of  $\lambda$ , to be determined later. (We will follow the MATLAB notational convention that ‘diag’ applied to a matrix returns the vector of diagonal entries of this matrix, while ‘diag’ applied to a vector returns the diagonal matrix with diagonal entries given by this vector.)

We then see that

$$\tilde{\rho}_\lambda \propto \exp\left(-\frac{1}{2}x^T Bx - \frac{\lambda}{8} \sum_{ij} v_{ij} (x_i^2 - G_{ii})(x_j^2 - G_{jj})\right).$$

Notice that in the limit as  $\lambda \rightarrow \infty$ , this measure concentrates around the minimizers of

$$f(x) := \frac{1}{8} \sum_{ij} v_{ij} (x_i^2 - G_{ii})(x_j^2 - G_{jj}).$$

Since  $v$  is positive definite, the minimum is attained if and only if  $x_i^2 = G_{ii}$  for all  $i$ , so the minimizers are precisely the vertices of a rectangular prism in dimension  $n$ , i.e., the points  $(\sigma_1\sqrt{G_{11}}, \dots, \sigma_n\sqrt{G_{nn}})$ , where each  $\sigma_i \in \{-1, 1\}$ . When  $G_{ii} = 1$  for all  $i$ , these are the vertices of the unit cube, i.e.,  $\{-1, 1\}^n$ . More generally, we can write this set as  $g\{-1, 1\}^n$ , where we define  $g = \text{diag}\sqrt{\text{diag}(G)}$ .

In fact,  $e^{-\lambda f}$ , suitably normalized, converges (weakly) to the singular measure

$$\sum_{x \in g\{-1, 1\}^n} [\det \nabla^2 f(x)]^{-1/2} \cdot \delta_x$$

supported on  $g\{-1, 1\}^n$ .

We compute  $\nabla^2 f$  explicitly:

$$\nabla^2 f(x) = v \odot (xx^T) + \frac{1}{2} \text{diag} \left[ \sum_j v_{1j}(x_j^2 - G_{jj}), \dots, \sum_j v_{nj}(x_j^2 - G_{jj}) \right].$$

Thus, for  $z \in g\{-1, 1\}^n$ , we have

$$\nabla^2 f(z) = v \odot (zz^T) = \text{diag}(z) v \text{diag}(z) \succ 0,$$

and

$$\det \nabla^2 f(z) = \det v \det g^2 = \det v \prod_{i=1}^n G_{ii}$$

which is independent of  $z$ . Thus,  $e^{-\lambda f}$ , suitably normalized, converges to the uniform measure on  $g\{-1, 1\}^n$ , written

$$\mu := 2^{-n} \sum_{x \in g\{-1, 1\}^n} \delta_x.$$

It follows that the Gibbs measure  $\tilde{\mu}_\lambda$ , which has density  $\tilde{\rho}_\lambda$  satisfying

$$\tilde{\rho}_\lambda(x) \propto e^{-\frac{1}{2}x^T Bx} e^{-\lambda f(x)},$$

converges weakly to

$$\mu_B := \frac{1}{\widehat{Z}[B]} \sum_{z \in g\{-1, 1\}^n} e^{-\frac{1}{2}z^T Bz} \cdot \delta_z,$$

where  $\widehat{Z}[B]$  is the appropriate normalizing constant. (We use hats in the notation here and in the following to distinguish from the  $Z$  and  $\Omega$  already defined above.)

Now  $\tilde{\mu}_\lambda \rightarrow \mu_B$  weakly, and one can show that in fact  $G_\lambda[\tilde{A}_\lambda] \rightarrow \int xx^T d\mu_B$ . Recall that we wanted to choose  $\tilde{A}_\lambda$  such that  $G_\lambda[\tilde{A}_\lambda] \rightarrow G$ . We have already chosen a  $\lambda$ -dependent (diagonal) part of  $\tilde{A}_\lambda$  that ensures that the measure  $\tilde{\mu}_\lambda$  concentrates on the set  $g\{-1, 1\}^n$ . We are then still free to choose the constant part  $B$  such that

$$G = \int xx^T d\mu_B = \frac{1}{\widehat{Z}[B]} \sum_{z \in g\{-1, 1\}^n} zz^T e^{-\frac{1}{2}z^T Bz}.$$

Notice that automatically  $\int x_i^2 d\mu_B = G_{ii}$ , so we are only concerned with matching the off-diagonal entries of  $G$ .

The details of the construction of  $B$  rely on a concave duality, entirely analogous to the one already introduced but in the setting of *spin systems*. Precise details will be accounted for in the following section, but for now we outline what is needed in order to complete our heuristic discussion of the asymptotics as  $\lambda \rightarrow \infty$ .

Define  $\widehat{\Omega}[B] = -\log \widehat{Z}[B]$ , and consider  $\widehat{\Omega}$  as a functional on the set  $\widehat{\mathcal{S}}^n$  of symmetric matrices with zeros on the diagonal. (The diagonal entries of  $B$  do not affect the corresponding Gibbs measure, so any arbitrary but fixed choice of diagonal entries makes sense. Notice that we only need to match the off-diagonal entries of  $G$ , so we expect to have ‘enough’ parameters in  $B$  to do so.)

Then define  $\widehat{G}[B] := \nabla_B \widehat{\Omega}[B]$ , where the gradient is defined appropriately within the set  $\widehat{\mathcal{S}}^n$  so that  $\widehat{G}[B]$  consists of a matrix with the off-diagonal entries ‘absent.’ Abusing notation somewhat and identifying  $\widehat{G}[B]$  also with the corresponding matrix with fixed diagonal entries  $G_{ii}$  filled in, we are trying to find  $B$  such that  $G = \widehat{G}[B]$ .

One can show that  $\widehat{\Omega}$  is smooth and strictly concave on  $\widehat{\mathcal{S}}^n$  and consider the concave conjugate  $\widehat{\Omega}^*$ . By general concave duality theory, one sees that  $\widehat{G}$  is a bijection  $\widehat{\mathcal{S}}^n \rightarrow \text{dom}(\widehat{\Omega}^*)$ , where  $\text{dom}(\widehat{\Omega}^*)$  the set on which  $\widehat{\Omega}^* > -\infty$ . However, as we shall see,  $\text{dom}(\widehat{\Omega}^*)$  is *not* the set of all off-diagonal parts of positive definite matrices, i.e., there exist positive definite matrices whose off-diagonal parts are *not* in  $\text{dom}(\widehat{\Omega}^*)$ . Such  $G$  must then be left out of the scope of the current inquiry.

We can nonetheless provide a fairly explicit description of the set  $\text{dom}(\widehat{\Omega}^*)$ . By similar considerations as in the Luttinger-Ward duality picture, we find that, as a functional of  $\widehat{G}$ ,

$$\widehat{\mathcal{F}}[G] := \widehat{\Omega}^*[\widehat{G}] = \sup_{p \rightarrow \widehat{G}} S(p),$$

where the supremum is over probability vectors

$$p = (p_z)_{z \in g\{-1,1\}^n}$$

such that

$$\widehat{G} = \sum_{z \in g\{-1,1\}^n} z z^T p_z,$$

and  $S(p)$  is the discrete entropy:

$$S(p) = - \sum_{z \in g\{-1,1\}^n} p_z \log p_z.$$

Thus  $\text{dom}(\widehat{\Omega}^*)$  is the set of  $\widehat{G}$  for which this supremum set is nonempty, i.e., the convex hull  $K_g$  of the set

$$V_g := \{z z^T : z \in g\{-1,1\}^n\}.$$

Let  $\mathcal{E}_g$  denote the set of positive definite matrices with diagonal  $g^2$ . When  $g = I_n$ , this set is referred to as the set of ‘correlation matrices,’ or, alternatively, the elliptope. This is a convex set, and there is a large literature on its geometry. It turns out that in dimension  $n \geq 3$ ,  $K_g \neq \mathcal{E}_g$ . In the language of convex geometry, the set of extremal points of  $\mathcal{E}_g$  is larger than  $V_g$ , though the set of vertices of  $\mathcal{E}_g$  turns out to be precisely  $V_g$ . For an example of a  $3 \times 3$  matrix that is in  $\mathcal{E}_{I_3}$  but not  $K_{I_3}$ , consider

$$\begin{pmatrix} 1 & -\frac{2}{5} & -\frac{2}{5} \\ -\frac{2}{5} & 1 & -\frac{2}{5} \\ -\frac{2}{5} & -\frac{2}{5} & 1 \end{pmatrix}.$$

Let  $K_n = \bigcup_g K_g$ , where the union is over all positive diagonal  $n \times n$  matrices. Henceforth, we assume that  $G \in K_n$ .

Returning to our discussion, under the condition that  $G \in K_n$  (and hence  $G \in K_g$  for  $g$  defined with respect to  $G$  as above), we have that there exists a unique  $B = B[G] \in \widehat{\mathcal{S}}^n$  such that  $\widehat{G}[B] = G$ . Then *this* shall be our particular choice of  $B$ , for a given  $G$ .

In summary, we have chosen

$$\widetilde{A}_\lambda = -\lambda D[G] + B[G],$$

where

$$D[G] := \frac{1}{2} \text{diag} \left[ \sum_j v_{1j} G_{jj}, \dots, \sum_j v_{nj} G_{jj} \right],$$

and we have shown that  $G_\lambda[\widetilde{A}_\lambda] \rightarrow G$ . Note that  $D$  depends only on the diagonal elements of  $G$ , and recall that  $v$  is fixed throughout, hence does not appear in the notation. Thus, as was the premise for this discussion, we expect that  $\widetilde{A}_\lambda$  models the asymptotics of  $A_\lambda$ .

## 4 Spin system duality

In the discussion above, we defined a duality related to Gibbs measures for spin systems. In fact, we have actually considered a duality for each possible choice of diagonal elements of  $G$ , which fix the vertices of a rectangular prism on which our measures are localized. To make this clear in the notation, we shall fix  $G_{ii}^* > 0$  and define  $g^*$  with respect to  $G_{ii}^*$  to be the diagonal matrix with  $g_{ii}^* = \sqrt{G_{ii}^*}$ . The possible choices of  $g^*$  will parametrize our dualities, which will be related to each other by appropriate scaling transformations. Note carefully that the ‘\*’ notation *does not* indicate the adjoint anywhere in this Part.

As suggested above, define

$$\widehat{Z}[B | g^*] = \sum_{z \in g^* \{-1, 1\}^n} e^{-\frac{1}{2} z^T B z}, \quad \widehat{\Omega}[B | g^*] = -\log \widehat{Z}[B | g^*]$$

for  $B \in \mathcal{S}^n$ . Then define

$$\widehat{G}[B | g^*] := \nabla_B \widehat{\Omega}[B | g^*] = \frac{1}{\widehat{Z}[B | g^*]} \sum_{z \in g^* \{-1,1\}^n} z z^T e^{-\frac{1}{2} z^T B z}.$$

Also define reference functionals associated to the unit cube via

$$\widehat{Z}[B] := \widehat{Z}[B | I_n], \quad \widehat{\Omega}[B] = \widehat{\Omega}[B | I_n], \quad \widehat{G}[B] := \widehat{G}[B | I_n].$$

Observe that the scaled functional can be related to the reference functionals via

$$\widehat{Z}[B | g^*] = \widehat{Z}[g^* B g^*], \quad \widehat{\Omega}[B | g^*] = \widehat{\Omega}[g^* B g^*], \quad \widehat{G}[B | g^*] := g^* \widehat{G}[g^* B g^*] g^*.$$

Now let  $\mathcal{P}^n$  be the set of probability vectors  $p = (p_\sigma)_{\sigma \in \{-1,1\}^n}$  on the vertices of the unit cube, and define  $\widehat{\mathcal{G}}(\cdot | g^*) : \mathcal{P}^n \rightarrow \mathcal{S}_+^n$  via

$$\widehat{\mathcal{G}}(p | g^*) = \sum_{\sigma \in \{-1,1\}^n} (g^* \sigma \sigma^T g^*) p_\sigma.$$

Finally, define a reference  $\widehat{\mathcal{G}}(p) := \widehat{\mathcal{G}}(p | I_n)$ , so  $\widehat{\mathcal{G}}(p | g^*) = g^* \widehat{\mathcal{G}}(p) g^*$ .

Then we can summarize our spin system duality in the following proposition:

**Proposition 1.** *The concave functional  $\widehat{\mathcal{F}}[\cdot | g^*] : K_{g^*} \rightarrow \mathbb{R}$  defined by*

$$\widehat{\mathcal{F}}[G | g^*] := \sup_{p \in \mathcal{P}^n : \widehat{\mathcal{G}}[p | g^*] = G} S(p)$$

is the concave conjugate of  $\widehat{\Omega}[\cdot | g^*]$  with respect to the inner product  $\langle B, G \rangle = \frac{1}{2} \text{Tr}[BG]$ , and it is  $C^\infty$ -smooth on  $\text{relint}(K_{g^*})$  as a function of the off-diagonal entries of its argument. Moreover  $\widehat{G}[\cdot | g^*]$  is a bijection  $\widehat{\mathcal{S}}^n \rightarrow \text{relint}(K_{g^*})$  with inverse  $B[\cdot | g^*] = \widehat{\nabla}_G \widehat{\mathcal{F}}[\cdot | g^*]$ , where  $\widehat{\nabla}$  is the gradient in the off-diagonal directions, and the supremum in the definition of  $\widehat{\mathcal{F}}[G | g^*]$  is attained uniquely at  $p$  defined by  $p_z \propto e^{-\frac{1}{2} z^T B z}$ . Finally, the functionals  $\widehat{\mathcal{F}}$  can be related to the reference functional  $\widehat{\mathcal{F}}[\cdot] := \widehat{\mathcal{F}}[\cdot | I_n]$  via

$$\widehat{\mathcal{F}}[G | g^*] = \widehat{\mathcal{F}}[(g^*)^{-1} G (g^*)^{-1}].$$

*Remark 2.* Note that we have extended the domain of  $\widehat{\Omega}$  from the domain considered in our heuristic discussion of the previous section, namely  $\widehat{\mathcal{S}}^n$ , to all of  $\mathcal{S}^n$ . This extension is then affine along any segment with direction given by a diagonal matrix, hence not *strictly* concave. On the dual side, the conjugate is (negative) infinite except on the set of  $G$  with  $G_{ii} = G_{ii}^*$ . In fact, the domain is even smaller: it is  $K_{g^*}$ , as discussed earlier.

*Proof.* Define  $T_{g^*} : \mathcal{S}^n \rightarrow \mathcal{S}^n$  via  $A \mapsto g^* A g^*$ . Then since  $\widehat{\Omega}[\cdot | g^*] = \widehat{\Omega} \circ T_{g^*}$ , it follows (as a general property of concave conjugation) that the concave conjugate of  $\widehat{\Omega}[\cdot | g^*]$  is  $\widehat{\Omega}^* \circ T_{g^*}^{-1}$ . This proves the relation to the reference functional, so for the remainder of the proof we can consider the case  $g^* = I_n$ .

Fix  $B \in \mathcal{S}^n$ , and let  $w \in \mathcal{P}^n$  be defined by  $w_\sigma = \frac{1}{Z[B]} e^{-\frac{1}{2}\sigma^T B \sigma}$ . Then for  $p \in \mathcal{P}^n$ , the relative entropy

$$S_w(p) = - \sum_{\sigma \in \{-1,1\}^n} p_\sigma \log \left( \frac{p_\sigma}{w_\sigma} \right)$$

satisfies

$$S_w(p) \leq 0$$

with equality if and only if  $p = w$ . But

$$S_w(p) = S(p) + \sum_{\sigma \in \{-1,1\}^n} p_\sigma \log w_\sigma = S(p) + \widehat{\Omega}[B] - \frac{1}{2} \sum_{\sigma \in \{-1,1\}^n} \sigma^T B \sigma p_\sigma.$$

Thus we have established that

$$\widehat{\Omega}[B] \leq \frac{1}{2} \text{Tr} [B \widehat{\mathcal{G}}(p)] - S(p)$$

with equality if and only if  $p = w$ , so

$$\widehat{\Omega}[B] = \inf_{p \in \mathcal{P}^n} \left( \frac{1}{2} \text{Tr} [B \widehat{\mathcal{G}}(p)] - S(p) \right) = \sup_G \left( \frac{1}{2} \text{Tr}[BG] - \sup_{p \in \mathcal{P}^n : \mathcal{G}(p)=G} S(p) \right),$$

i.e.,

$$\widehat{\Omega}[B] = \sup_G \left( \frac{1}{2} \text{Tr}[BG] - \widehat{\mathcal{F}}[G] \right).$$

The result then follows by further mimicking the techniques of Part III. (If we restrict  $\widehat{\Omega}$  to a subspace with fixed diagonal, then it is ‘of Legendre type’ in the sense of [91]. This means that the concave conjugate is as well, and their gradients are inverse mappings of one another.)  $\square$

## 5 Rigorous asymptotics

Let  $G_{ii}^* > 0$  and define  $g^*$  with respect to  $G_{ii}^*$  to be the diagonal matrix with  $g_{ii}^* = \sqrt{G_{ii}^*}$ . Then define  $f = f_{g^*}$  as above by

$$f_{g^*}(x) = \frac{1}{8} \sum_{ij} v_{ij} (x_i^2 - G_{ii}^*) (x_j^2 - G_{jj}^*).$$

Note that  $f_{g^*}$  is nonnegative with minimum value zero, and the minimizers are given by  $g^*\{-1, 1\}^n$ . Observe that for  $z \in g^*\{-1, 1\}^n$ ,  $\nabla^2 f_{g^*}(z) = v \odot (zz^T)$ , so  $\det \nabla^2 f_{g^*}(z) = \det v (\det g^*)^2 = \det(v) \prod_{i=1}^n G_{ii}^*$ .

Next define a functional with domain  $\mathcal{S}_{++}^n$  via

$$\overline{\mathcal{F}}_\lambda[G | g^*] := \sup_{\mu \in \mathcal{G}^{-1}(G)} \left[ S(\mu) - \lambda \int f_{g^*} d\mu \right].$$

By the results of Part III,  $\overline{\mathcal{F}}_\lambda[\cdot, g^*] = \overline{\Omega}_\lambda^*$ , where  $\overline{\Omega}_\lambda : \mathcal{S}^n \rightarrow \mathbb{R}$  is defined by

$$\overline{\Omega}_\lambda[B | g^*] := -\log \int e^{-\frac{1}{2}x^T Bx - \lambda f_{g^*}(x)} dx.$$

Notice that for  $\mu$  with  $\int x_i^2 d\mu = G_{ii}^*$ ,

$$\begin{aligned} \lambda \int f d\mu &= \int U_\lambda d\mu - \frac{\lambda}{4} \sum_{ij} v_{ij} G_{jj}^* \int x_i^2 d\mu + \frac{\lambda}{8} \sum_{ij} v_{ij} G_{ii}^* G_{jj}^* \\ &= \int U_\lambda d\mu - \frac{\lambda}{4} \sum_{ij} v_{ij} G_{ii}^* G_{jj}^* + \frac{\lambda}{8} \sum_{ij} v_{ij} G_{ii}^* G_{jj}^* \\ &= \int U_\lambda d\mu - \frac{\lambda}{8} \sum_{ij} v_{ij} G_{ii}^* G_{jj}^*. \end{aligned}$$

Therefore, for  $G \in \mathcal{S}_{++}^n$ ,  $g = \text{diag} \sqrt{\text{diag}(G)}$ , we have the equality

$$\overline{\mathcal{F}}_\lambda[G | g] = \mathcal{F}_\lambda[G] + \frac{\lambda}{8} \sum_{ij} v_{ij} G_{ii} G_{jj}.$$

Note that the last term on the right-hand side corresponds to the Hartree contribution to the Luttinger-Ward functional. (Note that the right-hand side is the sum of a concave function and a convex function, hence not necessarily concave on all of  $\mathcal{S}_{++}^n$ . However, it is concave in the off-diagonal part of  $G$  when the diagonal entries are fixed.)

Then we will prove:

**Theorem 3.** *For any  $G \in \text{int}(K_n)$  and  $g = \text{diag} \sqrt{\text{diag}(G)}$ ,*

$$\overline{\mathcal{F}}_\lambda[G | g] + \frac{n}{2} \log \left( \frac{\lambda}{2\pi} \right) = \widehat{\mathcal{F}}[G | g] - \frac{1}{2} \text{Tr} \log v - \frac{1}{2} \sum_{i=1}^n \log G_{ii} + O(\lambda^{-1}).$$

*Consequently*

$$\mathcal{F}_\lambda[G] + \frac{n}{2} \log \left( \frac{\lambda}{2\pi} \right) + \frac{\lambda}{8} \sum_{ij} v_{ij} G_{ii} G_{jj} = \widehat{\mathcal{F}}[G | g] - \frac{1}{2} \text{Tr} \log v - \frac{1}{2} \sum_{i=1}^n \log G_{ii} + O(\lambda^{-1}).$$

*In fact the constant in  $O(\lambda^{-1})$  can be taken to be locally uniform in  $G$ .*



*Remark 4.* In fact, for fixed  $g^*$ ,

$$\overline{\mathcal{F}}_\lambda[\cdot, g^*] + \frac{n}{2} \log \left( \frac{\lambda}{2\pi} \right) \rightarrow \widehat{\mathcal{F}}[\cdot, g^*] - \frac{1}{2} \text{Tr} \log v - \frac{1}{2} \sum_{i=1}^n \log G_{ii}^*$$

in the sense of hypo-convergence as well as pointwise convergence. This perspective (i.e., that of foliated dualities) will be helpful in proving one of the bounds necessary for the proof.

The remainder of this section will be dedicated to proving the theorem.

We divide the proof into a proof an upper bound and a lower bound for  $\overline{\mathcal{F}}_\lambda$ . Throughout we will use  $C$  to denote constants independent of  $\lambda$  that are sufficiently large in the given context; the precise meaning of  $C$  may change across usages. The reader should observe that constants used in the proof can be taken to be locally uniform in  $G$ , though we will not make explicit mention of this throughout.

## I. The upper bound.

Let  $G \in \text{int}(K_n)$  and  $g = \text{diag} \sqrt{\text{diag}(G)}$ . Recall that for any  $g^*$ ,  $\overline{\mathcal{F}}_\lambda[\cdot, g^*] = \overline{\Omega}_\lambda^*$ . Thus, in particular, setting  $g^* = g$ , we have the equation

$$\overline{\mathcal{F}}_\lambda[G | g] = \inf_{B \in \mathcal{S}^n} \left( \frac{1}{2} \text{Tr}[BG] - \overline{\Omega}_\lambda[B | g] \right).$$

Fix  $B$  to be the unique matrix in  $\widehat{\mathcal{S}}^n$  such that  $\widehat{G}[B | g] = G$ , furnished by the spin system duality of Proposition 1 above. The Legendre correspondence between  $B$  and  $G$  ensures  $\widehat{\Omega}[B | g] = \frac{1}{2} \text{Tr}[BG] - \widehat{\mathcal{F}}[G | g]$ . Then, making this particular choice of  $B$  in the above infimum, we have

$$\overline{\mathcal{F}}_\lambda[G | g] \leq \frac{1}{2} \text{Tr}[BG] - \overline{\Omega}_\lambda[B | g],$$

and we will use this inequality to prove the needed asymptotic upper bound.

We write  $f = f_g$  and observe via Laplace's method (see, e.g., p. 495 of [107] for a precise statement) that

$$\int e^{-\frac{1}{2}x^T Bx - \lambda f(x)} dx = \left( \frac{2\pi}{\lambda} \right)^{n/2} [\det \nabla^2 f(z)]^{-1/2} e^{-\frac{1}{2}z^T Bz} [1 + O(\lambda^{-1})].$$

Therefore we can compute

$$\overline{\Omega}_\lambda[B | g] = -\log \int e^{-\frac{1}{2}x^T Bx - \lambda f(x)} dx$$

$$\begin{aligned}
&= -\log \left( \sum_{z \in g\{-1,1\}^n} \left( \frac{2\pi}{\lambda} \right)^{n/2} [\det \nabla^2 f(z)]^{-1/2} e^{-\frac{1}{2}z^T B z} \right) + O(\lambda^{-1}) \\
&= -\log \left( \left( \frac{2\pi}{\lambda} \right)^{n/2} \frac{1}{\sqrt{\prod_{i=1}^n G_{ii}}} \frac{1}{\sqrt{\det v}} \sum_{z \in g\{-1,1\}^n} e^{-\frac{1}{2}z^T B z} \right) + O(\lambda^{-1}) \\
&= \frac{n}{2} \log \left( \frac{\lambda}{2\pi} \right) + \frac{1}{2} \sum_{i=1}^n \log G_{ii} + \frac{1}{2} \text{Tr} \log v + \widehat{\Omega}[B | g] + O(\lambda^{-1}).
\end{aligned}$$

Now recall that  $\widehat{\Omega}[B | g] = \frac{1}{2} \text{Tr}[BG] - \widehat{\mathcal{F}}[G | g]$ , so in fact the preceding implies

$$\overline{\mathcal{F}}_\lambda[G | g] + \frac{n}{2} \log \left( \frac{\lambda}{2\pi} \right) \leq \widehat{\mathcal{F}}[G | g] - \frac{1}{2} \text{Tr} \log v - \frac{1}{2} \sum_{i=1}^n \log G_{ii} + O(\lambda^{-1}).$$

This completes the proof of the upper bound.

## II. The lower bound.

Fix  $G \in \text{int}(K_n)$  and  $g = \text{diag} \sqrt{\text{diag}(G)}$ . Throughout we write  $f = f_g$ . In this section the goal is to show that

$$\overline{\mathcal{F}}_\lambda[G | g] + \frac{n}{2} \log \left( \frac{\lambda}{2\pi} \right) \geq \widehat{\mathcal{F}}[G | g] - \frac{1}{2} \text{Tr} \log v - \frac{1}{2} \sum_{i=1}^n \log G_{ii} + O(\lambda^{-1}).$$

The proof of this lower bound is considerably more involved, and we break it into steps.

To get such a lower bound, we will make a specific choice of  $\mu_\lambda$  in the supremizing set from the definition of  $\overline{\mathcal{F}}_\lambda$ . The next section deals with the construction of this measure, which is rather delicate.

### II.1. Choosing $\mu_\lambda \in \mathcal{G}^{-1}[G]$ .

Since  $G \in K_n$ , let  $B$  be the unique matrix in  $\widehat{\mathcal{S}}^n$  such that  $\widehat{G}[B | g] = G$ , furnished by the spin system duality of Proposition 1 above. Then as a first guess for a specific choice of  $\mu_\lambda$ , let  $\mu'_{z,\lambda} \sim \mathcal{N}(z, \lambda^{-1} H_z^{-1})$  for  $z \in g\{-1,1\}^n$ , where  $H_z := \nabla^2 f(z) = v \odot (zz^T)$ , and consider the convex combination  $\mu'_\lambda := \frac{1}{\widehat{Z}[B | g]} \sum_{z \in g\{-1,1\}^n} e^{-\frac{1}{2}z^T B z} \mu'_{z,\lambda}$ . This is a mixture of Gaussians that weakly converges to the singular measure

$$\mu := \frac{1}{\widehat{Z}[B | g]} \sum_{z \in g\{-1,1\}^n} e^{-\frac{1}{2}z^T B z} \cdot \delta_z.$$

While  $\int xx^T d\nu = G$  by construction, unfortunately  $\int xx^T d\mu'_\lambda \neq G$ . However, equality does hold approximately for large  $\lambda$ , and moreover, the inequality has a direction. Naively, one expects the error to be  $\sim \lambda^{-1/2}$ , hence that we must correct  $\mu'_\lambda$  by  $\sim \lambda^{-1/2}$  to obtain the desired  $G$ . In fact, due to a special cancellation, the error is only  $\sim \lambda^{-1}$ . (Remarkably, this fact will be essential in getting a lower bound with this approach.) Indeed, write

$$\begin{aligned} \int xx^T d\mu'_\lambda &= \frac{1}{\widehat{Z}[B|g]} \sum_{z \in g\{-1,1\}^n} e^{-\frac{1}{2}z^T Bz} \int xx^T d\mu'_{z,\lambda} \\ &= \frac{1}{\widehat{Z}[B|g]} \sum_{z \in g\{-1,1\}^n} e^{-\frac{1}{2}z^T Bz} \\ &\quad \int [zz^T + z(x-z)^T + (x-z)z^T + (x-z)(x-z)^T] d\mu'_{z,\lambda}. \end{aligned}$$

Now  $\mu'_{z,\lambda}$  is a Gaussian centered at  $z$ , so  $\int z(x-z)^T d\mu'_{z,\lambda} = \int (x-z)z^T d\mu'_{z,\lambda} = 0$ , and we have in fact that

$$\begin{aligned} \int xx^T d\mu'_\lambda &= \frac{1}{\widehat{Z}[B|g]} \sum_{z \in g\{-1,1\}^n} e^{-\frac{1}{2}z^T Bz} \int [zz^T + (x-z)(x-z)^T] d\mu'_{z,\lambda} \\ &= G + \frac{1}{\lambda} \frac{1}{\widehat{Z}[B|g]} \sum_{z \in g\{-1,1\}^n} e^{-\frac{1}{2}z^T Bz} H_z^{-1} \\ &= G + \lambda^{-1} \frac{1}{\widehat{Z}[B|g]} \sum_{z \in g\{-1,1\}^n} e^{-\frac{1}{2}z^T Bz} [v \odot (zz^T)]^{-1}. \end{aligned}$$

From here we can already see that  $\int xx^T d\mu'_\lambda$  must overestimate  $G$  by an error term of size  $O(\lambda^{-1})$ , but it is extremely worthwhile to continue simplifying this expression.

Now  $v \odot (zz^T) = \text{diag}(z) v \text{diag}(z)$ , so

$$[v \odot (zz^T)]^{-1} = [\text{diag}(z)]^{-1} v^{-1} [\text{diag}(z)]^{-1}.$$

If we write  $\sigma = g^{-1}z \in \{-1,1\}^n$ , then  $\text{diag}(z) = \text{diag}(\sigma)g$ , and  $[\text{diag}(\sigma)]^{-1} = \text{diag}(\sigma)$ , so

$$[v \odot (zz^T)]^{-1} = g^{-1} [v^{-1} \odot (\sigma\sigma^T)] g^{-1} = g^{-2} [v^{-1} \odot (g\sigma^T)] g^{-2}$$

Therefore

$$\begin{aligned} \int xx^T d\mu'_\lambda &= G + \lambda^{-1} g^{-2} \left( \frac{1}{\widehat{Z}[B|g]} \sum_{z \in g\{-1,1\}^n} e^{-\frac{1}{2}z^T Bz} [v^{-1} \odot (zz^T)] \right) g^{-2} \\ &= G + \lambda^{-1} g^{-2} \left[ v^{-1} \odot \left( \frac{1}{\widehat{Z}[B|g]} \sum_{z \in g\{-1,1\}^n} e^{-\frac{1}{2}z^T Bz} zz^T \right) \right] g^{-2} \end{aligned}$$

$$\begin{aligned}
&= G + \lambda^{-1}g^{-2} [v^{-1} \odot G] g^{-2} \\
&= G + (\lambda^{-1}g^{-2}v^{-1}g^{-2}) \odot G.
\end{aligned}$$

We will try to use these calculations to see whether a *different* choice of measure might have given the desired second moments. For now we give ourselves the freedom to choose  $G_\lambda$ , which shall ultimately be  $G + O(\lambda^{-1})$ . Then we will define  $g_\lambda$  and  $B_\lambda$  in terms of  $G_\lambda$  by letting  $g_\lambda = \text{diag} \sqrt{\text{diag}(G_\lambda)}$  and choosing  $B_\lambda$  to be the unique matrix in  $\widehat{\mathcal{S}}^n$  such that  $\widehat{G}[B_\lambda | g_\lambda] = G_\lambda$ . (This will be possible because for  $\lambda$  sufficiently large, we will have  $G_\lambda \in K_n$ .)

Then let  $f_\lambda := f_{g_\lambda}$ , and let  $\mu_{z,\lambda} \sim \mathcal{N}(z, \lambda^{-1}H_{z,\lambda}^{-1})$  for  $z \in g_\lambda\{-1, 1\}^n$ , where  $H_{z,\lambda} := \nabla^2 f_\lambda(z) = v \odot (zz^T)$ , so that  $\det H_{z,\lambda} = \det v \det g_\lambda^2$ . Consider the convex combination

$$\mu_\lambda := \frac{1}{\widehat{Z}[B_\lambda | g_\lambda]} \sum_{z \in g_\lambda\{-1, 1\}^n} e^{-\frac{1}{2}z^T B_\lambda z} \mu_{z,\lambda}.$$

The preceding calculations establish that

$$\int xx^T d\mu_\lambda = G_\lambda + [\lambda^{-1}g_\lambda^{-2}v^{-1}g_\lambda^{-2}] \odot G_\lambda,$$

so we want our choice of  $G_\lambda$  to be such that the latter expression is equal to the given  $G$ , i.e., such that

$$G_{ij} = G_{\lambda,ij} + \lambda^{-1}v^{ij} (G_{\lambda,ii})^{-1} (G_{\lambda,jj})^{-1} G_{\lambda,ij}.$$

The important thing to realize here is that first we can solve the ‘diagonal’ set of equations for  $G_{\lambda,ii}$ , leaving a set of uncoupled linear equations for the off-diagonal terms. Indeed, observe that the diagonal equations are

$$G_{ii} = G_{\lambda,ii} + \lambda^{-1}v^{ii} (G_{\lambda,ii})^{-1},$$

and we want to solve them for  $G_{\lambda,ii}$ . Fix  $i$ , and let  $b = G_{ii} > 0$ ,  $c = v^{ii} > 0$ . Then the desired  $G_{\lambda,ii}$  is a root of

$$0 = x^2 - bx + \lambda^{-1}c.$$

The discriminant is positive when  $b^2 > 4\lambda^{-1}c$ , which is the case for  $\lambda$  sufficiently large. Moreover, as  $\lambda \rightarrow \infty$ , the roots are

$$x = \frac{b \pm \sqrt{b^2 - 4\lambda^{-1}c}}{2} = \frac{b \pm (b - \frac{2c}{b}\lambda^{-1})}{2} + O(\lambda^{-2}).$$

We choose the larger root, which is  $b - \frac{c}{b}\lambda^{-1} + O(\lambda^{-2})$ , yielding

$$G_{\lambda,ii} = \frac{G_{ii} + \sqrt{(G_{ii})^2 - 4\lambda^{-1}v^{ii}}}{2} = G_{ii} - \frac{v^{ii}}{G_{ii}}\lambda^{-1} + O(\lambda^{-2}).$$

Then the off-diagonal equations,  $i \neq j$ , can be solved simply by choosing

$$G_{\lambda,ij} = \frac{G_{ij}}{1 + \lambda^{-1}v^{ij} (G_{\lambda,ii})^{-1} (G_{\lambda,jj})^{-1}},$$

which makes sense for  $\lambda$  sufficiently large.

In summary we have found  $G_\lambda = G + O(\lambda^{-1})$  such that  $\mu_\lambda$  as defined above satisfies  $\int xx^T d\mu_\lambda = G$ .

## II.2. Arranging terms appropriately.

Now we can (finally!) use the specific choice  $\mu_\lambda$  to get our lower bound. Indeed,

$$\bar{\mathcal{F}}_\lambda[G|g] \geq S(\mu_\lambda) - \int f d\mu_\lambda.$$

Next consider the convex combination  $\nu_\lambda := \frac{1}{2^n} \sum_{z \in g_\lambda \{-1,1\}^n} \mu_{z,\lambda}$ . Let the density of  $\mu_{z,\lambda}$  be denoted by  $\rho_{z,\lambda}$  and that of  $\nu_\lambda$  by  $\rho_\lambda$ . Then write

$$S(\mu_\lambda) = S(\mu_\lambda|\nu_\lambda) - \int \log \rho_\lambda d\mu_\lambda$$

Recall the construction of  $\mu_\lambda$  as the mixture of the same Gaussians  $\mu_{z,\lambda}$  but with weights  $p_{z,\lambda} := \left(\widehat{Z}[B_\lambda|g_\lambda]\right)^{-1} e^{-\frac{1}{2}z^T B_\lambda z}$  for  $z \in g_\lambda \{-1,1\}^n$ . Note that the weight vector  $p_\lambda$  converges to  $p_z := \left(\widehat{Z}[B|g]\right)^{-1} e^{-\frac{1}{2}z^T B z}$ . It is not hard to see (via the smoothness of the Legendre duality) that  $B_\lambda = B + O(\lambda^{-1})$ , so  $p_\lambda = p + O(\lambda^{-1})$ . The hope now is that  $S_{\nu_\lambda}(\mu_\lambda) \rightarrow S_w(p)$ , where the former and latter expressions indicates the continuous and discrete relative entropies, respectively, and  $w$  is the uniform probability vector. Indeed,

$$S_w(p) = S(p) - n \log 2 = \widehat{\mathcal{F}}[G|g] - n \log 2$$

by the spin system duality outlined earlier and our choice of  $B = B[G|g]$ , so this contains the important term that we wanted to end up with in the limit. The other terms will be yielded by the asymptotics of  $\int \log \rho_\lambda d\mu_\lambda$ , to which we now turn.

## II.3. Asymptotics of $\int \log \rho_\lambda d\mu_\lambda$ .

Now, with  $Q_{z,\lambda} := \frac{1}{2}(x-z)^T H_{z,\lambda}(x-z)$  for  $z \in g_\lambda \{-1,1\}^n$ , we have

$$\rho_{z,\lambda} = \left(\frac{\lambda}{2\pi}\right)^{n/2} [\det H_{z,\lambda}]^{1/2} \exp(-\lambda Q_{z,\lambda}),$$

so

$$\rho_\lambda = \left(\frac{\lambda}{2\pi}\right)^{n/2} \frac{1}{2^n} \sum_{z \in g_\lambda \{-1,1\}^n} [\det H_{z,\lambda}]^{1/2} \exp(-\lambda Q_{z,\lambda})$$

$$= \left(\frac{\lambda}{2\pi}\right)^{n/2} \frac{1}{2^n} [\det v \det g_\lambda^2]^{1/2} \sum_{z \in g_\lambda \{-1, 1\}^n} \exp(-\lambda Q_{z, \lambda})$$

Then

$$\begin{aligned} \int \log \rho_\lambda d\mu_\lambda &= \frac{n}{2} \log \left(\frac{\lambda}{2\pi}\right) - n \log 2 + \frac{1}{2} \text{Tr} \log v + \frac{1}{2} \sum_{i=1}^n \log G_{\lambda, ii} \\ &\quad + \int \log \sum_{z \in g_\lambda \{-1, 1\}^n} \exp(-\lambda Q_{z, \lambda}) d\mu_\lambda. \end{aligned}$$

Now roughly speaking, Laplace's principle would indicate that

$$h_\lambda(x) := \log \sum_{z \in g_\lambda \{-1, 1\}^n} \exp(-\lambda Q_{z, \lambda}(x)) \approx -\lambda \min_{z \in g_\lambda \{-1, 1\}^n} Q_{z, \lambda}(x).$$

We want to make this estimate precise to obtain an upper bound for  $\int h_\lambda d\mu_\lambda$ . First observe that  $h_\lambda \leq C$  globally for some constant, but we want to get a sharp bound for  $h_\lambda$  near  $g\{-1, 1\}^n$ .

To this end, let  $\mathcal{N}_z$  be fixed disjoint neighborhoods of the  $z \in g\{-1, 1\}^n$ . These are also neighborhoods of  $g_\lambda\{-1, 1\}^n$  for  $\lambda$  large. To relate our indexing of  $g\{-1, 1\}^n$  with that of  $g_\lambda\{-1, 1\}^n$ , for  $z \in g\{-1, 1\}^n$  we let  $z_\lambda := g^{-1}g_\lambda z$  denote the corresponding point in  $g_\lambda\{-1, 1\}^n$ . Observe that we can choose the neighborhoods  $\mathcal{N}_z$  and a constant  $\Delta > 0$  such that, when  $z \neq z'$ ,

$$Q_{z_\lambda, \lambda} + \Delta \leq Q_{z'_\lambda, \lambda}$$

on  $\mathcal{N}_z$  for all  $\lambda$  sufficiently large. Finally, define  $\mathcal{N} := \bigcup_z \mathcal{N}_z$ .

Next we introduce a quantitative statement of Laplace's principle that is suited to our needs.

**Lemma 5.** *Suppose  $a_1, \dots, a_k \in \mathbb{R}$  and  $a_i$  is strictly minimal among these numbers. Then*

$$\log \sum_{j=1}^k e^{-a_j} \leq -a_i + (k-1)e^{-c},$$

where  $c := \min_{j \neq i} (a_j - a_i) > 0$ .

*Proof.* Compute

$$\begin{aligned} \log \sum_{j=1}^k e^{-a_j} &= \log \sum_{j=1}^k e^{-a_i} e^{-(a_j - a_i)} \\ &= -a_i + \log \sum_{j=1}^k e^{-(a_j - a_i)} \end{aligned}$$

$$\begin{aligned}
&= -a_i + \log \left( 1 + \sum_{j \neq i} e^{-(a_j - a_i)} \right) \\
&\leq -a_i + \log (1 + (k-1)e^{-c}) \\
&\leq -a_i + (k-1)e^{-c},
\end{aligned}$$

using the inequality  $\log(1+x) \leq x$  for  $x \geq 0$ .  $\square$

Then the lemma implies that there exists a constant  $C$  such that, for any  $z \in g\{-1, 1\}^n$ ,

$$h_\lambda(x) \leq -\lambda Q_{z_\lambda, \lambda}(x) + Ce^{-\lambda \Delta}$$

for all  $x \in \mathcal{N}_z$ .

Proceeding, we obtain

$$\begin{aligned}
\int h_\lambda d\mu_\lambda &= \int_{\mathcal{N}^c} h_\lambda d\mu_\lambda + \sum_{z \in g\{-1, 1\}^n} \int_{\mathcal{N}_z} h_\lambda d\mu_\lambda \\
&\leq C\mu_\lambda(\mathcal{N}^c) - \lambda \sum_{z \in g\{-1, 1\}^n} \int_{\mathcal{N}_z} Q_{z_\lambda, \lambda} d\mu_\lambda + Ce^{-\lambda \Delta} \\
&\leq C\mu_\lambda(\mathcal{N}^c) - \lambda \sum_{z \in g\{-1, 1\}^n} p_{z_\lambda, \lambda} \int_{\mathcal{N}_z} Q_{z_\lambda, \lambda} d\mu_{z_\lambda, \lambda} + Ce^{-\lambda \Delta},
\end{aligned}$$

where in the last line we have used the nonnegativity of the  $Q_{z_\lambda, \lambda}$  and the definition of  $\mu_\lambda$  as a convex combination of the  $\mu_{z_\lambda, \lambda}$ . Note that  $\mu'_\lambda(\mathcal{N}^c) = O(e^{-C^{-1}\lambda})$  for some  $C > 0$ , so we in turn have the upper bound

$$\begin{aligned}
\int h_\lambda d\mu_\lambda &\leq -\lambda \sum_{z \in g\{-1, 1\}^n} p_{z_\lambda, \lambda} \int_{\mathcal{N}_z} Q_{z_\lambda, \lambda} d\mu_{z_\lambda, \lambda} + Ce^{-C^{-1}\lambda} \\
&= -\frac{\lambda}{2} \sum_{z \in g\{-1, 1\}^n} p_{z_\lambda, \lambda} \int_{\mathcal{N}_z} (x - z_\lambda)^T H_{z_\lambda, \lambda} (x - z_\lambda) d\mu_{z_\lambda, \lambda} + Ce^{-C^{-1}\lambda}.
\end{aligned}$$

Now want to replace  $H_{z_\lambda, \lambda} = v \odot (z_\lambda z_\lambda^T)$  with  $H_z = v \odot (zz^T)$  in this expression at a cost of only  $O(\lambda^{-1})$ . Indeed, for any  $z \in g\{-1, 1\}^n$ ,

$$\begin{aligned}
\left| \int_{\mathcal{N}_z} (x - z_\lambda)^T [H_{z_\lambda, \lambda} - H_z] (x - z_\lambda) d\mu_{z_\lambda, \lambda} \right| &\leq \|H_{z_\lambda, \lambda} - H_z\| \int_{\mathcal{N}_z} |x - z_\lambda|^2 d\mu_{z_\lambda, \lambda} \\
&\leq \|H_{z_\lambda, \lambda} - H_z\| \frac{1}{\lambda} \text{Tr}[H_{z_\lambda, \lambda}] \\
&= O(\lambda^{-2}),
\end{aligned}$$

so indeed we can make this replacement, which yields

$$\int h_\lambda d\mu_\lambda \leq -\frac{\lambda}{2} \sum_{z \in g\{-1, 1\}^n} p_{z_\lambda, \lambda} \int_{\mathcal{N}_z} (x - z_\lambda)^T H_z (x - z_\lambda) d\mu_{z_\lambda, \lambda} + O(\lambda^{-1}).$$

Next we want to replace  $z_\lambda$  with  $z$  at a cost of only  $O(\lambda^{-1})$ . Observe:

$$\begin{aligned}
& (x - z_\lambda)^T H_z(x - z_\lambda) \\
&= [(x - z) + (z - z_\lambda)]^T H_z [(x - z) + (z - z_\lambda)] \\
&= (x - z)^T H_z(x - z) + 2(z - z_\lambda)^T H_z(x - z) + (z - z_\lambda)^T H_z(z - z_\lambda) \\
&= (x - z)^T H_z(x - z) + 2(z - z_\lambda)^T H_z(x - z_\lambda) - (z - z_\lambda)^T H_z(z - z_\lambda),
\end{aligned}$$

i.e.,

$$(x - z_\lambda)^T H_{z,\lambda}(x - z_\lambda) - (x - z)^T H_{z,\lambda}(x - z) = 2(z - z_\lambda)^T H_{z,\lambda}(x - z_\lambda) - (z - z_\lambda)^T H_{z,\lambda}(z - z_\lambda).$$

The last term on the right-hand side is  $O(\lambda^{-2})$ , independent of  $x$ , hence can be ignored. The other term integrates exactly as

$$2(z - z_\lambda)^T H_{z,\lambda} \int_{\mathcal{N}_z} (x - z_\lambda) d\mu_{z_\lambda,\lambda} = 0.$$

Thus we can make the desired replacement, which yields

$$\begin{aligned}
\int h_\lambda d\mu_\lambda &\leq -\frac{\lambda}{2} \sum_{z \in g\{-1,1\}^n} p_{z_\lambda,\lambda} \int_{\mathcal{N}_z} (x - z)^T H_z(x - z) d\mu_{z_\lambda,\lambda} + O(\lambda^{-1}) \\
&= -\lambda \sum_{z \in g\{-1,1\}^n} p_{z_\lambda,\lambda} \int_{\mathcal{N}_z} Q_z d\mu_{z_\lambda,\lambda} + O(\lambda^{-1}),
\end{aligned}$$

where  $Q_z := \frac{1}{2}(x - z)^T H_z(x - z)$ .

#### II.4. Resynthesizing our bounds.

Combining with our previous work we obtain

$$\begin{aligned}
\bar{\mathcal{F}}_\lambda[G|g] + \frac{n}{2} \log \left( \frac{\lambda}{2\pi} \right) &\geq S_{\nu_\lambda}(\mu_\lambda) + n \log 2 - \frac{1}{2} \text{Tr} \log v - \frac{1}{2} \sum_{i=1}^n \log G_{\lambda,ii} \\
&\quad + \lambda \sum_{z \in g\{-1,1\}^n} p_{z_\lambda,\lambda} \int_{\mathcal{N}_z} Q_z d\mu_{z_\lambda,\lambda} \\
&\quad - \lambda \int f d\mu_\lambda + O(\lambda^{-1}).
\end{aligned}$$

Now

$$\int f d\mu_\lambda = \int_{\mathcal{N}^c} f d\mu_\lambda + \sum_{z \in g\{-1,1\}^n} \int_{\mathcal{N}_z} f d\mu_\lambda$$



$$= \sum_{z \in g\{-1,1\}^n} p_{z_\lambda, \lambda} \int_{\mathcal{N}_z} f d\mu_{z_\lambda, \lambda} + O(\lambda^{-1}),$$

and moreover observe that  $G_{\lambda, ii} = G_{ii} + O(\lambda^{-1})$ , so in fact we have that

$$\begin{aligned} \overline{\mathcal{F}}_\lambda[G | g] + \frac{n}{2} \log \left( \frac{\lambda}{2\pi} \right) &\geq S_{\nu_\lambda}(\mu_\lambda) + n \log 2 - \frac{1}{2} \text{Tr} \log v - \frac{1}{2} \sum_{i=1}^n \log G_{ii} \\ &\quad - \lambda \sum_{z \in g\{-1,1\}^n} p_{z_\lambda, \lambda} \int_{\mathcal{N}_z} (f - Q_z) d\mu_{z_\lambda, \lambda} + O(\lambda^{-1}). \end{aligned}$$

Now, for  $z \in g\{-1, 1\}^n$ , we focus on the term

$$\int_{\mathcal{N}_z} (f - Q_z) d\mu_{z_\lambda, \lambda}$$

with the aim of showing that it is  $O(\lambda^{-2})$ .

## II.5. Showing that $\int_{\mathcal{N}_z} (f - Q_z) d\mu_{z_\lambda, \lambda}$ is negligible.

The key is that the Taylor series for  $P_z := f - Q_z$  vanishes through second order at  $z$ , so  $P_z(x - z) = O(|x - z|^4)$ , and that  $|z_\lambda - z| = O(\lambda^{-1})$ . We want to use these facts to control the Taylor series of  $f - Q_z$  about  $z_\lambda$ . Since the Gaussian  $\mu_{z_\lambda, \lambda}$  is centered at  $z_\lambda$ , the odd part of the Taylor polynomial of  $f - Q_z$  about  $z_\lambda$  does not contribute to the integral against  $\mu_{z_\lambda, \lambda}$ , so we need only control the zeroth, second, and fourth order terms (noting that the Taylor series terminates at fourth order since  $f - Q_z$  is quartic).

The zeroth order term at  $z_\lambda$  is simply  $P_z(z_\lambda) = O(|z_\lambda - z|^3) = O(\lambda^{-3})$ .

For second order term, we must compute  $\nabla^2 P_z(z_\lambda)$ . Now  $\nabla^2 P_z$  vanishes through zeroth order at  $z$ , so  $\nabla^2 P_z(x) = O(|x - z|)$ , and in particular  $\nabla^2 P_z(z_\lambda) = O(\lambda^{-1})$ . The second order term of the Taylor series of  $P_z$  at  $z_\lambda$  is then  $\frac{1}{2}(x - z_\lambda)^T \nabla^2 P_z(z_\lambda)(x - z_\lambda)$ , which integrates against  $\mu_{z_\lambda, \lambda}$  to  $O(\lambda^{-2})$ .

Finally, we consider the fourth order term. Now the fourth derivatives of  $P_z$  are constant, and the fourth order term of the Taylor series of  $P_z$  at  $z_\lambda$  is a linear combination (with coefficients constant in  $\lambda$ ) of terms of the form

$$(x_i - z_{\lambda, i})(x_j - z_{\lambda, j})(x_k - z_{\lambda, k})(x_l - z_{\lambda, l}),$$

which integrate against  $\mu_{z_\lambda, \lambda}$  to  $O(\lambda^{-2})$ .

Thus we have established that

$$\int_{\mathcal{N}_z} (f - Q_z) d\mu_{z_\lambda, \lambda} = O(\lambda^{-2}),$$

for all  $z \in g\{-1, 1\}^n$ , as desired.

## II.6. Controlling the relative entropy term.

The preceding result implies that

$$\overline{\mathcal{F}}_\lambda[G | g] + \frac{n}{2} \log \left( \frac{\lambda}{2\pi} \right) \geq S_{\nu_\lambda}(\mu_\lambda) + n \log 2 - \frac{1}{2} \text{Tr} \log v - \frac{1}{2} \sum_{i=1}^n \log G_{ii} + O(\lambda^{-1}).$$

For the rest of the lower bound, as discussed above, we only need to upper-bound  $S_{\nu_\lambda}(\mu_\lambda)$  by  $S_w(p) = \widehat{\mathcal{F}}[G | g] - n \log 2$  plus some vanishing error term. (Unfortunately, the joint upper semi-continuity of the entropy with respect to the topology of weak convergence gives us an inequality in the wrong direction, but in any case we can be more quantitative.)

Recall:  $\mu_\lambda = \sum_{z \in g\{-1,1\}^n} p_{z_\lambda, \lambda} \mu_{z_\lambda, \lambda}$  and  $\nu = \frac{1}{2^n} \sum_{z \in g\{-1,1\}^n} \mu_{z_\lambda, \lambda}$ , where  $\mu_{z_\lambda, \lambda} \sim \mathcal{N}(z_\lambda, \lambda^{-1} H_{z_\lambda, \lambda}^{-1})$ . Then

$$S_{\nu_\lambda}(\mu_\lambda) = - \int \log \left[ \frac{\sum_{z \in g\{-1,1\}^n} p_{z_\lambda, \lambda} \exp(-\lambda Q_{z_\lambda, \lambda})}{\sum_{z \in g\{-1,1\}^n} \frac{1}{2^n} \exp(-\lambda Q_{z_\lambda, \lambda})} \right] d\mu_\lambda.$$

Notice that there exists  $C > 0$  such that  $C^{-1} \leq p_{z_\lambda, \lambda} 2^n \leq C$ , hence

$$\begin{aligned} C^{-1} \sum_{z \in g\{-1,1\}^n} \frac{1}{2^n} \exp(-\lambda Q_z) &\leq \sum_{z \in g\{-1,1\}^n} p_{z_\lambda, \lambda} \exp(-\lambda Q_z) \\ &\leq C \sum_{z \in g\{-1,1\}^n} \frac{1}{2^n} \exp(-\lambda Q_z), \end{aligned}$$

i.e.,

$$C^{-1} \leq \frac{\sum_{z \in g\{-1,1\}^n} p_{z_\lambda, \lambda} \exp(-\lambda Q_{z_\lambda, \lambda})}{\sum_{z \in g\{-1,1\}^n} \frac{1}{2^n} \exp(-\lambda Q_{z_\lambda, \lambda})} \leq C.$$

Then

$$\begin{aligned} S_{\nu_\lambda}(\mu_\lambda) &= - \sum_{z'} \int_{\mathcal{N}_{z'}} \log \left[ \frac{\sum_{z \in g\{-1,1\}^n} p_{z_\lambda, \lambda} \exp(-\lambda Q_{z_\lambda, \lambda})}{\sum_{z \in g\{-1,1\}^n} \frac{1}{2^n} \exp(-\lambda Q_{z_\lambda, \lambda})} \right] d\mu_\lambda + O(e^{-C^{-1}\lambda}) \\ &\leq - \sum_{z'} \int_{\mathcal{N}_{z'}} \log \left[ \frac{p_{z'_\lambda, \lambda} \exp(-\lambda Q_{z'_\lambda, \lambda})}{\sum_{z \in g\{-1,1\}^n} \frac{1}{2^n} \exp(-\lambda Q_{z_\lambda, \lambda})} \right] d\mu_\lambda + O(e^{-C^{-1}\lambda}). \end{aligned}$$

Furthermore, as proved earlier,

$$\log \sum_{z \in g\{-1,1\}^n} \exp(-\lambda Q_{z_\lambda, \lambda}) \leq -\lambda Q_{z'_\lambda, \lambda}(x) + C e^{-\lambda \Delta} = \log \exp(-\lambda Q_{z'_\lambda, \lambda}(x) + C e^{-\lambda \Delta})$$

on  $\mathcal{N}_{z'}$ , so

$$\begin{aligned}
S_{\nu_\lambda}(\mu_\lambda) &\leq -\sum_{z'} \int_{\mathcal{N}_{z'}} \log \left[ \frac{p_{z',\lambda} \exp(-\lambda Q_{z',\lambda})}{\frac{1}{2^n} \exp(-\lambda Q_{z',\lambda}(x) + C e^{-\lambda \Delta})} \right] d\mu_\lambda + O(e^{-C^{-1}\lambda}) \\
&= -\sum_z \int_{\mathcal{N}_z} \log \left[ \frac{p_{z,\lambda}}{2^{-n}} \right] d\mu_\lambda + O(e^{-C^{-1}\lambda}) \\
&= -\sum_z \log \left[ \frac{p_{z,\lambda}}{2^{-n}} \right] \mu_\lambda(\mathcal{N}_z) + O(e^{-C^{-1}\lambda}).
\end{aligned}$$

But  $\mu_\lambda(\mathcal{N}_z) = p_z + O(e^{-C^{-1}\lambda})$ , and  $p_{z,\lambda} = p_z + O(\lambda^{-1})$ , so we have

$$S_{\nu_\lambda}(\mu_\lambda) \leq S_w(p) + O(\lambda^{-1}).$$

In summary, we have shown

$$\bar{\mathcal{F}}_\lambda[G|g] + \frac{n}{2} \log \left( \frac{\lambda}{2\pi} \right) \geq \hat{\mathcal{F}}[G|g] - \frac{1}{2} \text{Tr} \log v - \frac{1}{2} \sum_{i=1}^n \log G_{ii} + O(\lambda^{-1}).$$

as desired.

This completes the proof.  $\square$

## 6 Gradient asymptotics

We want to understand not only the asymptotic behavior of  $\mathcal{F}_\lambda[G]$  as  $\lambda \rightarrow \infty$ , but also the behavior of  $A_\lambda[G]$  as  $\lambda \rightarrow \infty$ .

In order to separate the behavior on the diagonal from the behavior on the off-diagonal, we define an operator  $\mathcal{Z}$  which replaces the diagonal entries of its argument with zeros and an operator  $\mathcal{D}$  which replaces the off-diagonal entries with zeros. Thus  $\mathcal{D} + \mathcal{Z} = \text{Id}$ .

Recall that  $A_\lambda[G] = \nabla \mathcal{F}_\lambda[G]$ . Now the main theorem of the last section implies in particular that

$$\frac{1}{\lambda} \mathcal{F}_\lambda[G] \rightarrow -\frac{1}{8} \sum_{ij} v_{ij} G_{ii} G_{jj}$$

pointwise for  $G \in \text{int}(K_n)$ . The pointwise convergence of concave functions (cf. Theorem 23 of Appendix C) implies the pointwise convergence of their supergradients (in this case, ordinary gradients). Therefore

$$\frac{1}{\lambda} A_\lambda[G] \rightarrow -D[G]$$

for all  $G \in \text{int}(K_n)$ , where  $D[G]$  is defined as before by

$$D[G] = \frac{1}{2} \text{diag} \left[ \sum_j v_{1j} G_{jj}, \dots, \sum_j v_{nj} G_{jj} \right].$$

We also saw that

$$\mathcal{F}_\lambda[G] + \frac{n}{2} \log \left( \frac{\lambda}{2\pi} \right) + \frac{\lambda}{8} \sum_{ij} v_{ij} G_{ii} G_{jj} \rightarrow \widehat{\mathcal{F}}[G|g] - \frac{1}{2} \text{Tr} \log v - \frac{1}{2} \sum_{i=1}^n \log G_{ii}$$

pointwise, though this is no longer the limit of convex functions on  $\mathcal{S}_{++}^n$ . However, this is a limit of convex functions on matrices with fixed diagonal  $G_{ii} = G_{ii}^*$ , so taking gradients in the off-diagonal directions implies that

$$L_\lambda[G] := \mathcal{Z}(A_\lambda[G]) \rightarrow B[G].$$

Next, it would be desirable to show that

$$A_\lambda[G] + \lambda D[G]$$

converges, or at least first that this quantity bounded. What we don't yet know is whether this holds on the diagonal, i.e., whether the 'on-diagonal remainder'

$$M_\lambda[G] := \mathcal{D}(A_\lambda[G]) + \lambda D[G]$$

is bounded. This can be established by combining concavity with our quantitative error bounds from the last section.

**Proposition 6.** *For  $G \in \text{int}(K_n)$ ,  $\mathcal{D}(A_\lambda[G]) = -\lambda D[G] + O(1)$ , i.e.,  $M_\lambda[G] = O(1)$ .*

*Proof.* This is the same as saying that  $M_\lambda[G] = O(1)$ . For  $G \in \mathcal{S}_{++}^n$ , define

$$\Gamma_\lambda[G] := \frac{1}{\lambda} \left[ \mathcal{F}_\lambda[G] + \frac{n}{2} \log \left( \frac{\lambda}{2\pi} \right) + \frac{1}{2} \text{Tr} \log v \right],$$

so  $\Gamma_\lambda$  is concave on  $\mathcal{S}_{++}^n$ , and  $\nabla \Gamma_\lambda = \frac{1}{\lambda} A_\lambda$ . Then for  $G \in \text{int}(K_n)$ , by Theorem 3 we have

$$\Gamma_\lambda[G] = \Psi_\lambda[G] + O(\lambda^{-2}),$$

where

$$\Psi_\lambda[G] := -\frac{1}{8} \sum_{ij} v_{ij} G_{ii} G_{jj} + \lambda^{-1} \widehat{\mathcal{F}}[G|g] - \frac{1}{2} \lambda^{-1} \sum_{i=1}^n \log G_{ii}$$

and the constant in  $O(\lambda^{-2})$  can be taken to be locally uniform in  $G$ . Let us make the dependence of  $g = g(G) = \text{diag} \sqrt{\text{diag}(G)}$  on  $G$  explicit and use the relation of  $\widehat{\mathcal{F}}[G|g]$  to the reference spin system via Proposition 1 to write

$$\widehat{\mathcal{F}}[G|g] = \widehat{\mathcal{F}}[g(G)^{-1} G g(G)^{-1}].$$

By the smoothness of  $\widehat{\mathcal{F}}$ , it is then clear that  $\Psi_\lambda[G]$  has derivatives that are locally uniformly bounded on  $\text{int}(K_n)$ , with bounds independent of  $\lambda$  (sufficiently large).

Now fix  $G \in \text{int}(K_n)$ , and consider  $T \in \mathcal{S}^n$  and  $\delta$  sufficiently small so that in particular  $G + \delta T \in \text{int}(K_n)$ , we have by concavity that

$$\Gamma_\lambda[G] + \delta \text{Tr}(\lambda^{-1} A_\lambda[G]T) \geq \Gamma_\lambda[G + \delta T].$$

Therefore

$$\begin{aligned} \text{Tr}(\lambda^{-1} A_\lambda[G]T) &\geq \frac{\Gamma_\lambda[G + \delta T] - \Gamma_\lambda[G]}{\delta} \\ &\geq \frac{\Psi_\lambda[G + \delta T] - \Psi_\lambda[G]}{\delta} - C\delta^{-1}\lambda^{-2} \\ &\geq \text{Tr}(\nabla\Psi_\lambda[G]T) - C\delta - C\delta^{-1}\lambda^{-2}, \end{aligned}$$

where we have used local  $\lambda$ -independent uniform bounds on the second derivative of  $\Psi_\lambda$  to bound the error of the difference quotient in the last step.

Then choose  $\delta = \delta(\lambda) = \lambda^{-1}$  to obtain

$$\text{Tr}(\lambda^{-1} A_\lambda[G]T) \geq \text{Tr}(\nabla\Psi_\lambda[G]T) - C\lambda^{-1}.$$

Notice, however, that (since  $G$  is fixed),  $\nabla\Psi_\lambda[G] = -D[G] + O(\lambda^{-1})$ , so we have shown that

$$\text{Tr}[T(\lambda^{-1} A_\lambda[G] + D[G])] \geq O(\lambda^{-1}),$$

i.e.,

$$\text{Tr}[T(A_\lambda[G] + \lambda D[G])] \geq O(1),$$

for any fixed  $T \in \mathcal{S}^n$ . Simply choosing  $T = \pm e_i e_i^T$  yields the result.  $\square$

## 7 Polarization relation

Let  $G \in \mathcal{S}_{++}^n$ , and let  $\langle \cdot \rangle_{\lambda, G}$  denote the expectation with respect to the Gibbs measure  $\mu_{\lambda, G}$  with density

$$d\mu_\lambda = (Z_\lambda[A_\lambda])^{-1} e^{-\frac{1}{2}x^T A_\lambda[G]x - \lambda U(x)} dx.$$

In particular,  $\langle x_i x_j \rangle_{\lambda, G} = G_{ij}$ .

In the following, when considering a fixed  $G \in \mathcal{S}_{++}^n$  we sometimes omit the dependence on  $G$  from the notation, denoting, e.g.,  $A_\lambda = A_\lambda[G]$ ,  $\langle \cdot \rangle_\lambda = \langle \cdot \rangle_{\lambda, G}$ , etc., when the context is clear.

Now define the *polarization*  $P_\lambda = P_\lambda[G]$  by

$$P_{\lambda, ij} = \langle (x_i^2 - \langle x_i^2 \rangle_\lambda)(x_j^2 - \langle x_j^2 \rangle_\lambda) \rangle_\lambda = \langle x_i^2 x_j^2 \rangle_\lambda - G_{ii} G_{jj}.$$

If  $X_\lambda$  is the random vector with distribution  $\mu_\lambda$ , then  $P_\lambda$  is the covariance matrix of  $X_\lambda$ . Evidently  $P_\lambda \in \mathcal{S}_{++}^n$ .

**Proposition 7.** For all  $G \in \mathcal{S}_{++}^n$ ,  $\lambda > 0$ , and  $i, j = 1, \dots, n$ ,

$$\begin{aligned} (A_\lambda[G]G)_{ij} &= \delta_{ij} - \frac{1}{2}\lambda \sum_{k=1}^n v_{ik} \langle x_i x_j x_k^2 \rangle_{\lambda, G} \\ &= \delta_{ij} - \lambda (D[G]G)_{ij} - \frac{1}{2}\lambda \sum_{k=1}^n v_{ik} \langle x_i x_j (x_k^2 - G_{kk}) \rangle_{\lambda, G}. \end{aligned}$$

Consequently,

$$\begin{aligned} (A_\lambda[G]G)_{ii} &= 1 - \frac{1}{2}\lambda \sum_{j=1}^n v_{ik} G_{ii} G_{kk} - \frac{1}{2}\lambda (P_\lambda[G]v)_{ii} \\ &= 1 - \lambda (D[G]G)_{ii} - \frac{1}{2}\lambda (P_\lambda[G]v)_{ii}, \end{aligned}$$

so

$$M_{\lambda, ii}[G]G_{ii} = 1 - [L_\lambda[G]G]_{ii} - \frac{1}{2}\lambda (P_\lambda[G]v)_{ii},$$

and

$$\text{Tr} \left[ \left( A_\lambda[G] + \frac{1}{2}\lambda D[G] - I \right) G \right] = \frac{1}{2}\lambda \text{Tr} (P_\lambda[G]v).$$

*Proof.* Fix  $G$  and write  $A_\lambda = A_\lambda[G]$ , etc. Let

$$J_\lambda(x) = \exp \left[ -\frac{1}{2} \sum_{kl} \left( A_{\lambda, kl} x_k x_l + \frac{1}{4} \lambda v_{ij} x_k^2 x_l^2 \right) \right],$$

so

$$\frac{\partial J_\lambda}{\partial x_i} = - \sum_k \left( A_{\lambda, ik} x_k + \frac{1}{2} \lambda v_{ik} x_k^2 x_i \right) J(x).$$

Then we can integrate by parts to obtain

$$\begin{aligned} 1 &= \frac{1}{Z_\lambda[A_\lambda]} \int 1 \cdot J_\lambda(x) dx \\ &= \frac{1}{Z_\lambda[A_\lambda]} \int x_k \sum_i \left( A_{\lambda, ik} x_i + \frac{1}{2} \lambda v_{ik} x_i^2 x_k \right) J(x) dx \\ &= \sum_i A_{\lambda, ik} \frac{1}{Z_\lambda[A_\lambda]} \int x_k x_i J(x) dx + \sum_i \frac{1}{2} \lambda v_{ik} \frac{1}{Z_\lambda[A_\lambda]} \int x_i^2 x_k^2 J(x) dx \\ &= \sum_i A_{\lambda, ik} G_{ki} + \frac{1}{2} \lambda \sum_i v_{ik} \langle x_i^2 x_k^2 \rangle_\lambda. \end{aligned}$$

This establishes the first statement in the case  $i = j$ .

For the case  $i \neq j$ , we apply integration by parts again:

$$\begin{aligned}
\delta_{ij} &= \frac{1}{Z_\lambda[A_\lambda]} \int \delta_{ij} \cdot J_\lambda(x) dx \\
&= \frac{1}{Z_\lambda[A_\lambda]} \int \frac{\partial x_j}{\partial x_i} \cdot J_\lambda(x) dx \\
&= \frac{1}{Z_\lambda[A_\lambda]} \int x_j \sum_k \left( A_{\lambda,ik} x_k + \frac{1}{2} \lambda v_{ik} x_k^2 x_i \right) J(x) dx \\
&= \sum_k A_{\lambda,ik} \frac{1}{Z_\lambda[A_\lambda]} \int x_k x_j J(x) dx + \sum_k \frac{1}{2} \lambda v_{ik} \frac{1}{Z_\lambda[A_\lambda]} \int x_i x_j x_k^2 J(x) dx \\
&= \sum_k A_{\lambda,ik} G_{kj} + \frac{1}{2} \lambda \sum_i v_{ik} \langle x_i x_j x_k^2 \rangle_\lambda.
\end{aligned}$$

This establishes the first statement. The others follow by considering the case  $i = j$  and rearranging terms.  $\square$

We have shown that  $M_\lambda = O(1)$ , but we have not identified a limit for  $M_\lambda$  (nor shown that one exists). By contrast, we do have a formula for the limit of the off-diagonal part of  $A_\lambda + \lambda D[G]$ , namely  $B = B[G]$ .

The formula

$$M_{\lambda,ii} G_{ii} = 1 - [L_\lambda G]_{ii} - \frac{1}{2} \lambda (P_\lambda v)_{ii}$$

suggests a way to remedy this. Note that  $[L_\lambda G]_{ii} \rightarrow (B[G]G)_{ii}$ . In fact, we expect  $\lambda P_\lambda$  to have a limit as well. If we can compute this limit, then we can simply read off a limiting formula for  $M_\lambda$ .

To see why  $\lambda P_\lambda$  should have a limit, write  $P_{\lambda,ij} = \langle f_{ij} \rangle_\lambda$ , where

$$f_{ij}(x) = (x_i^2 - G_{ii})(x_j^2 - G_{jj}).$$

Now  $f_{ij}$  has critical points at  $z \in g\{-1, 1\}^n$ , hence is locally quadratic near  $g\{-1, 1\}^n$ . We expect  $\mu_\lambda$  to look Gaussian near  $g\{-1, 1\}^n$  with covariance  $\sim \frac{1}{\lambda}$ . Thus  $\langle f_{ii} \rangle_\lambda$  should be  $O(\lambda^{-1})$ , and  $\langle f_{ij} \rangle_\lambda$  should be even smaller. Thus as claimed, we expect  $\frac{1}{2} \lambda (P_\lambda v)_{ii}$  to have a limit.

## 8 Asymptotic polarization and the limit of $M_\lambda[G]$

We now carry out the strategy outlined at the end of the last section.

**Proposition 8.** *Let  $G \in \text{int}(K_n)$ . Then  $\lambda P_\lambda[G] \rightarrow 4v^{-1}$  as  $\lambda \rightarrow \infty$ . Consequently  $M_{\lambda,ii}[G] \rightarrow -(G_{ii})^{-1} [1 + (B[G]G)_{ii}]$ , i.e.,*

$$M_\lambda[G] \rightarrow -g^{-1} [I + \mathcal{D}(B[G]G)] g^{-1}$$

as  $\lambda \rightarrow \infty$ .

**Corollary 9.** *Let  $G \in \text{int}(K_n)$ . Then  $M_{\lambda,ii}[G] \rightarrow -(G_{ii})^{-1} [1 + (B[G]G)_{ii}]$ , i.e.,*

$$M_\lambda[G] \rightarrow -g^{-1} [I + \mathcal{D}(B[G]G)] g^{-1}$$

as  $\lambda \rightarrow \infty$ . Consequently

$$A_\lambda[G] = -\lambda D[G] + B[G] + g^{-1} [I + \mathcal{D}(B[G]G)] g^{-1} + o(1).$$

*Proof.* Throughout we suppose  $G \in \text{int}(K_n)$  is fixed. Before proceeding, recall that we know that

$$A_\lambda = -\lambda D[G] + L_\lambda + M_\lambda,$$

where  $L_\lambda \rightarrow B = B[G]$  has zero diagonal and  $M_\lambda = O(1)$ . Then there exists  $C$  such that  $-CI_n \preceq M_\lambda \preceq CI_n$ . Assume  $C$  is large enough so that  $B_{\frac{1}{2}\sqrt{C}}(0)$  contains  $g\{-1, 1\}^n$ .

Throughout this discussion we will use  $C$  to denote constants independent of  $\lambda$  that are sufficiently large in the given context; the precise meaning of  $C$  may change across usages.

Let  $\delta > 0$ , and assume that  $\lambda$  is large enough so that  $-\delta I_n \preceq L_\lambda - B \preceq \delta I_n$ . Also, let  $\mathcal{N} \subset B_{\sqrt{C}}(0)$  be a small enough neighborhood of  $g\{-1, 1\}^n$  such that

$$\left| \frac{1}{2} g^T M_\lambda g - \frac{1}{2} x^T M_\lambda x \right| \leq \delta$$

on  $\mathcal{N}$  for all  $\lambda$ . (This is possible because  $M_\lambda$  is diagonal!)

Let

$$f(x) = f_g(x) = \sum_{ij} v_{ij} (x_i^2 - G_{ii})(x_j^2 - G_{jj}),$$

and recall

$$f(x) = \frac{1}{2} D[G] - U(x) - \frac{1}{8} \sum_{ij} v_{ij} G_{ii} G_{jj}.$$

Finally, define  $Z_\lambda^H := \exp\left(\frac{\lambda}{8} \sum_{ij} v_{ij} G_{ii} G_{jj}\right)$  to ease the notation later on. ('H' is for 'Hartree'.)

Then compute

$$\begin{aligned} Z_\lambda[A_\lambda] &= \int e^{-\frac{1}{2} x^T A_\lambda x - \lambda U(x)} dx \\ &= Z_\lambda^H \int e^{-\frac{1}{2} x^T (L_\lambda + M_\lambda) x} e^{-\lambda f(x)} dx \\ &= Z_\lambda^H e^{-\frac{1}{2} g^T M_\lambda g} \int e^{-\frac{1}{2} x^T L_\lambda x} e^{\frac{1}{2} g^T M_\lambda g - \frac{1}{2} x^T M_\lambda x} e^{-\lambda f(x)} dx \\ &= Z_\lambda^H e^{-\frac{1}{2} g^T M_\lambda g} \left[ \int_{\mathcal{N}} e^{-\frac{1}{2} x^T L_\lambda x} e^{\frac{1}{2} g^T M_\lambda g - \frac{1}{2} x^T M_\lambda x} e^{-\lambda f(x)} dx + O(e^{-C^{-1}\lambda}) \right] \end{aligned}$$



$$\begin{aligned}
&\leq Z_\lambda^H e^{-\frac{1}{2}g^T M_\lambda g} \left[ \int_{\mathcal{N}} e^{\delta|x|^2} e^{-\frac{1}{2}x^T Bx} e^{-\lambda f(x)} dx + O(e^{-C^{-1}\lambda}) \right] \\
&\leq Z_\lambda^H e^{-\frac{1}{2}g^T M_\lambda g} e^{C\delta} \left[ \int_{\mathcal{N}} e^{-\frac{1}{2}x^T Bx} e^{-\lambda f(x)} dx + O(e^{-C^{-1}\lambda}) \right] \\
&= Z_\lambda^H e^{-\frac{1}{2}g^T M_\lambda g} e^{C\delta} \\
&\quad \left( \frac{2\pi}{\lambda} \right)^{n/2} \frac{1}{\sqrt{\prod_{i=1}^n G_{ii}}} \frac{1}{\sqrt{\det v}} \sum_{z \in g\{-1,1\}^n} e^{-\frac{1}{2}z^T Bz} [1 + O(\lambda^{-1})],
\end{aligned}$$

where in the last step we have used Laplace's method. Meanwhile,

$$\begin{aligned}
\langle f_{ij} \rangle_\lambda &= Z_\lambda^H e^{-\frac{1}{2}g^T M_\lambda g} \int f_{ij}(x) e^{-\frac{1}{2}x^T L_\lambda x} e^{\frac{1}{2}g^T M_\lambda g - \frac{1}{2}x^T M_\lambda x} e^{-\lambda f(x)} dx \\
&\geq Z_\lambda^H e^{-\frac{1}{2}g^T M_\lambda g} \left[ \int_{\mathcal{N}} e^{-\delta|x|^2} f_{ij}(x) e^{-\frac{1}{2}x^T Bx} e^{-\lambda f(x)} dx + O(e^{-C^{-1}\lambda}) \right] \\
&\geq Z_\lambda^H e^{-\frac{1}{2}g^T M_\lambda g} e^{-C\delta} \left[ \int_{\mathcal{N}} e^{-\frac{1}{2}x^T Bx} f_{ij}(x) e^{-\lambda f(x)} dx + O(e^{-C^{-1}\lambda}) \right].
\end{aligned}$$

Now for  $z \in g\{-1,1\}^n$ , we have  $f_{ij}(z) = 0$ ,  $\nabla(f_{ij})(z) = 0$ , and  $\nabla^2(f_{ij})(z) = 4E_{ij} \odot (zz^T)$ , where  $(E_{ij})_{kl} = \delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk}$ . Then expanding  $e^{-\frac{1}{2}x^T Bx} f_{ij}(x)$  to second order near  $z$ , we have

$$e^{-\frac{1}{2}x^T Bx} f_{ij}(x) = e^{-\frac{1}{2}z^T Bz} \frac{1}{2}(x-z)^T [E_{ij} \odot (zz^T)] (x-z) + \dots$$

Note that  $e^{-\lambda f(x)}$ , normalized by the appropriate factor, looks near  $z$  like the density of the Gaussian distribution centered at  $z$  with covariance  $\lambda^{-1} [v \odot (zz^T)]^{-1}$ .

Omitting some details, the technique of Laplace's method then yields

$$\begin{aligned}
&\int_{\mathcal{N}} e^{-\frac{1}{2}x^T Bx} f_{ij}(x) e^{-\lambda f(x)} dx \\
&\sim \left( \frac{2\pi}{\lambda} \right)^{n/2} \frac{1}{\sqrt{\prod_{i=1}^n G_{ii}}} \frac{1}{\sqrt{\det v}} \\
&\quad \sum_{z \in \{-1,1\}^n} e^{-\frac{1}{2}z^T Bz} 2\lambda^{-1} \text{Tr} \left( [E_{ij} \odot (zz^T)] [v \odot (zz^T)]^{-1} \right).
\end{aligned}$$

But, writing  $\zeta = \text{diag}(z)$ ,

$$[E_{ij} \odot (zz^T)] [v \odot (zz^T)]^{-1} = [\zeta E_{ij} \zeta] [\zeta v \zeta]^{-1} = \zeta E_{ij} v^{-1} \zeta^{-1},$$

so

$$\text{Tr} \left( [E_{ij} \odot (zz^T)] [v \odot (zz^T)]^{-1} \right) = \text{Tr} [E_{ij} v^{-1}] = 2v^{ij}.$$

Therefore

$$\int_{\mathcal{N}} e^{-\frac{1}{2}x^T Bx} f_{ij}(x) e^{-\lambda f(x)} dx \sim 4\lambda^{-1} \left(\frac{2\pi}{\lambda}\right)^{n/2} \frac{1}{\sqrt{\prod_{i=1}^n G_{ii}}} \frac{1}{\sqrt{\det v}} \sum_{z \in \{-1,1\}^n} e^{-\frac{1}{2}z^T Bz} v^{ij}.$$

Combining our results then reveals

$$\lambda P_{\lambda,ij} = \frac{\langle \lambda f_{ij} \rangle_{\lambda}}{Z_{\lambda}[A_{\lambda}]} \geq e^{-2C\delta} 4v^{ij} + o(1).$$

Since  $\delta > 0$  was arbitrary, we have shown

$$\lambda P_{\lambda,ij} \geq 4v^{ij} + o(1).$$

Similar reasoning guarantees the reverse inequality, so  $\lambda P_{\lambda} \rightarrow 4v^{-1}$ , as was to be shown.  $\square$

If we define the *screened Coulomb interaction* as  $W_{\lambda}[G] := v - \frac{1}{4}vP_{\lambda}[G]v$ , then we see that  $\lambda^{-1}W_{\lambda} \rightarrow -v$ , so the screened Coulomb interaction is  $o(\lambda)$ .

## Part V

# The 1-RDM theory and fermionic embedding

## 1 Introduction

The fermionic 1-RDM theory (one-body reduced density matrix) is an analog of sorts to the classical Luttinger-Ward (LW) theory as presented in Part III. In the LW theory, we started with the concavity of the free energy functional, considered as a function of the quadratic part of the Hamiltonian (with the rest of the Hamiltonian, i.e., the interaction, fixed). The quadratic part on its own yields a Gaussian measure, which can be thought of as a completely solved model.

Meanwhile, in the fermionic 1-RDM theory, we start with the concavity of the quantum (fermionic) free energy functional, considered as a function of the quadratic part of the Hamiltonian, where the interaction is again fixed. The quadratic part on its own yields a free fermionic ensemble, which can also be thought of as a completely solved model.

Proceeding by analogy to Part III, we can define a suitable concave duality and examine the behavior of the dual functional (namely, the 1-RDM functional) on the boundary of its effective domain. Unfortunately, the 1-RDM functional does not enjoy a diagrammatic expansion, but its boundary behavior can, like that of the LW functional, still be understood in terms of the 1-RDM functional of a smaller system with an effective interaction, which we derive explicitly in the case of the two-body interaction. To prove this result, we first prove a structural result that constrains the form of fermionic ensembles whose 1-RDMs are on the boundary of the domain of the 1-RDM functional. In addition to helping us understand the behavior of the 1-RDM functional on the boundary, this result offers some perspective on fermionic embedding methods such as the density matrix embedding theory (DMET) [49].

## 2 The fermionic 1-RDM theory

Our setting is a fermionic Fock space denoted  $\mathcal{F}_d$  with a finite number  $d$  of states. (Refer to section 4 of Part I for background on second quantization.) The reader may consult [38] for a related presentation of the 1-RDM theory.

Now consider a Hamiltonian  $\hat{H}$ , i.e., a Hermitian operator  $\mathcal{F}_d \rightarrow \mathcal{F}_d$ . In general we can write

$$\hat{H} = \hat{H}_0 + \hat{V},$$

where

$$\hat{H}_0 = \sum_{i,j=1}^d A_{ij} c_i^\dagger c_j,$$

is the single-body (or noninteracting) part of the Hamiltonian, specified by a  $d \times d$  Hermitian matrix  $A = (A_{ij})$ , and  $\hat{V}$  is the *interacting* part, also a Hermitian operator. Though we need not define  $\hat{V}$  more explicitly for now, one might keep in mind the two-body interaction

$$\hat{V} = \frac{1}{2} \sum_{i,j,k,l=1}^d (ij|kl) c_i^\dagger c_k^\dagger c_l c_j.$$

In quantum statistical mechanics at inverse temperature  $\beta \in (0, \infty)$ , we are concerned with the *partition function*

$$Z := \text{Tr} \left[ e^{-\beta(\hat{H} - \mu \hat{N})} \right],$$

where  $\mu \in \mathbb{R}$  is the chemical potential and ‘Tr’ denotes the trace operation for operators on the Fock space. The *density operator*  $\rho := \frac{1}{Z} e^{-\beta(\hat{H} - \mu \hat{N})}$  defines the so-called *grand canonical ensemble* by specifying how to compute ensemble averages of operators. To wit, for an operator  $\mathcal{O} : \mathcal{F}_d \rightarrow \mathcal{F}_d$ , the ensemble average of  $\mathcal{O}$  is denoted by

$$\langle \mathcal{O} \rangle = \frac{1}{Z} \text{Tr} [\mathcal{O} \rho],$$

For now we will not think of  $\mu$  as variable, we can assume without loss of generality (by appropriately redefining  $A \leftarrow A - \mu I_d$ ) that  $\mu = 0$ , so

$$Z = \text{Tr} \left[ e^{-\beta \hat{H}} \right], \quad \rho = \frac{1}{Z} e^{-\beta \hat{H}}$$

In fact, we will think of  $H$ ,  $Z$ , and  $\rho$  as *functions* of  $A$ , writing

$$\hat{H}[A] = \sum_{i,j} A_{ij} c_i^\dagger c_j + \hat{V}, \quad Z[A] = \text{Tr} \left[ e^{-\hat{H}[A]} \right], \quad \rho[A] = \frac{1}{Z[A]} e^{-\beta \hat{H}[A]}$$

and further defining the *free energy* by

$$\Omega[A] := -\log Z[A].$$

It is sometimes convenient to let  $c = (c_1, \dots, c_d)^\top$  denote the vector of annihilation operators, so  $\hat{H}[A] = c^\dagger A c + \hat{V}$ .

One can also think of all of these quantities as depending on  $\hat{V}$ , and we can accordingly write, e.g.  $\Omega[A; \hat{V}]$ . However, when the context is clear we omit the dependence on  $\hat{V}$  from the notation.

For clarity, we now establish our conventions for taking derivatives on the set of Hermitian matrices.

**Definition 1.** For  $i, j = 1, \dots, d$ , let  $E^{(ij)} \in \mathbf{H}^d$  be defined by  $E_{kl}^{(ij)} = \frac{1}{2}(1+i)\delta_{ik}\delta_{jl} + \frac{1}{2}(1-i)\delta_{il}\delta_{jk}$ . For a differentiable function  $f : \mathbf{H}^d \rightarrow \mathbb{R}$ , define the gradient  $\nabla f : \mathbf{H}^d \rightarrow \mathbf{H}^d$  by

$$(\nabla f(A))_{ij} := \nabla_{ij} f(A) := \lim_{\delta \rightarrow 0} \frac{f(A + \delta E^{(ij)}) - f(A)}{\delta}.$$

If  $f$  is obtained by restriction from a function  $f : \mathbb{C}^{M \times M} \rightarrow \mathbb{R}$ , then equivalently

$$\nabla_{ij} f = \frac{1}{2} \frac{\partial f}{\partial A_{ij}} + \frac{1}{2} \frac{\partial f}{\partial A_{ji}}.$$

Then observe that the gradient map  $\nabla \Omega$  is given by

$$\nabla_{ij} \Omega[A] = \frac{1}{Z[A]} \text{Tr} \left[ \frac{1}{2} \left( c_i^\dagger c_j + c_j^\dagger c_i \right) e^{-\hat{H}[A]} \right].$$

Notice that

$$\begin{aligned} \text{Tr} \left[ c_i^\dagger c_j e^{-\hat{H}[A]} \right] &= \text{Tr} \left[ c_j e^{-\hat{H}[A]} c_i^\dagger \right] \\ &= \text{Tr} \left[ \left( c_j e^{-\hat{H}[A]} c_i^\dagger \right)^\dagger \right] = \text{Tr} \left[ c_i e^{-\hat{H}[A]} c_j^\dagger \right] = \text{Tr} \left[ c_j^\dagger c_i e^{-\hat{H}[A]} \right], \end{aligned}$$

so in fact

$$\nabla_{ij} \Omega[A] = \text{Tr} \left[ c_i^\dagger c_j \rho[A] \right].$$

Then  $D[A] := \nabla \Omega[A]$  is the *1-particle density matrix (1-RDM)*. This map will define a correspondence between single-particle Hamiltonians and 1-RDMs via concave duality, suggested by the following lemma.

**Lemma 2.**  $\Omega$  is strictly concave on  $\mathbf{H}^d$ .

The proof follows from the following more general fact:

**Lemma 3.** Let  $M$  be a positive integer. Then the function  $f_M : \mathbf{H}^M \rightarrow \mathbb{R}$  defined by  $f(X) = \log \text{Tr} [e^X]$  is strictly convex.

*Proof.* One can prove this fact by taking derivatives, but we present much quicker proof that makes use of nice inequalities. Indeed, let  $X_1, X_2 \in \mathbf{H}^M$ , and let  $\theta \in [0, 1]$ . Then

$$f(\theta X_1 + (1-\theta)X_2) = \log \text{Tr} [e^{\theta X_1 + (1-\theta)X_2}] \leq \log \text{Tr} [e^{\theta X_1} e^{(1-\theta)X_2}]$$

by the Golden-Thompson inequality. Now Hölder's inequality for the Schatten norms implies that

$$\text{Tr} [e^{\theta X_1} e^{(1-\theta)X_2}] \leq \|e^{\theta X_1}\|_{\theta^{-1}} \|e^{(1-\theta)X_2}\|_{(1-\theta)^{-1}} = (\text{Tr} [e^{X_1}])^\theta (\text{Tr} [e^{X_2}])^{1-\theta}.$$

Therefore

$$f(\theta X_1 + (1 - \theta)X_2) \leq \theta f(X_1) + (1 - \theta)f(X_2),$$

which proves convexity. In fact Hölder's inequality is strict if  $X_1 \neq X_2$ , so we have strict convexity.  $\square$

*Proof.* (Of Lemma 2.) We will apply Lemma 3 in the case  $M = 2^d$ . Observe that

$$\Omega[A] = -f_{2^d} \left( - \sum_{ij} A_{ij} c_i^\dagger c_j - \hat{V} \right).$$

Since the composition of a strictly convex function with a nondegenerate affine transformation is strictly convex, it follows that  $\Omega$  is strictly concave.  $\square$

Now we turn to finding the concave conjugate (Legendre transform) of  $\Omega$ . The answer comes to us via a contraction of the quantum Gibbs variational principle (Lemma 9 below), analogous to the classical Gibbs variational principle from classical statistical mechanics. In the quantum setting, the role of probability measures is played by density operators:

**Definition 4.** A *density operator* on  $\mathcal{F}_d$  (or for short, density operator) is a Hermitian positive semidefinite operator  $\mathcal{F}_d \rightarrow \mathcal{F}_d$  of unit trace. The set of density operators on  $\mathcal{F}_d$  is denoted  $\mathcal{D}_d$ .

*Remark 5.* Note that  $\rho[A]$  as defined above is an example of a density operator.

Moreover, the roles of the classical entropy and relative entropy are played by the von Neumann entropy and quantum relative entropy, defined as follows.

**Definition 6.** For  $\rho, \sigma \in \mathcal{D}_d$ , let

$$S(\rho) := \text{Tr}[\rho \log \rho]$$

denote the von Neumann entropy of  $\rho$ , and let

$$S(\rho||\sigma) := \text{Tr}[\rho \log \rho] - \text{Tr}[\rho \log \sigma]$$

denote the quantum relative entropy of  $\rho$  with respect to  $\sigma$ .

We collect some facts [78] about the von Neumann entropy and the relative entropy:

**Fact 7.** *The von Neumann entropy is  $\rho \mapsto S(\rho)$  is a concave function on  $\mathcal{D}_d$ , and  $0 \leq S(\rho) \leq \log(2^d)$  for all  $\rho \in \mathcal{D}_d$ .*

**Fact 8.** *For any  $\rho, \sigma \in \mathcal{D}_d$ ,  $S(\rho||\sigma) \geq 0$  with equality if and only if  $\rho = \sigma$ .*

Now our first step to identifying the concave conjugate of  $\Omega$  is the following lemma:

**Lemma 9.** For  $A \in \mathbf{H}^d$ ,

$$\Omega[A] = \inf_{\rho \in \mathcal{D}_d} \left[ \text{Tr}(\rho \hat{H}[A]) - S(\rho) \right].$$

Moreover, the infimum is uniquely attained at  $\rho[A] = \frac{1}{Z[A]} e^{-\hat{H}[A]}$ .

*Proof.* For  $\rho \in \mathcal{D}_d$ ,

$$\text{Tr}(\rho \hat{H}[A]) - S(\rho) = -\log Z[A] - \text{Tr}(\rho \log \rho[A]) - S(\rho) = \Omega[A] + S(\rho \| \rho_A).$$

But  $S(\rho \| \rho_A) \geq 0$  with equality if and only if  $\rho = \rho_A$ . This completes the proof.  $\square$

For notational convenience, we define a general notion of 1-RDM, as well as the map that associates to every density matrix its 1-RDM.

**Definition 10.** Let  $\mathcal{D}_d^{(1)} := \{D \in \mathbf{H}^d : 0 \preceq D \preceq 1\}$ . We refer to elements of  $\mathcal{D}_d^{(1)}$  as *one-body reduced density matrices*, or *1-RDMs*. Define  $\gamma : \mathcal{D}_d \rightarrow \mathcal{D}_d^{(1)}$  by  $\gamma_{ij}(\rho) = \text{Tr}[\rho c_i^\dagger c_j]$ .

For this definition, we have to check that  $\gamma$  as defined actually maps  $\mathcal{D}_d$  into  $\mathcal{D}_d^{(1)}$ :

**Lemma 11.** For  $\rho \in \mathcal{D}_d$ , the matrix  $D \in \mathbf{H}^d$  defined by  $D_{ij} = \text{Tr}[\rho c_i^\dagger c_j]$  is an element of  $\mathcal{D}_d^{(1)}$ .

*Proof.* To see this, let  $\rho \in \mathcal{D}_d$ , and let  $z \in \mathbb{C}^d$  with  $\|z\|^2 = \sum_{i=1}^d |z_i|^2 = 1$ . Then

$$z^* \gamma(\rho) z = \sum_{ij} \bar{z}_i z_j \text{Tr}[c_i^\dagger c_j \rho] = \text{Tr}[\tilde{c}^\dagger \tilde{c} \rho],$$

where  $\tilde{c} := \sum_{i=1}^d z_i c_i$ . Now  $\tilde{c}^\dagger \tilde{c}$  and  $\rho$  are positive semidefinite operators, so  $\text{Tr}[\tilde{c}^\dagger \tilde{c} \rho] \geq 0$ . It remains to show that  $\text{Tr}[\tilde{c}^\dagger \tilde{c} \rho] \leq 1$ . To see this observe that

$$\begin{aligned} \tilde{c}^\dagger \tilde{c} + \tilde{c} \tilde{c}^\dagger &= \sum_{i,j=1}^d \bar{z}_i z_j c_i^\dagger c_j + \sum_{i,j=1}^d z_i \bar{z}_j c_i c_j^\dagger \\ &= \sum_{i,j=1}^d \bar{z}_i z_j \left( c_i^\dagger c_j + c_j c_i^\dagger \right) \\ &= \sum_{i,j=1}^d \bar{z}_i z_j (\delta_{ij} \text{Id}_{\mathcal{F}_d}) \\ &= \text{Id}_{\mathcal{F}_d}. \end{aligned}$$

Since  $\tilde{c}^\dagger \tilde{c}, \tilde{c} \tilde{c}^\dagger \succeq 0$ , it follows that  $\tilde{c}^\dagger \tilde{c} \preceq \text{Id}_{\mathcal{F}_d}$ . Therefore  $\text{Tr}[\tilde{c}^\dagger \tilde{c} \rho] \leq \text{Tr}[\rho] = 1$ , as was to be shown.  $\square$

Finally, we obtain the concave conjugate of  $\Omega$  via a ‘contraction’ principle:

**Lemma 12.** For  $A \in \mathbf{H}^d$ ,

$$\Omega[A] = \inf_{D \in \mathcal{D}_d^{(1)}} (\text{Tr}[AD] - \mathcal{F}[D]),$$

where

$$\mathcal{F}[D] := \sup_{\rho \in \gamma^{-1}(D)} [S(\rho) - \text{Tr}(\rho \hat{V})]$$

is the 1-RDM functional.  $\mathcal{F}$  can be thought of as a function on all of  $\mathbf{H}^d$  via the convention that  $\mathcal{F}[D] = -\infty$  if  $\gamma^{-1}(D)$  is empty.

*Proof.* Write

$$\begin{aligned} \Omega[A] &= \inf_{D \in \mathcal{D}_d^{(1)}} \inf_{\rho \in \gamma^{-1}(D)} \left[ \sum_{i,j=1}^d A_{ij} \text{Tr}(\rho c_i^\dagger c_j) + \text{Tr}(\rho \hat{V}) - S(\rho) \right] \\ &= \inf_{D \in \mathcal{D}_d^{(1)}} \left( \text{Tr}[AD] + \inf_{\rho \in \gamma^{-1}(D)} [\text{Tr}(\rho \hat{V}) - S(\rho)] \right). \end{aligned}$$

□

Then we have the following properties of  $\mathcal{F}$ :

**Lemma 13.**  $\mathcal{F}$  is finite and concave on  $\mathcal{D}_d^{(1)}$ .

*Remark 14.* Since  $\mathcal{F}$  is finite on  $\mathcal{D}_d^{(1)}$  and  $\mathcal{F} \equiv -\infty$  on  $\mathbf{H}^d \setminus \mathcal{D}_d^{(1)}$  (since  $\gamma^{-1}(D) = \emptyset$  for all  $D$  in this set), in the language of convex analysis we say that  $\mathcal{D}_d^{(1)}$  is the *effective domain* of  $\mathcal{F}$  and write  $\mathcal{D}_d^{(1)} = \text{dom}\mathcal{F}$ .

*Proof.* First we show that  $\mathcal{F}$  is finite. Let  $D \in \mathcal{D}_d^{(1)}$ . The fact that  $\mathcal{F}[D] < \infty$  follows from the facts that  $S(\rho) \leq d \log 2$  and  $\text{Tr}[\rho \hat{V}] \leq \|V\|_{\text{op}}$  for all  $\rho \in \mathcal{D}_d$ . Then to show that  $\mathcal{F}[D] > -\infty$  is to show that  $\gamma^{-1}(D)$  is nonempty. In fact, for such  $D$ , we can always find a *noninteracting* density matrix  $\rho$  such that  $\gamma(\rho) = d$ .

Inspired by Appendix F, we let  $A = \log(D^{-1} - 1)$ . If  $D$  is on the boundary of  $\mathcal{D}_d^{(1)}$ , then formally we allow  $A$  to have possibly eigenvalues  $\pm\infty$ . Then let  $\rho = \rho_0[A] = \frac{1}{Z_0[A]} e^{-c^\dagger A c}$ . Even for  $A$  understood formally to have infinite eigenvalues,  $\rho$  is well-defined as a density matrix, and  $\gamma(\rho_0[A]) = D$  (see Appendix F).

Now we prove convexity. Let  $D_1, D_2 \in \mathcal{D}_d^{(1)}$  and  $\theta \in [0, 1]$ . Let  $\varepsilon \geq 0$ , and choose  $\rho_1 \in \gamma^{-1}(D_1)$  and  $\rho_2 \in \gamma^{-1}(D_2)$  such that  $\mathcal{F}[D_i] \leq S(\rho_i) - \text{Tr}(\rho_i \hat{V}) + \varepsilon$  for  $i = 1, 2$ . Then  $\gamma(\theta D_1 + (1 - \theta) D_2) = \theta D_1 + (1 - \theta) D_2$ , so we have from the definition of  $\mathcal{F}$  that

$$\mathcal{F}[\theta D_1 + (1 - \theta) D_2] \geq S(\theta \rho_1 + (1 - \theta) \rho_2) - \left( \text{Tr}[\theta \rho_1 + (1 - \theta) \rho_2] \hat{V} \right).$$



Then by the concavity of  $S$ ,

$$\begin{aligned} \mathcal{F}[\theta D_1 + (1 - \theta)D_2] &\geq \theta \left[ S(\rho_1) - \text{Tr}(\rho_1 \hat{V}) \right] + (1 - \theta) \left[ S(\rho_2) - \text{Tr}(\rho_2 \hat{V}) \right] \\ &= \theta \mathcal{F}[D_1] + (1 - \theta) \mathcal{F}[D_2] - \varepsilon. \end{aligned}$$

Since  $\varepsilon > 0$  was arbitrary, the concavity of  $\mathcal{F}$  follows.  $\square$

Now in the language of convex analysis, Lemma 12 precisely means that  $\Omega$  is the concave conjugate of  $\mathcal{F}$ , and we write  $\Omega = \mathcal{F}^*$ . But because  $\mathcal{F}$  is concave,  $\Omega^* = \mathcal{F}^{**}$  is the upper semicontinuous hull of  $\mathcal{F}$  (see Appendix C). Since  $\mathcal{F}$  is concave on  $\mathcal{D}_d^{(1)}$ , it follows that  $\mathcal{F}$  is in fact continuous on  $\text{int } \mathcal{D}_d^{(1)}$  (Appendix C), hence agrees with  $\mathcal{F}^{**}$  on this set. Therefore  $\Omega^* = \mathcal{F}$  on  $\text{int } \mathcal{D}_d^{(1)}$ . In fact, our later analysis (see Remark 18 below) implies in particular that  $\mathcal{F}$  is continuous up to the boundary of  $\mathcal{D}_d^{(1)}$ , hence  $\mathcal{F}^{**} = \mathcal{F}$ , and  $\Omega^* = \mathcal{F}$ . In summary, we have

**Theorem 15.**  $\Omega^* = \mathcal{F}$ .

Meanwhile, tentatively defining  $A[D] = \nabla \mathcal{F}[D]$  (pending the proof of the differentiability of  $\mathcal{F}$ ), we expect that  $D[A] = \nabla \Omega[A]$  and  $A[D]$  are inverses of one another (on appropriate sets), as is guaranteed by the following lemma.

**Lemma 16.**  $D[A] = \nabla \Omega[A]$  is a bijection  $\mathbf{H}^d \rightarrow \text{int } \mathcal{D}_d^{(1)}$  with inverse given by  $A[D] = \nabla \mathcal{F}[D]$ .

*Proof.* In particular we need to show that  $\mathcal{F}$  is differentiable on  $\text{int } \mathcal{D}_d^{(1)}$ . Since  $\Omega$  is differentiable and strictly convex on  $\mathbf{H}^d$ , it follows that  $\Omega^* = \mathcal{F}^{**}$  is differentiable and strictly convex on  $\text{int } \mathcal{D}_d^{(1)}$  and that  $\nabla \Omega$  is a bijection  $\mathbf{H}^d \rightarrow \text{int } \mathcal{D}_d^{(1)}$  with inverse given by  $\nabla \mathcal{F}^{**}$ ; see Appendix C) or Theorem 26.5 of [91]. Since  $\mathcal{F}^{**} = \mathcal{F}$  on  $\text{int } \mathcal{D}_d^{(1)}$ , the result follows.  $\square$

Lemma 16 specifies, roughly speaking, a one-to-one correspondence between the 1-RDM and the single-particle part of the Hamiltonian.

## 2.1 The noninteracting case

We pause now to consider the noninteracting case, in which  $\hat{V} = 0$ . To indicate this case, we write  $\hat{H}_0[A] = c^\dagger A c$ ,  $Z_0[A] = Z[A; 0]$ ,  $\rho_0[A] = \frac{1}{Z_0[A]} e^{-\hat{H}_0[A]}$ ,  $\Omega_0[A] = \Omega[A; 0]$ ,  $D_0[A] = D[A; 0]$ ,  $\mathcal{F}_0[D] = \mathcal{F}[D; 0]$ , and  $A_0[D] = A[D; 0]$ .

In Appendix F, we prove that

$$Z_0[A] = \det(1 + e^{-A}),$$

from which it follows that

$$\Omega_0[A] = -\log \det(1 + e^{-A}) = -\text{Tr}[\log(1 + e^{-A})],$$

so

$$D_0[A] = \nabla \Omega_0[A] = \frac{e^{-A}}{1 + e^{-A}} = (1 + e^A)^{-1}.$$

Therefore the inverse map  $A_0[D]$  is given by

$$A_0[D] = \log(D^{-1} - 1).$$

Then via Legendre duality, we compute  $\mathcal{F}_0$  as

$$\begin{aligned} \mathcal{F}_0[D] &= \text{Tr}[DA_0[D]] - \Omega_0[A_0[D]] \\ &= \text{Tr}[D \log(D^{-1} - 1)] + \text{Tr}\left[\log\left(1 + e^{-\log(D^{-1} - 1)}\right)\right] \\ &= \text{Tr}[D \log(D^{-1}(1 - D))] + \text{Tr}\left[\log\left(1 + \frac{1}{D^{-1} - 1}\right)\right] \\ &= -\text{Tr}[D \log D] + \text{Tr}[D \log(1 - D)] + \text{Tr}\left[\log\left(\frac{D^{-1}}{D^{-1} - 1}\right)\right] \\ &= -\text{Tr}[D \log D] + \text{Tr}[D \log(1 - D)] + \text{Tr}\left[\log\left(\frac{1}{1 - D}\right)\right] \\ &= -\text{Tr}[D \log D] + \text{Tr}[D \log(1 - D)] - \text{Tr}[\log(1 - D)] \\ &= -\text{Tr}[D \log D] - \text{Tr}[(1 - D) \log(1 - D)], \end{aligned}$$

so

$$\mathcal{F}_0[D] = -\text{Tr}[D \log D] - \text{Tr}[(1 - D) \log(1 - D)]$$

is a sort of ‘matrix binary entropy.’ In fact we have proved this latter fact only for  $D \in \text{int } \mathcal{D}_d^{(1)}$ , but it follows on all of  $\mathcal{D}_d^{(1)}$  by the continuity of  $\mathcal{F}$  up to the boundary (Remark 18 below).

## 2.2 Comparison with the Luttinger-Ward formalism

The developments here can be viewed as a ‘quantum-statistical-mechanical’ analog of the development of the Luttinger-Ward formalism in the ‘classical-statistical-mechanical’ setting (i.e., the setting of Gibbs measures); see Part III.

Now to identify the analog of ‘Luttinger-Ward functional’ in this setting, one should take the difference of the interacting and noninteracting  $\mathcal{F}$  functionals. Indeed, we may define

$$\begin{aligned} \Phi[D] &:= \mathcal{F}[D] - \mathcal{F}_0[D] \\ &= \mathcal{F}[D] + \text{Tr}[D \log D] + \text{Tr}[(1 - D) \log(1 - D)]. \end{aligned}$$

Taking gradients, this yields a ‘Dyson equation’

$$A = \nabla \mathcal{F}_0[D] + \Sigma[D] = \log(D^{-1} - 1) + \Sigma[D]$$

where  $\Sigma := \nabla\Phi$  is the analog of the ‘self-energy,’ so

$$D^{-1} = 1 + e^{A-\Sigma[D]}.$$

As in the classical case, the self-energy defines an effective single-particle Hamiltonian (via  $A - \Sigma$ ) whose noninteracting 1-RDM is precisely the interacting 1-RDM function for  $A$ .

An important difference between this setting and the classical setting is that *there is no transformation rule*, i.e., for an arbitrary, linear transformation  $T : \mathbb{R}^d \rightarrow \mathbb{R}^d$ , there is no simple recipe for computing  $\mathcal{F}[TDT^*]$  or  $\Phi[TDT^*]$  in terms of  $\mathcal{F}[D]$  or  $\Phi[D]$ . (However, there *is* a transformation rule for *unitary* linear transformations, which can be established via the canonical transformation of the creation and annihilation operators.) Notably, the failure of the transformation rule entails the failure of the projection rule; this prevents us from simply downfolding the bath of an impurity problems to derive an effective single-particle matrix. (Instead, as can be understood from the path integral perspective in which the formal similarity to the classical field theory is visible, one obtains a *dynamical* effective single-particle matrix. However, this development is orthogonal to the discussion of this section.)

### 3 The embedding lemma and boundary analysis

Motivated by developments for the classical Luttinger-Ward functional, we want to describe the behavior of  $\mathcal{F}$  on the boundary of its domain, i.e.,  $\partial\mathcal{D}_d^{(1)}$ .

Consider block-diagonal  $D$  of the form

$$D = \begin{pmatrix} D_p & 0 & 0 \\ 0 & I_q & 0 \\ 0 & 0 & 0_r \end{pmatrix}, \quad (3.1)$$

where  $G_p \in \mathcal{D}_p^{(1)}$  and  $p + q + r = d$ . Via canonical transformation, to determine the boundary values of  $\mathcal{F}$ , it suffices to determine the value of  $\Phi$  for  $G$  of this form.

Identify the first  $p$  lattice sites as  $X$  and the last  $q + r$  as  $Y$ , and let the corresponding Fock spaces be  $\mathcal{F}_X$  and  $\mathcal{F}_Y$ , respectively, so  $\mathcal{F} = \mathcal{F}_X \otimes \mathcal{F}_Y$ , where the tensor product is considered with respect to the occupation number representations of the Fock spaces in question.

Finally, define

$$|\Psi_Y\rangle := c_{p+1}^\dagger \cdots c_{p+q}^\dagger |-\rangle \in \mathcal{F}_Y.$$

Then we have the following fundamental result, which constrains the form of any density operator that yields a 1-RDM of the form (3.1).

**Lemma 17** (Embedding lemma). *Suppose  $\rho \in \mathcal{D}_d$  and  $\gamma(\rho) = D$ , where  $D$  is of the form (3.1). Then*

$$\rho = \rho_X \otimes |\Psi_Y\rangle\langle\Psi_Y|.$$

*Proof.* First define  $\rho_X := \text{Tr}_Y[\rho]$  and  $\rho_Y := \text{Tr}_X[\rho]$ . (We do not yet know that  $\rho$  factorizes as  $\rho = \rho_X \otimes \rho_Y$ .) Then evidently

$$\gamma(\rho_Y) = \begin{pmatrix} I_p & 0 \\ 0 & 0_r \end{pmatrix}.$$

We claim that this implies that  $\rho_Y$  is a pure state, to wit, the projector onto the Slater determinant  $|\Psi_Y\rangle$ .

To this end, write

$$\rho_Y = \sum_i \alpha_i |Y_i\rangle \langle Y_i|,$$

where  $Y_i \in \mathcal{F}_Y$  are orthonormal and  $\alpha_i > 0$  with  $\sum_i \alpha_i = 1$ . Then for  $j = p + 1, \dots, p + q$ ,

$$1 = \text{Tr}[a_j^\dagger a_j \rho_Y] = \sum_i \alpha_i \langle Y_i | a_j^\dagger a_j | Y_i \rangle. \quad (3.2)$$

Now  $\langle Y_i | a_j^\dagger a_j | Y_i \rangle = \|a_j |Y_i\rangle\|^2 \geq 0$ , but also  $\|a_j^\dagger a_j\| = 1$ , so  $\langle Y_i | a_j^\dagger a_j | Y_i \rangle \leq 1$ . Therefore, in order for (3.2) to hold, it must be the case that

$$\langle Y_i | a_j^\dagger a_j | Y_i \rangle = 1.$$

Since  $\|a_j^\dagger a_j\| = 1$ , it must be the case that  $|Y_i\rangle$  is a 1-eigenstate of  $a_j^\dagger a_j$  for all  $j = p + 1, \dots, p + q$  and all  $i$ .

Meanwhile for all  $j > p + 1$ , we have

$$0 = \sum_i \alpha_i \langle Y_i | a_j^\dagger a_j | Y_i \rangle,$$

so it must be the case that

$$\|a_j |Y_i\rangle\|^2 = \langle Y_i | a_j^\dagger a_j | Y_i \rangle = 0.$$

Then  $a_j |Y_i\rangle = 0$ , and  $|Y_i\rangle$  is a 0-eigenstate of  $a_j^\dagger a_j$  for all  $j > p + q$  and all  $i$ .

Since each  $|Y_i\rangle$  is a simultaneous eigenstate of all the  $a_j^\dagger a_j$  (with eigenvalue 1 for  $j = p + 1, \dots, p + q$  and 0 for  $j > p + q$ ), it follows that in the particle number representation (up to scaling by a complex number of unit modulus) we can write

$$|Y_i\rangle = |\underbrace{11 \dots 1}_q \underbrace{00 \dots 0}_r\rangle = |\Psi_Y\rangle$$

for all  $i$ , so in fact we have

$$\rho_Y = |\Psi_Y\rangle \langle \Psi_Y|,$$

as claimed.

Now since  $\rho_Y$  is pure,  $S(\rho_Y) = 0$ , and we have by the ‘triangle inequality’ for the von Neumann entropy [4] that

$$S(\rho) \leq S(\rho_X) + S(\rho_Y) = S(\rho_X) - S(\rho_Y) \leq S(\rho),$$

Therefore

$$S(\rho) = S(\rho_X) + S(\rho_Y) = S(\rho_X),$$

i.e., subadditivity holds with equality, which implies that in fact  $\rho = \rho_X \otimes \rho_Y$ .  $\square$

*Remark 18.* A more refined argument in the case that  $\gamma(\rho)$  not necessarily equal to  $D$  yields  $\|\rho_Y - |\Psi_Y\rangle\langle\Psi_Y|\| = o(\|\gamma(\rho) - D\|)$  and in turn that  $\|\rho - \rho_X \otimes |\Psi_Y\rangle\langle\Psi_Y|\| = o(\|\gamma(\rho) - D\|)$ . This fact, together with the boundary formula for  $\mathcal{F}$  proved in Proposition 19 below, can be used without great difficulty to establish the continuity of  $\mathcal{F}$  up to the boundary.

Now, armed with an understanding of the states that can produce boundary elements of  $\mathcal{D}_d^{(1)}$ , we will show that

$$\mathcal{F}[D; \hat{V}] = \mathcal{F}\left[D_p; \hat{V}_X\right]$$

for  $D$  as in (3.1), where  $\hat{V}_X := \langle\Psi_Y|\hat{V}|\Psi_Y\rangle$  is an operator on  $\mathcal{F}_X$ . More precisely, if one writes  $\hat{V} = \sum_k O_X^{(k)} \otimes O_Y^{(k)}$  for operators  $O_X^{(k)}$  and  $O_Y^{(k)}$  on  $\mathcal{F}_X$  and  $\mathcal{F}_Y$ , then  $\hat{V}_X = \sum_k O_X^{(k)} \langle\Psi_Y|O_Y^{(k)}|\Psi_Y\rangle$ .

One can check that this construction is well-defined. Indeed, suppose that we could write

$$\hat{V} = \sum_k O_X^{(k)} \otimes O_Y^{(k)} = \sum_k Q_X^{(k)} \otimes Q_Y^{(k)}.$$

Then we must show that

$$\sum_k O_X^{(k)} \langle\Psi_Y|O_Y^{(k)}|\Psi_Y\rangle = \sum_k Q_X^{(k)} \langle\Psi_Y|Q_Y^{(k)}|\Psi_Y\rangle$$

as operators on  $\mathcal{F}_X$ . This holds if and only if

$$\langle X_1 | \left( \sum_k O_X^{(k)} \langle\Psi_Y|O_Y^{(k)}|\Psi_Y\rangle \right) | X_2 \rangle = \langle X_1 | \left( \sum_k Q_X^{(k)} \langle\Psi_Y|Q_Y^{(k)}|\Psi_Y\rangle \right) | X_2 \rangle$$

for all  $|X_1\rangle, |X_2\rangle \in \mathcal{F}_X$ . But both sides are equal to  $\langle X_1 \Psi_Y | \hat{V} | X_2 \Psi_Y \rangle$ .

We comment that one can equivalently define

$$\hat{V}_X = \langle\Psi_Y|\hat{V}|\Psi_Y\rangle = \text{Tr}_Y \left[ \hat{V} (\text{Id}_{\mathcal{F}_X} \otimes |\Psi_Y\rangle\langle\Psi_Y|) \right].$$

Then we are prepared to state:

**Proposition 19.** *If  $D$  is of the form (3.1), then*

$$\mathcal{F}[D; \hat{V}] = \mathcal{F} \left[ D_p; \langle \Psi_Y | \hat{V} | \Psi_Y \rangle \right],$$

where  $\Psi_Y \in \mathcal{F}_Y$  is defined

$$|\Psi_Y\rangle := a_{p+1}^\dagger \cdots a_{p+q}^\dagger |-\rangle \in \mathcal{F}_Y.$$

*Proof.* By Lemma 17, if  $\gamma(\rho) = D$ , then  $\rho = \rho_X \otimes \rho_Y$ , where  $\rho_Y = |\Psi_Y\rangle\langle\Psi_Y|$ . Therefore

$$\begin{aligned} \mathcal{F}[D; \hat{V}] &= \sup_{\rho \in \gamma^{-1}(D)} \left[ S(\rho) - \text{Tr}(\rho \hat{V}) \right] \\ &= \sup_{\rho_X \in \gamma^{-1}(D_p)} \left( S(\rho_X) - \text{Tr} \left[ (\rho_X \otimes |\Psi_Y\rangle\langle\Psi_Y|) \hat{V} \right] \right). \end{aligned}$$

But writing  $\hat{V} = \sum_k O_X^{(k)} \otimes O_Y^{(k)}$ , we have

$$\begin{aligned} \text{Tr} \left[ (\rho_X \otimes |\Psi_Y\rangle\langle\Psi_Y|) \hat{V} \right] &= \sum_k \langle \Psi_Y | \left( O_X^{(k)} \otimes O_Y^{(k)} \right) (\rho_X \otimes |\Psi_Y\rangle\langle\Psi_Y|) | \Psi_Y \rangle \\ &= \sum_k \langle \Psi_Y | \left( \left[ O_X^{(k)} \rho_X \right] \otimes \left[ O_Y^{(k)} |\Psi_Y\rangle\langle\Psi_Y| \right] \right) | \Psi_Y \rangle \\ &= \sum_k O_X^{(k)} \rho_X \langle \Psi_Y | O_Y^{(k)} | \Psi_Y \rangle \\ &= \hat{V}_X \rho_X, \end{aligned}$$

where  $\hat{V}_X = \langle \Psi_Y | \hat{V} | \Psi_Y \rangle$ . Thus

$$\mathcal{F}[D; \hat{V}] = \sup_{\rho_X \in \gamma^{-1}(D_p)} \left[ S(\rho_X) - \text{Tr} \left( \rho_X \hat{V}_X \right) \right] = \mathcal{F}[D_p; \hat{V}_X],$$

as was to be shown. □

## 4 The zero-temperature limit

Thus far we have simply assumed the inverse temperature to be given by  $\beta = 1$ . However, now we consider scaled interaction  $\beta V$ . We will add an extra argument  $\beta$  to all of our functionals. Notice that  $\nabla_D \mathcal{F}[D; \beta \hat{V}]$  corresponds to  $\beta A[D; \hat{V}; \beta]$ . Therefore we define  $\mathcal{F}[D; \hat{V}; \beta] := \beta^{-1} \mathcal{F}[D; \beta \hat{V}]$ , i.e.,

$$\mathcal{F}[D; \hat{V}; \beta] = \sup_{\rho \in \gamma^{-1}(D)} \left[ \beta^{-1} S(\rho) - \text{Tr}(\rho \hat{V}) \right].$$

Then in the  $\beta \rightarrow \infty$  limit we find

$$\mathcal{F}[D; \hat{V}; \infty] = \sup_{\rho \in \gamma^{-1}(D)} \text{Tr}(\rho \hat{V}).$$

Note that the supremum will be attained by  $\rho$  on the boundary of the set of density matrices, whereas in the finite temperature case, the supremum will always be attained on the interior.

The free-energy functional

$$\Omega[A; \hat{V}; \beta] = \mathcal{F}^*[A; \hat{V}; \beta] = \beta^{-1} \Omega[\beta A; \beta \hat{V}]$$

can be written

$$\Omega[A; \hat{V}; \beta] = -\beta^{-1} \log \text{Tr} \left[ e^{-\beta(c^\dagger A c + \hat{V})} \right].$$

For finite temperature,  $\Omega$  and  $\mathcal{F}$  are both smooth and strictly convex. However, in the zero temperature limit,  $\Omega$  develops singularities and  $\mathcal{F}$  develops degenerate (flat) regions. The flat regions represent 1-RDMs that are not attainable by any single-particle Hamiltonian with a unique ground state. In the case  $\hat{V} = 0$ , we discover  $\mathcal{F}[D; 0; \infty] \equiv 0$ . This reflects the fact that, in the noninteracting case, the only 1-RDMs that are attainable in this sense are on the boundary of  $\mathcal{D}_d^{(1)}$ . When  $\hat{V} \neq 0$ , however, scaling does affect the 1-RDM, and there are attainable 1-RDMs that are not in  $\partial \mathcal{D}^M$ . However, not all 1-RDMs are attainable. By smoothly varying the single-particle Hamiltonian, we can cause the ground state of the Hamiltonian to become the state of second-lowest energy. The discontinuous jump to the new ground state reflects a singularity in  $\Omega[\cdot; V; \infty]$  or a jump in the 1-RDM (which at this point passes over a flat, i.e., unattainable region of  $\mathcal{D}^M$ ).

By the same proof as in the preceding section, we have the following:

**Proposition 20.** *Let  $\beta \in (0, \infty]$ . If  $D$  is of the form (3.1), then*

$$\mathcal{F}[D; \hat{V}; \beta] = \mathcal{F} \left[ D_p; \langle \Psi_Y | \hat{V} | \Psi_Y \rangle; \beta \right],$$

where  $\Psi_Y \in \mathcal{F}_Y$  is defined

$$|\Psi_Y\rangle := c_{p+1}^\dagger \cdots c_{p+q}^\dagger |-\rangle \in \mathcal{F}_Y.$$

## 5 Two-body embedding Hamiltonian

Motivated by the previous section, we discuss the computation of the *embedding Hamiltonian*  $\hat{V}_X = \langle \Psi_Y | \hat{V} | \Psi_Y \rangle$  in the case of the two-body interaction

$$\hat{V} = \frac{1}{2} \sum_{ijkl=1}^M (ij|kl) c_i^\dagger c_k^\dagger c_l c_j,$$

where  $|\Psi_Y\rangle \in \mathcal{F}_Y$ .

To this end, it suffices to compute  $\langle \Psi_Y | c_i^\dagger c_k^\dagger c_l c_j | \Psi_Y \rangle$ . Recall that, in the occupation number representation,

$$c_i^\dagger = s \otimes \cdots \otimes s \otimes a^\dagger \otimes I_2 \otimes \cdots \otimes I_2,$$

where

$$s = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad a = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}.$$

Therefore

$$c_i^\dagger c_k^\dagger c_l c_j = O_X \otimes c_{Y,i}^\dagger c_{Y,k}^\dagger c_{Y,l} c_{Y,j},$$

where on the RHS, the  $c_{Y,i}$  are interpreted as operators on  $\mathcal{F}_Y$ , with  $c_i = \text{Id}_{\mathcal{F}_Y}$  for  $i \leq p$  and  $c_{Y,i} = c_i$  for  $i > p$ , and  $O_X$  is an operator on  $\mathcal{F}_X$ . Note that we will have  $\langle \Psi_Y | c_i^\dagger c_k^\dagger c_l c_j | \Psi_Y \rangle = 0$  except in the following three scenarios:

1. none of the operators have index  $i > p$ ,
2. exactly one creation operator and exactly one annihilation operator have index  $i > p$ ,
3. all four of the operators have index  $i > p$ .

In scenario (1), we merely obtain

$$\langle \Psi_Y | c_i^\dagger c_k^\dagger c_l c_j | \Psi_Y \rangle = c_i^\dagger c_k^\dagger c_l c_j \in \text{End}(\mathcal{F}_X)$$

where  $i, j, k, l \leq p$ .

In scenario (3), the

$$\langle \Psi_Y | c_i^\dagger c_k^\dagger c_l c_j | \Psi_Y \rangle = \text{Id}_{\mathcal{F}_X} \langle \Psi_Y | c_i^\dagger c_k^\dagger c_l c_j | \Psi_Y \rangle,$$

where the  $c_i$  are interpreted on the RHS as operators on  $\mathcal{F}_Y$ .

In scenario (2), we can have, e.g.,  $i, j \leq p$  and  $k, l > p$ , in which case we have  $c_i^\dagger c_k^\dagger c_l c_j = c_i^\dagger c_j \otimes c_k^\dagger c_l$  and

$$\langle \Psi_Y | c_i^\dagger c_k^\dagger c_l c_j | \Psi_Y \rangle = c_i^\dagger c_j D_{kl},$$

or we can have  $i, l \leq p$  and  $k, j > p$ , in which case  $c_i^\dagger c_k^\dagger c_l c_j = -c_i^\dagger c_l \otimes c_k^\dagger c_j$  and

$$\langle \Psi_Y | c_i^\dagger c_k^\dagger c_l c_j | \Psi_Y \rangle = -c_i^\dagger c_l D_{kj},$$

or we can have  $k, l \leq p$  and  $i, j > p$ , in which case  $c_i^\dagger c_k^\dagger c_l c_j = c_k^\dagger c_l \otimes c_i^\dagger c_j$  and

$$\langle \Psi_Y | c_i^\dagger c_k^\dagger c_l c_j | \Psi_Y \rangle = c_k^\dagger c_l D_{ij},$$



or we can have  $k, j \leq p$  and  $i, l > p$ , in which case  $c_i^\dagger c_k^\dagger c_l c_j = -c_k^\dagger c_j \otimes c_i^\dagger c_l$  and

$$\langle \Psi_Y | c_i^\dagger c_k^\dagger c_l c_j | \Psi_Y \rangle = -c_k^\dagger c_j D_{il}.$$

Then the sum over all contributing terms of this form is

$$\sum_{ij=1}^p c_i^\dagger c_j \sum_{kl=p+1}^M [(ij|kl) - (il|kj)] D_{kl}.$$

Therefore, we have computed:

$$\hat{V}_X = E_Y + \sum_{ijkl=1}^p (ij|kl) c_i^\dagger c_k^\dagger c_l c_j + \sum_{ij}^p c_i^\dagger c_j \sum_{kl=p+1}^M [(ij|kl) - (il|kj)] D_{kl},$$

where  $E_Y$  is a constant. In the case  $|\Psi_Y\rangle = c_{p+1}^\dagger \cdots c_{p+q}^\dagger |-\rangle$ , we have

$$\hat{V}_X = E_Y + \sum_{ijkl=1}^p (ij|kl) c_i^\dagger c_k^\dagger c_l c_j + \sum_{ij}^p c_i^\dagger c_j \sum_{k=p+1}^{p+q} [(ij|kk) - (ik|kj)].$$

Not coincidentally, this formula recovers (up to a constant) the embedding Hamiltonian of the density matrix embedding theory (DMET) [49] and more generally of complete active space methods.

## Part VI

# Classical and quantum impurity problems

## 1 Introduction

To begin we recall the setting of Part III. Consider the second-moment matrix  $G \in \mathbb{R}^{d \times d}$  of a Gibbs measure defined by a Hamiltonian  $H : \mathbb{R}^d \rightarrow \mathbb{R}$ , i.e.,

$$G = \frac{1}{Z} \int_{\mathbb{R}^d} xx^T e^{-H(x)} dx. \quad (1.1)$$

Here the partition function

$$Z = \int_{\mathbb{R}^d} e^{-H(x)} dx$$

is the appropriate normalization factor.

We will write  $H$  in the form  $H = H_0 + U$ , where  $H_0 = \frac{1}{2}x^T Ax$  is a quadratic form. Assume that  $A$  and  $U$  are such that both  $Z$  and  $G$  are finite. Via the analogy with quantum many-body physics that will be discussed below, we refer to  $U$  as the *interacting* part of the Hamiltonian, or simply the *interaction*. Meanwhile  $H_0$  represents the *non-interacting* part.

If  $U \equiv 0$  and  $A$  is a positive definite matrix, then immediately we have  $G = A^{-1}$ . One seeks a generalization of this fact to the case in which  $U(x)$  depends only on a subset of the variables. We refer to this setting as the (classical) *impurity model* (cf. section 3.4 of Part III), by analogy to the quantum impurity model to be discussed below. Perhaps surprisingly, we have the following result:

**Theorem 1.** *Let  $p \leq d$ , and let  $A \in \mathbb{R}^{d \times d}$  be a symmetric matrix whose lower-right  $(d-p) \times (d-p)$  block is positive definite. Let  $U : \mathbb{R}^d \rightarrow \mathbb{R}$  be a function that depends only on its first  $p$  arguments, i.e.,  $U(x) = U_1(x_1, \dots, x_p)$  for some  $U_1 : \mathbb{R}^p \rightarrow \mathbb{R}$ , and assume that  $U_1$  satisfies sufficient growth conditions such that the Gibbs measure with density proportional to  $e^{-\frac{1}{2}x^T Ax - U(x)}$  has finite second-order moments. Then, with  $G$  defined as in (1.1),*

$$\Sigma := A - G^{-1} = \begin{pmatrix} \Sigma_p & 0 \\ 0 & 0 \end{pmatrix},$$

where  $\Sigma_p \in \mathbb{R}^{p \times p}$  is a symmetric matrix.

In fact, Theorem 1 can be generalized by considering an arbitrary measure  $d\mu_1(\mathbf{x}_1)$  of sufficient decay in the place of  $e^{-U_1(\mathbf{x}_1)} d\mathbf{x}_1$ , where we denote  $\mathbf{x}_1 = (x_1, \dots, x_p)^T$

and  $\mathbf{x}_2 = (x_{p+1}, \dots, x_d)^T$ . In this setting the partition function is defined

$$Z = \int_{\mathbb{R}^p} \int_{\mathbb{R}^{d-p}} e^{-\frac{1}{2}x^T A x} d\mathbf{x}_2 d\mu_1(\mathbf{x}_1),$$

and the Green's function is defined accordingly. The case

$$\mu_1(\mathbf{x}_1) = e^{-\sum_{i,j=1}^p J_{ij} x_i x_j} \sum_{\sigma \in \{-1,1\}^p} \delta(\cdot - \sigma)$$

defines a notion of a classical impurity model for spin systems, in which a spin system is coupled to a Gaussian ‘bath.’ For such a spin impurity model, we can assume without loss of generality that the upper-left  $p \times p$  block of  $A$  is zero, and the ensemble is specified by the partition function

$$Z = \sum_{\sigma \in \{-1,1\}^p} e^{-\frac{1}{2} \sum_{i,j=1}^p J_{ij} \sigma_i \sigma_j} \int_{\mathbb{R}^{d-p}} e^{-\frac{1}{2} y^T A_{22} y - y^T A_{21} \sigma} dy,$$

where  $A_{21}$  and  $A_{22}$  denote the appropriate blocks of  $A$ . We will stick to the original setting, in which the impurity is specified by a function  $U_1$ , to emphasize the analogy with the setting of the quantum many-body problem, but we comment that the proof of Theorem 1 is exactly the same in this broader context.

In statistics,  $G^{-1}$  is sometimes called the precision matrix. In our setting, if  $A$  is positive definite and  $U \equiv 0$ , then  $A$  is the precision matrix of the distribution in question. Hence Theorem 1 states that the difference of the precision matrices in the ‘interacting’ and ‘non-interacting’ settings, namely  $A - G^{-1}$ , is a *sparse* matrix if the interaction  $U$  only depends on a subset of variables. The proof of the theorem is non-perturbative, and in fact  $A$  need not be positive definite (though, when  $U$  is independent of the last  $d - p$  variables, the lower-right  $(d - p) \times (d - p)$  block of  $A$  must be positive definite to ensure that  $e^{-\frac{1}{2}x^T A x - U(x)}$  is normalizable). To the best of our knowledge, other than from the perspective of the Luttinger-Ward formalism presented in Part III, this basic linear-algebraic fact about Gibbs measures was not previously present in the literature.

As a matter of fact, we first observed a result of this type in a more complex setting, namely that of quantum impurity problems at zero temperature (as we shall discuss below, the analogous result is also true at finite temperature). Consider the Hamiltonian, denoted by  $\hat{H}$ , for a system of interacting fermions or bosons. Throughout we shall distinguish the cases of fermions and bosons via a parameter  $\zeta$  given by  $\zeta = -1$  in the case of fermions and  $\zeta = +1$  in the case of bosons. In the second-quantized representation [33],  $\hat{H}$  can be generally written as  $\hat{H} = \hat{H}_0 + \hat{U}$ , where

$$\hat{H}_0 = \sum_{i,j=1}^d h_{ij} a_i^\dagger a_j \tag{1.2}$$

is viewed as the Hamiltonian for a system of non-interacting fermions or bosons. Here  $a_i^\dagger, a_j$  are the creation and annihilation operators, respectively, and  $h \in \mathbb{C}^{d \times d}$  is a Hermitian matrix. (Refer to section 4 of Part I for a brief introduction to second quantization.)

Meanwhile,  $\hat{U}$  is the interacting part of the Hamiltonian. Although  $\hat{U}$  can be far more general, usually we have in mind the two-body interaction

$$\hat{U} = \sum_{i,j,k,l} (ij|U|kl) a_i^\dagger a_j^\dagger a_l a_k. \quad (1.3)$$

In this case, if there exists  $p < d$  so that  $(ij|U|kl) \neq 0$  only if  $i, j, k, l \in \{1, \dots, p\}$ , then we call the Hamiltonian  $\hat{H}$  an impurity Hamiltonian. More generally, we say that  $\hat{H}$  is an impurity Hamiltonian if  $\hat{U}$  can be written as a polynomial of the creation and annihilation operators  $a_i^\dagger$  and  $a_i$  for  $i = 1, \dots, p$  and is particle-number-conserving (see section 4 of Part I for details). At a glance there is no connection between this impurity Hamiltonian and the type of Gibbs measure discussed earlier. Nonetheless, we claim that there is an analogy under which  $h$  maps to  $A$ , the Green's function of the quantum many-body problem maps to  $G$ , and the self-energy matrix associated with the Green's function maps to  $\Sigma$ . Then the counterpart of Theorem 1 can be stated in words as: *the self-energy matrix of a quantum impurity problem is a sparse matrix, with nonzero entries only on the block associated with the impurity sites.*

The connection between the classical impurity model and the quantum impurity problem can be understood formally by writing quantum Green's functions in terms of the coherent state path integral [77], which formally resembles a Gibbs measure. We remark that in the case of fermions, the resemblance should be noted with special caution because the coherent state path integral involves Grassmann integrals. In this sense, the setting of Theorem 1 can indeed be understood as the 'classical impurity problem.'

Unlike the corresponding result for Gibbs measures, the quantum result has been well-known in the quantum physics literature since Feynman and Vernon in 1963 [35] at the latest, and it plays a central role in numerical algorithms for solving the quantum impurity problem, such as the quantum Monte Carlo (QMC) method [41]. The sparsity of the impurity self-energy matrix is also the starting point of various approximate methods—such as the dynamical mean field theory (DMFT) [37, 53] and its extensions [106, 58]—for solving general (i.e., non-impurity) quantum systems, especially those that are strongly correlated. Again somewhat surprisingly, this important statement is to the best of our knowledge a 'folk theorem,' in that we cannot find a rigorous proof of this result in the literature.

In this Part, we fill this gap by providing a rigorous proofs of the sparsity of the self-energy matrix of quantum impurity problems, in both the fermionic and bosonic cases at zero and finite temperature. We will also cover the non-equilibrium setting via the consideration of arbitrary contour-ordered Green's functions, as well as the anomalous setting, which is relevant to superconductivity. Excellent introductions to

the non-equilibrium and anomalous formalisms can be found in [100, 14], respectively.

Our results in the non-equilibrium setting should be compared to those in a recent work [25], in which *advanced/retarded* non-equilibrium self-energies are rigorously constructed in the case of fermions. Though not noted explicitly in the work, the appropriate sparsity results for these quantities can be seen to follow from the construction itself. By contrast, our non-equilibrium sparsity result concerns the *contour-ordered* self-energy (for both fermions and bosons) and in particular recovers sparsity results for the advanced/retarded Green's functions. Moreover, our result holds for arbitrary contour. However, we do not actually construct the contour-ordered self-energy, but rather phrase our sparsity result in terms of operators that we suggestively name ' $G\Sigma$ ' and ' $\Sigma G$ .'<sup>9</sup> In so doing we sidestep a considerable analytical challenge such as that encountered in [25]. Thus our result can be viewed as trying to parsimoniously illustrate the broadest possible formal picture of sparsity results for the self-energy, rather than focusing on the analytical question of the construction of the self-energy itself. Incidentally, in our view a rigorous construction of the contour-ordered self-energy (for arbitrary contour) seems to be an interesting and non-trivial matter.

The reader is directed to section 7 of Part I for a brief introduction to the theory of Green's functions, both fermionic and bosonic, in the zero-temperature and finite-temperature settings. Below we also include introduction to the non-equilibrium and anomalous settings. In all of these settings, the impurity model with  $p = 0$  is precisely the non-interacting model, and our results on the sparsity pattern of the self-energy, applied in this special case, yield formulas for the non-interacting Green's functions. In the non-equilibrium setting especially, such a formula seems to be non-trivial to establish by other means. Readers new to the subject may find this presentation of the non-interacting Green's functions, as well as its embedding into a unified perspective, to be appealing in its own right. Please note that this Part is based on [63] (joint work with Lin Lin).

## 1.1 Other formal perspectives

We discuss several other ways of understanding the sparsity pattern of the self-energy for impurity problems. First, we remark by considering the coherent-state path integral representation [77] (in any of the quantum settings discussed in this Part), one can formally view the quantum many-body ensemble as a Gibbs measure. The proof of Theorem 1 can be mimicked in these settings at the formal level to derive the appropriate sparsity results, but we omit such formal manipulations here.

Secondly, the sparsity pattern can be most intuitively understood via the Feynman diagrammatic expansion, which provides another viewpoint on the formal unification

---

<sup>9</sup>The reader will find that from the point of view adopted in this Part, these objects can be thought of as more natural than the non-equilibrium self-energy itself, and indeed all of our sparsity results are proved by considering their analogs.

of the classical and quantum settings. Indeed, due to the connection between the classical setting of Gibbs measures and the coherent state path integral, we limit our discussion the case of Gibbs models here for simplicity. The reader may refer to Part II for a self-contained introduction to the diagrammatic expansion.

As before, define the partition function

$$Z = \int_{\mathbb{R}^d} e^{-\frac{1}{2}x^T A x - \varepsilon U(x)} dx, \quad (1.4)$$

where  $A$  is a positive definite matrix and where we have introduced the parameter  $\varepsilon > 0$  as a prefactor for the interaction (referred to as the coupling constant). Then formally we may apply Taylor expansion for  $e^{-\varepsilon U(x)}$  to obtain a series expansion for  $Z$ , as in

$$Z = \int_{\mathbb{R}^d} \sum_{n=0}^{\infty} \frac{\varepsilon^n}{n!} (-U(x))^n e^{-\frac{1}{2}x^T A x} dx \sim \sum_{n=0}^{\infty} \frac{\varepsilon^n}{n!} \int_{\mathbb{R}^d} (-U(x))^n e^{-\frac{1}{2}x^T A x} dx, \quad (1.5)$$

where the ‘ $\sim$ ’ is meant to indicate that the series is valid only in the asymptotic sense.

Assuming  $U(x)$  is a polynomial of  $x$ , then each term on the right hand side of Eq. (1.5) requires the evaluation of a possibly large, but finite, number of moments of a Gaussian distribution. The expansion can be organized in terms of Feynman diagrams.

Feynman diagrammatic expansions can also be obtained for  $G$  and  $\Sigma$ . In particular, recall from Part II that the self-energy diagrams are truncated, one-particle irreducible Feynman diagrams. To be concrete, one can keep in mind the quartic interaction

$$U(x) = \frac{1}{8} \sum_{i,j=1}^N v_{ij} x_i^2 x_j^2, \quad (1.6)$$

which mimics the two-body Coulomb interaction of quantum many-body physics. Here  $v$  is a symmetric positive definite matrix. In order to specify an impurity problem with fragment specified by indices  $1, \dots, p$  we take  $v_{ij} = 0$  if  $i > p$  or  $j > p$ . Then, it can readily be read from the diagrammatic expansion of  $\Sigma$  as in Part II that for each term in the expansion of  $\Sigma_{ij}$ , the corresponding matrix element is nonzero only if  $1 \leq i, j \leq p$ . This observation suggests that the self-energy matrix  $\Sigma$ , as the infinite sum of all of these terms, should follow the same sparsity pattern. We remark that the above diagrammatic argument can be applied to Gibbs models with rather general interaction form  $U(x)$ , as well as in the quantum many-body setting, where the diagrammatic series can be derived directly in the second-quantized representation or via the coherent state path integral.

The major caveat to this argument is that the Feynman diagrammatic expansion often has zero radius of convergence and maintains validity only in the asymptotic sense. This is the case at least for the Gibbs models as well as bosonic systems. Hence the sparsity for each term of the expansion does not necessarily imply that the same

is true of the self-energy itself when  $\varepsilon$  is positive. Even when the series does converge (such as for fermionic systems with finitely many states), the convergence radius may only be finite. Bootstrapping a positive radius of convergence via resummation or analytic continuation arguments [77] is one possible route to proving the sparsity result in such a setting, though the details seem to be cumbersome and the proof is not as simple or general as others considered above.

Finally, we discuss a route to the sparsity of the self-energy matrix via the so-called Luttinger-Ward formalism [65], which expresses the self-energy as a functional derivative

$$\Sigma = \frac{\delta\Phi[G]}{\delta G}. \quad (1.7)$$

Here  $\Phi[G]$  is a functional of the Green's function, called the Luttinger-Ward functional. We proved in Part III that for the Gibbs model,  $\Phi[G]$  is a well-defined for positive semidefinite  $G$ . In particular, we established a *projection rule*, which states that for the classical impurity problem when  $U(x) = U_p(x_1, \dots, x_p)$ , we have

$$\Phi[G] = \Phi_p[G_p]. \quad (1.8)$$

Here  $G_p$  is the upper-left  $p \times p$  block of  $G$ , and  $\Phi_p$  is the Luttinger-Ward functional for the  $p$ -dimensional model. Combining Eq. (1.7) and (1.8), one immediately obtains the sparsity pattern for  $\Sigma$ .

However, establishing the existence of the Luttinger-Ward functional  $\Phi[G]$  and its projection rule required a significant amount of work, and the rigorous proof is so far only applicable to the Gibbs model. In fact, the very existence of the Luttinger-Ward functional fermionic systems has been challenged over the past few years [54, 32, 42]. Although the Luttinger-Ward perspective offers additional insight, the direct proofs provided in this Part are at this point more generally applicable, and certainly much simpler.

## 1.2 Outline

This Part is organized as follows. We use the classical impurity problem as a motivating example and prove Theorem 1 in section 2. Section 3 treats the quantum many-body case, including the settings of fermions and bosons in the equilibrium setting at zero and finite temperature, as well as the non-equilibrium setting specified by an arbitrary contour in the complex plane and the anomalous setting relevant to superconductivity.

Note that in Appendix G we discuss the technical conditions needed to define the appropriate objects in the bosonic non-equilibrium setting and provide some background on main non-equilibrium setting of interest, specified by the Kadanoff-Baym contour.

## 2 The classical impurity problem

We now embark upon the proof of Theorem 1 stated above.

Recall the definitions:

$$Z = \int_{\mathbb{R}^n} e^{-\frac{1}{2}x^T Ax - U(x)} dx, \quad G = \frac{1}{Z} \int_{\mathbb{R}^n} xx^T e^{-\frac{1}{2}x^T Ax - U(x)} dx,$$

where the interaction  $U$  only depends on the first  $p \leq d$  variables. Let  $q = d - p$ . It is not hard to see that  $G$  is positive definite, hence invertible.

We will indicate the blocks of  $A$  via

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix},$$

where the upper-left block is  $p \times p$ . For various integrals considered below to be convergent, we will require that  $A_{22} \succ 0$ . More generally, we adopt the notation that for any  $d \times d$  matrix  $M$ , the notation  $M_{21}$  indicates the lower-left block of  $M$  (with respect to the above block structure), etc.

Then for the theorem, we want to show that the self-energy  $\Sigma := A - G^{-1}$  satisfies  $\Sigma_{12} = 0$ ,  $\Sigma_{21} = 0$ , and  $\Sigma_{22} = 0$ . In other words, we want to show that  $(G^{-1})_{12} = A_{12}$ ,  $(G^{-1})_{21} = A_{21}$ , and  $(G^{-1})_{22} = A_{22}$ . Since  $G$  and  $A$  are symmetric, it suffices to show that  $(G^{-1})_{12} = A_{12}$  and  $(G^{-1})_{22} = A_{22}$ , i.e., that

$$\begin{pmatrix} (G^{-1})_{12} \\ (G^{-1})_{22} \end{pmatrix} = \begin{pmatrix} A_{12} \\ A_{22} \end{pmatrix}.$$

Left-multiplying both sides by  $G$  (invertible), we see that this is in turn equivalent to showing that  $(GA)_{12} = 0_{p \times q}$  and  $(GA)_{22} = I_q$ .

In the following our notation will make use of the splitting

$$x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix},$$

where  $x \in \mathbb{R}^d$ ,  $x_1 \in \mathbb{R}^p$ , and  $x_2 \in \mathbb{R}^q$ . (For notational convenience, we *do not* use the notation  $\mathbf{x}_i$  as in the introduction. In this section, we will make no reference to the individual entries of  $x$ , so the notation is clear.) Then we can write  $U(x) = U_1(x_1)$ . Abusing notation slightly, we write  $U_1 = U$ .

Roughly speaking, the goal is to ‘integrate out’ the lower variables (i.e., the last  $q$  variables). To this end, we expand  $G$  as

$$G = \frac{1}{Z} \int_{\mathbb{R}^p} e^{-U(x_1)} \int_{\mathbb{R}^q} xx^T \exp \left[ -\frac{1}{2} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}^T \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \right] dx_2 dx_1.$$



Observe that

$$\begin{aligned}
& \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}^T \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \\
&= (x_2 + A_{22}^{-1}A_{21}x_1)^T A_{22} (x_2 + A_{22}^{-1}A_{21}x_1) + x_1^T (A_{11} - A_{12}A_{22}^{-1}A_{21})x_1, \\
&= (x_2 + A_{22}^{-1}A_{21}x_1)^T A_{22} (x_2 + A_{22}^{-1}A_{21}x_1) + x_1^T A_{11}^S x_1,
\end{aligned} \tag{2.1}$$

where  $A_{22}^{-1}$  is understood always to indicate  $(A_{22})^{-1}$  and where we have defined the Schur complement

$$A_{11}^S := A_{11} - A_{12}A_{22}^{-1}A_{21}.$$

Then it follows that

$$\begin{aligned}
G &= \frac{1}{Z} \int_{\mathbb{R}^p} e^{-\frac{1}{2}x_1^T A_{11}^S x_1 - U(x_1)} \times \\
&\quad \int_{\mathbb{R}^q} xx^T \exp \left[ -\frac{1}{2} (x_2 + A_{22}^{-1}A_{21}x_1)^T A_{22} (x_2 + A_{22}^{-1}A_{21}x_1) \right] dx_2 dx_1.
\end{aligned} \tag{2.2}$$

Recall that we want to show that  $(GA)_{12} = 0$  and  $(GA)_{22} = I_q$ . Right-multiplying the integral in (2.2) by  $A$ , this motivates computing the upper-right and upper-left blocks of  $xx^T A$ , as in

$$(xx^T A)_{12} = x_1 (x_1^T \ x_2^T) \begin{pmatrix} A_{12} \\ A_{22} \end{pmatrix}, \quad (xx^T A)_{22} = x_2 (x_1^T \ x_2^T) \begin{pmatrix} A_{12} \\ A_{22} \end{pmatrix}.$$

Now

$$(x_1^T \ x_2^T) \begin{pmatrix} A_{12} \\ A_{22} \end{pmatrix} = x_1^T A_{12} + x_2^T A_{22} = (x_1^T A_{12} A_{22}^{-1} + x_2^T) A_{22} = y_2^T A_{22},$$

where we have defined a new variable  $y_2 = x_2 + A_{22}^{-1}A_{21}x_1$ , so

$$(xx^T A)_{12} = x_1 y_2^T A_{22}, \quad (xx^T A)_{22} = x_2 y_2^T A_{22}. \tag{2.3}$$

The remarkable thing is that  $x_2$  only appears in the exponent in the inner integrand of (2.2) via the expression  $x_2 + A_{22}^{-1}A_{21}x_1 = y_2$ . This motivates us to eliminate  $x_2$  from the second equation of (2.3) to obtain

$$(xx^T A)_{12} = x_1 y_2^T A_{22}, \quad (xx^T A)_{22} = y_2 y_2^T A_{22} - A_{22}^{-1}A_{21}x_1 y_2^T A_{22}. \tag{2.4}$$

Then consider the change of variables from  $x_1, x_2$  to  $x_1, y_2$ , yielded by the linear transformation

$$\begin{pmatrix} x_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} I_p & 0 \\ A_{22}^{-1}A_{21} & I_q \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

Since the Jacobian determinant of this transformation is one, it follows from (2.2) and (2.4) that

$$(GA)_{12} = \frac{1}{Z} \int_{\mathbb{R}^p} e^{-\frac{1}{2}x_1^T A_{11}^S x_1 - U(x_1)} x_1 \left( \int_{\mathbb{R}^q} y_2^T e^{-\frac{1}{2}y_2^T A_{22}y_2} dy_2 \right) A_{22} dx_1.$$

But evidently the inner integrand is zero, so  $(GA)_{12} = 0$ , as desired. It also follows from (2.2) and (2.4) that

$$\begin{aligned} (GA)_{22} &= \frac{1}{Z} \int_{\mathbb{R}^p} e^{-\frac{1}{2}x_1^T A_{11}^S x_1 - U(x_1)} \left( \int_{\mathbb{R}^q} y_2 y_2^T e^{-\frac{1}{2}y_2^T A_{22}y_2} dy_2 \right) A_{22} dx_1 \\ &\quad - \frac{1}{Z} \int_{\mathbb{R}^p} e^{-\frac{1}{2}x_1^T A_{11}^S x_1 - U(x_1)} A_{22}^{-1} A_{21} x_1 \left( \int_{\mathbb{R}^q} y_2^T e^{-\frac{1}{2}y_2^T A_{22}y_2} dy_2 \right) A_{22} dx_1. \end{aligned}$$

The inner integrand in the second term of the last expression is once again zero. Meanwhile, the inner integrand of the first term yields  $Z_2 A_{22}^{-1}$ , where

$$Z_2 := \int_{\mathbb{R}^q} e^{-\frac{1}{2}y_2^T A_{22}y_2} dy_2.$$

Then we have established

$$(GA)_{22} = \frac{I_p}{Z} \int_{\mathbb{R}^p} \int_{\mathbb{R}^q} e^{-\frac{1}{2}x_1^T A_{11}^S x_1 - \frac{1}{2}y_2^T A_{22}y_2 - U(x_1)} dy_2 dx_1.$$

Changing variables back to  $x_1, x_2$  and recalling from (2.1) that  $x_1^T A_{11}^S x_1 + y_2^T A_{22}y_2 = x^T A x$ , we see that

$$(GA)_{22} = \frac{I_p}{Z} \int_{\mathbb{R}^d} e^{-\frac{1}{2}x^T A x - U(x)} dx = I_p,$$

which completes the proof.  $\square$

### 3 The quantum impurity problem

Our setting then is the Fock space  $\mathcal{F}_{\zeta,d}$  of fermions ( $\zeta = -1$ ) or bosons ( $\zeta = +1$ ) with a finite number  $d$  of states. The annihilation and creation operators are denoted  $a_1, \dots, a_d$  and  $a_1^\dagger, \dots, a_d^\dagger$ , respectively. We refer the reader to section 4 of Part I for further details of the construction of  $\mathcal{F}_{\zeta,d}$  as well as other details of second quantization. For convenience we shall let  $a = (a_1, \dots, a_d)^T$  denote the vector of annihilation operators, and accordingly  $a^\dagger = (a_1^\dagger, \dots, a_d^\dagger)$ .

For now<sup>10</sup> we consider a particle-number-conserving<sup>11</sup> self-adjoint Hamiltonian  $\hat{H}$  on  $\mathcal{F}_{\zeta,d}$ , and we write  $\hat{H}$  of the form

$$\hat{H} = \hat{H}_0 + \hat{U},$$

<sup>10</sup>In sections 3.3 and 3.4 below, the notion of the Hamiltonian will be somewhat modified.

<sup>11</sup>See section 4 of Part I for details.

where

$$\hat{H}_0 := a^\dagger h a = \sum_{i,j=1}^d h_{ij} a_i^\dagger a_j$$

is the *single-particle* (or *non-interacting*) part of the Hamiltonian, specified by a Hermitian  $d \times d$  matrix  $h$ , and  $\hat{U}$  is the *interacting* part, which is itself a self-adjoint operator on  $\mathcal{F}_{\zeta,d}$  that conserves particle number.

In the case that  $\hat{U}$  can be written as a polynomial of the  $a_i^\dagger, a_i$  for  $i = 1, \dots, p$ , we say that  $\hat{H}$  is an *impurity Hamiltonian*, with a *fragment* specified by the indices  $1, \dots, p$ . The rest of the indices correspond to the *environment*. In this case, since  $\hat{U}$  conserves particle number, it follows that  $\hat{U}$  commutes with  $a_j$  and  $a_j^\dagger$  for  $j > p$ .

Before proceeding, we state and prove a simple but useful lemma that will be used repeatedly throughout the following discussion.

**Lemma 2.**  $[a^\dagger h a, a_j^\dagger] = \sum_{k=1}^d h_{kj} a_k^\dagger$  and  $[a_j, a^\dagger h a] = \sum_{l=1}^d h_{jl} a_l$ .

*Proof.* Simply compute

$$\begin{aligned} (a^\dagger h a) a_j^\dagger &= \sum_{k,l=1}^d h_{kl} a_k^\dagger a_l a_j^\dagger \\ &= \sum_{k,l=1}^d h_{kl} a_k^\dagger (\zeta a_j^\dagger a_l + \delta_{jl}) \\ &= \sum_{k,l=1}^d h_{kl} a_j^\dagger a_k^\dagger a_l + \sum_{k=1}^d h_{kj} a_k^\dagger \\ &= a_j^\dagger (a^\dagger h a) + \sum_{k=1}^d h_{kj} a_k^\dagger, \end{aligned}$$

which proves the first statement of the lemma. Similarly,

$$\begin{aligned} (a^\dagger h a) a_j &= \sum_{k,l=1}^d h_{kl} \zeta a_k^\dagger a_j a_l \\ &= \sum_{k,l=1}^d h_{kl} (a_j a_k^\dagger - \delta_{jk}) a_l \\ &= a_j (a^\dagger h a) - \sum_{l=1}^d h_{jl} a_l \end{aligned}$$

which proves the lemma. □

### 3.1 Zero temperature

We consider the setting of zero temperature and fixed particle number  $N$ . Let  $|\Psi_0^{(N)}\rangle$  denote a normalized  $N$ -particle ground state of  $\hat{H}$ , and let the corresponding eigenvalue be  $E_0^{(N)}$ . Then in this setting, the single-particle Green's function can be understood as a rational function  $G : \mathbb{C} \rightarrow \mathbb{C}^{d \times d}$  defined by  $G(z) = G^+(z) + G^-(z)$ , where  $G^\pm$  are themselves rational functions<sup>12</sup> defined by

$$G_{ij}^+(z) := \langle \Psi_0^{(N)} | a_i \frac{1}{z - (\hat{H} - E_0^{(N)})} a_j^\dagger | \Psi_0^{(N)} \rangle$$

$$G_{ij}^-(z) := -\zeta \langle \Psi_0^{(N)} | a_j^\dagger \frac{1}{z + (\hat{H} - E_0^{(N)})} a_i | \Psi_0^{(N)} \rangle.$$

The self-energy is the rational function  $\Sigma : \mathbb{C} \rightarrow \mathbb{C}^{d \times d}$  defined by

$$\Sigma(z) := z - h - G(z)^{-1}.$$

As we recover in Theorem 3,  $z - h$  is in fact the inverse of the non-interacting Green's function, so this self-energy is defined analogously to the classical self-energy of Theorem 1. The reader should consult Appendix 5.1 for further details and justification of these definitions.

**Theorem 3.** *Suppose that  $\hat{H}$  is an impurity Hamiltonian, with a fragment specified by the indices  $1, \dots, p$ . Then the self-energy  $\Sigma : \mathbb{C} \rightarrow \mathbb{C}^{d \times d}$  is (up to the resolution of removable discontinuities) of the form*

$$\Sigma(z) = \begin{pmatrix} \Sigma_p(z) & 0 \\ 0 & 0 \end{pmatrix}.$$

*Remark 4.* Observe that if the fragment is of size zero, i.e.,  $p = 0$ , then we are in the non-interacting setting, and Theorem 3 implies that  $\Sigma(z) \equiv 0$ , i.e., that  $G(z) = (z - h)^{-1}$ . Thus we recover a clean proof of the formula for the non-interacting Green's function. Usually this formula is proved by assuming, via a canonical transformation, that  $h$  is diagonal and then performing explicit computations [33].

*Proof.* We can write  $\hat{H} = \hat{H}_0 + \hat{U}$ , where  $\hat{H}_0 = a^\dagger h a$  and  $\hat{U}$  commutes with  $a_j$  and  $a_j^\dagger$  for  $j > p$  is the *interacting* part, which is itself a self-adjoint operator on  $\mathcal{F}_{\zeta, d}$  that conserves particle number.

It suffices to prove that the  $j$ -th column of  $G(z)\Sigma(z)$  is zero for  $j > p$  and that the  $i$ -th row of  $\Sigma(z)G(z)$  is zero for  $i > p$ . We will only prove the first claim; the second follows by symmetric reasoning.

---

<sup>12</sup>Usually  $G^\pm$  carry the extra information that their poles are viewed as being located infinitesimally below/above the real axis. The choices that yield the 'time-ordered' Green's function are described in Appendix 7.2. However, this extra information is irrelevant for the purpose of our results.

Now  $G(z)\Sigma(z) = zG(z) - G(z)h - I_d$ , so we want to show that  $zG_{ij}(z) = [G(z)h]_{ij} + \delta_{ij}$  for  $j > p$ .

Then we compute, using the fact that  $(\hat{H} - E_0^{(N)})|\Psi_0^{(N)}\rangle = 0$ ,

$$\begin{aligned} zG_{ij}^+(z) &= \langle \Psi_0^{(N)} | a_i \frac{1}{z - (\hat{H} - E_0^{(N)})} a_j^\dagger (z - (\hat{H} - E_0^{(N)})) | \Psi_0^{(N)} \rangle \\ &= \langle \Psi_0^{(N)} | a_i \frac{1}{z - (\hat{H} - E_0^{(N)})} (z - (\hat{H} - E_0^{(N)})) a_j^\dagger | \Psi_0^{(N)} \rangle \\ &\quad + \langle \Psi_0^{(N)} | a_i \frac{1}{z - (\hat{H} - E_0^{(N)})} [a_j^\dagger, z - (\hat{H} - E_0^{(N)})] | \Psi_0^{(N)} \rangle \\ &= \langle \Psi_0^{(N)} | a_i a_j^\dagger | \Psi_0^{(N)} \rangle + \langle \Psi_0^{(N)} | a_i \frac{1}{z - (\hat{H} - E_0^{(N)})} [a_j^\dagger, z - (\hat{H} - E_0^{(N)})] | \Psi_0^{(N)} \rangle. \end{aligned}$$

Now

$$[a_j^\dagger, z - (\hat{H} - E_0^{(N)})] = [\hat{H}, a_j^\dagger] = [a^\dagger h a, a_j^\dagger] + [\hat{U}, a_j^\dagger] = \sum_{k=1}^d h_{kj} a_k^\dagger,$$

where we have used Lemma 2 as well as the fact that  $j > p$  (so  $[\hat{U}, a_j^\dagger] = 0$ ).

Then it follows that

$$zG_{ij}^+(z) = \langle \Psi_0^{(N)} | a_i a_j^\dagger | \Psi_0^{(N)} \rangle + [G^+(z)h]_{ij}.$$

Similarly, we compute

$$\begin{aligned} zG_{ij}^-(z) &= -\zeta \langle \Psi_0^{(N)} | (z + (\hat{H} - E_0^{(N)})) a_j^\dagger \frac{1}{z + (\hat{H} - E_0^{(N)})} a_i | \Psi_0^{(N)} \rangle \\ &= -\zeta \langle \Psi_0^{(N)} | a_j^\dagger (z + (\hat{H} - E_0^{(N)})) \frac{1}{z + (\hat{H} - E_0^{(N)})} a_i | \Psi_0^{(N)} \rangle \\ &\quad + (-\zeta) \langle \Psi_0^{(N)} | [z + (\hat{H} - E_0^{(N)}), a_j^\dagger] \frac{1}{z + (\hat{H} - E_0^{(N)})} a_i | \Psi_0^{(N)} \rangle \\ &= -\zeta \langle \Psi_0^{(N)} | a_j^\dagger a_i | \Psi_0^{(N)} \rangle \\ &\quad - \zeta \langle \Psi_0^{(N)} | [z + (\hat{H} - E_0^{(N)}), a_j^\dagger] \frac{1}{z + (\hat{H} - E_0^{(N)})} a_i | \Psi_0^{(N)} \rangle. \end{aligned}$$

Now

$$[z + (\hat{H} - E_0^{(N)}), a_j^\dagger] = [\hat{H}, a_j^\dagger] = \sum_{k=1}^d h_{kj} a_k^\dagger.$$

Then it follows that

$$zG_{ij}^-(z) = -\zeta \langle \Psi_0^{(N)} | a_i a_j^\dagger | \Psi_0^{(N)} \rangle + [G^-(z)h]_{ij}.$$

Therefore

$$zG_{ij}(z) = [G(z)h]_{ij} + \langle \Psi_0^{(N)} | a_i a_j^\dagger - \zeta a_j^\dagger a_i | \Psi_0^{(N)} \rangle = [G(z)h]_{ij} + \delta_{ij},$$

as was to be shown.  $\square$

### 3.2 Finite temperature

Now we consider the setting of finite inverse temperature  $\beta \in (0, \infty)$  and chemical potential  $\mu \in \text{int dom } Z$ , where  $Z(\mu) = \text{Tr}[e^{-\beta(\hat{H}-\mu\hat{N})}]$  (see Appendix 5.2 for further details). Note that  $\text{int dom } Z$  is guaranteed to be non-empty under Assumption 4.

We also let  $|\Psi_m\rangle$  denote the normalized eigenstates of  $\hat{H}$ , where  $m$  ranges from 0 to  $2^d - 1$  in the case of fermions and from 0 to  $\infty$  in the case of bosons. In this setting, the single-particle Green's function can be understood as a rational function  $G : \mathbb{C} \rightarrow \mathbb{C}^{d \times d}$  defined by  $G(z) = G^+(z) + G^-(z)$ , where  $G^\pm$  are themselves rational functions<sup>13</sup> defined by

$$G_{ij}^+(z) := \frac{1}{Z} \sum_m e^{-\beta(E_m - \mu N_m)} \langle \Psi_m | a_i \frac{1}{z - (\hat{H} - E_m)} a_j^\dagger | \Psi_m \rangle$$

$$G_{ij}^-(z) := \frac{-\zeta}{Z} \sum_m e^{-\beta(E_m - \mu N_m)} \langle \Psi_m | a_j^\dagger \frac{1}{z + (\hat{H} - E_m)} a_i | \Psi_m \rangle,$$

and these sums are absolutely convergent away from the poles. Here

$$Z = \text{Tr}[e^{-\beta(\hat{H}-\mu\hat{N})}] = \sum_m e^{-\beta(E_m - \mu N_m)}.$$

Once again the self-energy is the rational function  $\Sigma : \mathbb{C} \rightarrow \mathbb{C}^{d \times d}$  defined by

$$\Sigma(z) := z - h - G(z)^{-1}.$$

The reader should consult section 7 of Part I for further details and justification of these definitions.

**Theorem 5.** *Suppose that  $\hat{H}$  is an impurity Hamiltonian, with a fragment specified by the indices  $1, \dots, p$ . Then the self-energy  $\Sigma : \mathbb{C} \rightarrow \mathbb{C}^{d \times d}$  is (up to the resolution of removable discontinuities) of the form*

$$\Sigma(z) = \begin{pmatrix} \Sigma_p(z) & 0 \\ 0 & 0 \end{pmatrix}.$$

*Remark 6.* Once again (cf. Remark 4), we recover in the non-interacting setting the formula  $G(z) = (z - h)^{-1}$ .

*Remark 7.* There is a further object known as the *Matsubara Green's function* [77], which in turn yields the Matsubara self-energy. Although it is not usually defined this way, the Matsubara Green's function can be shown to be obtained from the finite-temperature Green's function, as defined above, by restriction to points  $i\omega_m + \mu$ , where  $\omega_m$  are the fermionic/bosonic *Matsubara frequencies* [77]. The Matsubara self-energy can be obtained from the finite-temperature self-energy defined above via similar restriction. Therefore Theorem 5 implies the same sparsity pattern for the Matsubara self-energy.

<sup>13</sup>The same comments as in section 3.1 apply here as well, though instead see Appendix 7.3 for details relevant to this setting.

*Proof.* The proof is essentially the same as that of Theorem 3. Once again we want to show that the  $j$ -th column of  $G(z)\Sigma(z)$  is zero for  $j > p$  and that the  $i$ -th row of  $\Sigma(z)G(z)$  is zero for  $i > p$ . We will only prove the first claim; the second follows by symmetric reasoning.

Define  $G_m(z)$  by

$$G_{m,ij}(z) := \langle \Psi_m | a_i \frac{1}{z - (\hat{H} - E_m)} a_j^\dagger | \Psi_m \rangle - \zeta \langle \Psi_m | a_j^\dagger \frac{1}{z + (\hat{H} - E_m)} a_i | \Psi_m \rangle.$$

Then by the same reasoning as in the proof of Theorem 3 (with the roles of  $E_0^{(N)}$  and  $|\Psi_0^{(N)}\rangle$  played by  $E_m$  and  $|\Psi_m\rangle$ ), we find that

$$zG_{m,ij}(z) = [G_m(z)h]_{ij} + \delta_{ij}.$$

Now  $G(z) = \frac{1}{Z} \sum_m e^{-\beta(E_m - \mu N_m)} G_{m,ij}(z)$ , so the desired result follows.  $\square$

### 3.3 Arbitrary contour

There is a more general perspective in which the time-ordering operation used in Appendices 5.1 and 5.2 to derive the Green's functions considered above is generalized to an ordering operation on an arbitrary contour in the complex plane. This perspective adds significant value in the non-equilibrium setting, in which one considers a time-dependent Hamiltonian. For such time-dependent problems, passage to the frequency representation is not possible. Instead we consider kernels on the contour.

Let  $\mathcal{C}$  denote a piecewise smooth contour in the complex plane (not necessarily closed). Technically one should think of  $\mathcal{C}$  not as a subset of  $\mathbb{C}$ , but as a parametrized path,  $\gamma : I \rightarrow \mathbb{C}$ , where  $I = (s_0, s_1)$  is some interval. Then for  $s, s' \in I$  with  $s < s'$ , we define  $\mathcal{C}(s, s')$  to be the 'sub-contour' defined by restriction of  $\gamma$  to the interval  $(s, s')$ . If  $s > s'$ , we define  $\mathcal{C}(s, s')$  to be the contour obtained from  $\mathcal{C}(s', s)$  by reversing its orientation.

Additionally let  $\hat{H}(z)$  denote an operator-valued function on a neighborhood of  $\mathcal{C} = \gamma(I)$ . Here  $\hat{H}(z) = a^\dagger h(z) a + \hat{U}(z)$  is particle-number-conserving, and we say that  $\hat{H}(z)$  is an impurity Hamiltonian with a fragment specified by indices  $1, \dots, p$  if, for every  $z \in \mathcal{C}$ ,  $\hat{U}(z)$  can be written as a polynomial of the  $a_i^\dagger, a_i$  for  $i = 1, \dots, p$ . As above, since  $\hat{U}(z)$  must conserve particle number, it follows that  $\hat{U}(z)$  commutes with  $a_j$  and  $a_j^\dagger$  for  $j > p$ . It is convenient to denote  $z(s) := \gamma(s)$ , and abusing notation slightly we will write  $\hat{H}(s) = \hat{H}(z(s))$ .

As a technical point, we assume that  $\hat{H}(s)$  is piecewise continuous. Since the Fock space is finite dimensional in the case of fermions, the meaning of this statement is unambiguous. In the case of bosons, note that since  $\hat{H}(s)$  is particle-number-conserving, we can sensibly consider its restriction to each of the  $N$ -particle subspaces (see Appendix 4), each of which is finite-dimensional. Then by the continuity of  $\hat{H}(s)$  we mean the continuity of all of these restrictions individually.

Now define a (not necessarily unitary) evolution operator from contour time  $s' \in I$  to  $s \in I$  as the time-ordered exponential

$$U(s, s') = \mathcal{T} \left\{ e^{-i \int_{c(s, s')} \hat{H}(z) dz} \right\}.$$

This simply means that  $U(s, s')$  is taken as the solution of the differential equation

$$\partial_s U(s, s') = -i \dot{z}(s) \hat{H}(s) U(s, s'), \quad U(s', s') = \text{Id}. \quad (3.1)$$

This initial-value problem indeed admits a unique solution in the bosonic case because the ODE can be viewed as describing the evolution of an operator on each of the (finite-dimensional)  $N$ -particle subspace separately.

From this definition it follows that

$$U(s, s'') U(s'', s') = U(s, s')$$

for all  $s, s', s'' \in I$  and moreover that

$$\partial_{s'} U(s, s') = i \dot{z}(s) U(s, s') \hat{H}(s'). \quad (3.2)$$

Abusing notation slightly by pretending that we can invert  $s = s(z)$ , we can more cleanly write

$$\partial_z U(z, z') = -i \hat{H}(z) U(z, z'), \quad \partial_{z'} U(z, z') = i U(z, z') \hat{H}(z'),$$

where  $\partial_z = (\dot{z}(s))^{-1} \partial_s$ . We will sometimes adopt this notational convention, and the meaning should be clear from context.

The following assumption is adopted to ensure that the Green's function can be defined in the bosonic case:

**Assumption 8.** *We assume that for all  $s > s'$ ,  $U(s, s')$  is a bounded operator. Moreover, we assume that there exists  $s > s'$  such that the operator norm of the restriction of  $U(s, s')$  to the  $N$ -particle subspace decays exponentially in  $N$ .*

Define the partition function

$$Z = \text{Tr}[U(s_1, s_0)].$$

Note that Assumption 8 guarantees that  $U(s_1, s_0)$  is trace class, so  $Z$  is indeed well-defined. In order to define our ensemble, we must be able to divide by  $Z$ . Hence we assume:

**Assumption 9.**  $Z \neq 0$ .



We show in Appendix G how Assumptions 8 and 9 are naturally satisfied in the major non-equilibrium setting of interest, which features the Kadanoff-Baym contour.

Then we define ‘pseudo-Heisenberg’ representations of the annihilation and creation operators via

$$a_i(s) = U(s_0, s)a_iU(s, s_0), \quad a_i^\dagger(s) = U(s_0, s)a_i^\dagger U(s, s_0).$$

The *contour-ordered, single-body Green’s function* (which we call the Green’s function for short when the context is clear) is a function  $G : I \times I \rightarrow \mathbb{C}^{d \times d}$  defined by

$$G_{ij}(s, s') = \frac{-i}{Z} \text{Tr} [\mathcal{T} \{ a_i(s) a_j^\dagger(s') \} U(s_1, s_0)],$$

where  $\mathcal{T}$  is the *contour-ordering operator*, formally defined by

$$\mathcal{T} \{ a_i(s) a_j^\dagger(s') \} = \begin{cases} a_i(s) a_j^\dagger(s'), & s' < s \\ \zeta a_j^\dagger(s') a_i(s), & s' \geq s. \end{cases}$$

In other words we can write  $G = G^+ + G^-$ , where

$$iG_{ij}^+(s, s') = \frac{1}{Z} \text{Tr} [U(s_1, s) a_i U(s, s') a_j^\dagger U(s', s_0)] \theta(s - s')$$

and

$$iG_{ij}^-(s, s') = \frac{\zeta}{Z} \text{Tr} [U(s_1, s') a_j^\dagger U(s', s) a_i U(s, s_0)] (1 - \theta(s - s')).$$

Here

$$\theta(s) := \begin{cases} 1, & s > 0 \\ 0, & s \leq 0. \end{cases}$$

In the bosonic case, Assumption 8 guarantees that the traces needed for this definition do indeed exist. For later reference, note that we can define a product of suitable functions  $A, B : I \times I \rightarrow \mathbb{C}^{d \times d}$  (with an appropriate notion of multiplicative inverse, at least formally) via

$$(AB)(s, s') = \int_{s_0}^{s_1} A(s, s'') B(s'', s') \dot{z}(s'') ds'',$$

chosen so that formally we have

$$(AB)(z, z') = \int_{z_0}^{z_1} A(z, z'') B(z'', z') dz''.$$

Notice that the appropriate identity  $\delta(z, z')$  is then given by  $\delta(z, z') = (\dot{z}(s))^{-1} \delta(s - s')$ .

We remark that the zero-temperature and Matsubara Green’s functions discussed in sections 3.1 and 7, respectively, can be recovered as contour-ordered Green’s functions. By contrast, the real-time Green’s function at finite temperature considered in

section 3.2 *cannot* be recovered directly as a contour-ordered Green's function, though it can be obtained indirectly via analytic continuation of the Matsubara Green's function. For this reason, diagrammatic expansion techniques at finite temperature are limited to the Matsubara Green's function and must be carried over to the real-time Green's function via analytic continuation. For further details, see [100].

One now wants to define the self-energy as

$$\Sigma(z, z') = i\partial_z - h(z) \delta(z, z') - G^{-1}(z, z').$$

However, this definition is not rigorous without further justification. Indeed, note that  $G$  can be viewed as an integral operator on  $L^2(I)$ , and under reasonable assumptions  $G$  is Hilbert-Schmidt, hence in particular compact. Therefore its inverse is guaranteed to be an unbounded operator, if it can be constructed. Formally, one expects that the  $i(\dot{z}(s))^{-1}\partial_s$  in our definition of the self-energy will cancel an analogous term in the formal inverse  $G^{-1}$  and that the self-energy can be written as a sum of a static and dynamic part as

$$\Sigma(s, s') = \Sigma_{\text{stat}} \delta(s - s') + \Sigma_{\text{dyn}}(s, s'),$$

where  $\Sigma_{\text{dyn}}$  is a properly defined integral operator.

In our view the mathematical construction of the self-energy seems to be a non-trivial matter, and we will sidestep it in this work. (By contrast, the construction in the equilibrium setting is more straightforward in the frequency domain; see Appendices 5.1 and 5.2.)

How then to discuss the sparsity pattern of the self-energy? Observe that formally, we should have

$$\begin{aligned} (\Sigma G)(z, z') &= i\partial_z G(z, z') - h(z)G(z, z') - I_d \delta(z, z') \\ (G\Sigma)(z, z') &= -i\partial_{z'} G(z, z') - G(z, z')h(z') - I_d \delta(z, z'), \end{aligned}$$

or, more rigorously,

$$\begin{aligned} (\Sigma G)(s, s') &= i(\dot{z}(s))^{-1}\partial_s G(s, s') - h(s)G(s, s') - I_d (\dot{z}(s))^{-1} \delta(s - s') \\ (G\Sigma)(s, s') &= -i(\dot{z}(s'))^{-1}\partial_{s'} G(s, s') - G(s, s')h(s') - I_d (\dot{z}(s))^{-1} \delta(s - s'). \end{aligned} \quad (3.3)$$

Now instead of constructing the self-energy, we can *define* operators  $\Sigma G$  and  $G\Sigma$  via (3.3) (in the sense of distributions), with the ' $\Sigma$ ' appearing here merely as a notation. Now the desired sparsity pattern of  $\Sigma$  is formally equivalent to the statement that  $[\Sigma G]_{ij} = 0$  (as a distribution on  $I$ ) for  $i > p$  and  $[G\Sigma]_{ij} = 0$  for  $j > p$ .

**Theorem 10.** *With notation and assumptions as in the preceding, if  $\hat{H}(z)$  is an impurity Hamiltonian with a fragment specified by the indices  $1, \dots, p$ , then  $[\Sigma G]_{ij} = 0$  for  $i > p$  and  $[G\Sigma]_{ij} = 0$  for  $j > p$ .*

*Remark 11.* In the non-interacting setting  $p = 0$ , we recover the formulas

$$i\partial_z G(z, z') - h(z) G(z, z') = I_d \delta(z, z'), \quad -i\partial_{z'} G(z, z') - G(z, z') h(z') = I_d \delta(z, z'),$$

where we have abused notation slightly in the manner described above. These formulas seem to be non-trivial to establish by any other means. By contrast with the equilibrium case, this formula cannot be established simply via a canonical transformation because it may not be possible to simultaneously diagonalize the  $h(z)$  for all  $z$ . In fact, in [100], the non-interacting Green's function is *defined* via this formula (subject to certain boundary conditions) and shown to give the appropriate perturbation theory within the Martin-Schwinger hierarchy.

*Proof.* We prove only the first statement, as the second follows from similar arguments. Recall

$$iG_{ij}^+(s, s') = \frac{1}{Z} \text{Tr} \left[ U(s_1, s) a_i U(s, s') a_j^\dagger U(s', s_0) \right] \theta(s - s').$$

Then compute, using Eqs. (3.1) and (3.2),

$$\begin{aligned} i(\dot{z}(s))^{-1} \partial_s G_{ij}^+(s, s') &= \frac{i}{Z} \text{Tr} \left[ U(s_1, s) \hat{H}(s) a_i U(s, s') a_j^\dagger U(s', s_0) \right] \theta(s - s') \\ &\quad - \frac{i}{Z} \text{Tr} \left[ U(s_1, s) a_i \hat{H}(s) U(s, s') a_j^\dagger U(s', s_0) \right] \theta(s - s') \\ &\quad + (\dot{z}(s))^{-1} \frac{1}{Z} \text{Tr} \left[ U(s_1, s) a_i a_j^\dagger U(s, s_0) \right] \delta(s - s') \\ &= \frac{-i}{Z} \text{Tr} \left[ U(s_1, s) [a_i, \hat{H}(s)] U(s, s') a_j^\dagger U(s', s_0) \right] \theta(s - s') \\ &\quad + (\dot{z}(s))^{-1} \frac{1}{Z} \text{Tr} \left[ U(s_1, s) a_i a_j^\dagger U(s, s_0) \right] \delta(s - s'). \end{aligned}$$

Now for  $i > p$ ,  $[a_i, \hat{U}(s)] = 0$ , so  $[a_i, \hat{H}(s)] = [a_i, a^\dagger h(s) a] = \sum_{l=1}^d h_{il}(s) a_l$ , by Lemma 2. Therefore

$$\begin{aligned} i(\dot{z}(s))^{-1} \partial_s G_{ij}^+(s, s') &= \frac{-i}{Z} \sum_{l=1}^d h_{il}(s) \text{Tr} \left[ U(s_1, s) a_l U(s, s') a_j^\dagger U(s', s_0) \right] \theta(s - s') \\ &\quad + (\dot{z}(s))^{-1} \frac{1}{Z} \text{Tr} \left[ U(s_1, s) a_i a_j^\dagger U(s, s_0) \right] \delta(s - s') \\ &= [h(s) G^+(s, s')]_{ij} \\ &\quad + (\dot{z}(s))^{-1} \frac{1}{Z} \text{Tr} \left[ U(s_1, s) a_i a_j^\dagger U(s, s_0) \right] \delta(s - s'). \end{aligned}$$

Similarly,

$$i(\dot{z}(s))^{-1} \partial_s G_{ij}^-(s, s') = \frac{i\zeta}{Z} \text{Tr} \left[ U(s_1, s') a_j^\dagger U(s', s) \hat{H}(s) a_i U(s, s_0) \right] (1 - \theta(s - s'))$$

$$\begin{aligned}
& - \frac{i\zeta}{Z} \left[ U(s_1, s') a_j^\dagger U(s', s) a_i \hat{H}(s) U(s, s_0) \right] (1 - \theta(s - s')) \\
& - (\dot{z}(s))^{-1} \frac{\zeta}{Z} \text{Tr} \left[ U(s_1, s) a_j^\dagger a_i U(s, s_0) \right] \delta(s - s') \\
= & \frac{-i\zeta}{Z} \text{Tr} \left[ U(s_1, s) [a_i, \hat{H}(s)] U(s, s') a_j^\dagger U(s', s_0) \right] \theta(s - s') \\
& - \frac{\zeta}{Z} \text{Tr} \left[ U(s_1, s) a_j^\dagger a_i U(s, s_0) \right] \delta(s - s') \\
= & \frac{-i\zeta}{Z} \sum_{l=1}^d h_{il}(s) \text{Tr} \left[ U(s_1, s) a_l U(s, s') a_j^\dagger U(s', s_0) \right] \theta(s - s') \\
& - (\dot{z}(s))^{-1} \frac{\zeta}{Z} \text{Tr} \left[ U(s_1, s) a_j^\dagger a_i U(s, s_0) \right] \delta(s - s') \\
= & [h(s)G^-(s, s')]_{ij} \\
& - (\dot{z}(s))^{-1} \frac{\zeta}{Z} \text{Tr} \left[ U(s_1, s) a_j^\dagger a_i U(s, s_0) \right] \delta(s - s').
\end{aligned}$$

Therefore, since  $G = G^+ + G^-$ , we have

$$\begin{aligned}
i(\dot{z}(s))^{-1} \partial_s G_{ij}(s, s') &= [h(s)G(s, s')]_{ij} \\
&+ \frac{1}{Z} \text{Tr} \left[ U(s_1, s) (a_i a_j^\dagger - \zeta a_j^\dagger a_i) U(s, s_0) \right] (\dot{z}(s))^{-1} \delta(s - s') \\
&= [h(s)G(s, s')]_{ij} + \delta_{ij} (\dot{z}(s))^{-1} \delta(s - s'),
\end{aligned}$$

which completes the proof.  $\square$

### 3.4 Anomalous setting

Finally we will consider a sparsity result for the self-energy of *anomalous* impurity problems. These are impurity problems in which the Hamiltonian does not conserve particle number. Since the anomalous setting is of most interest for the study of superconductivity in *fermions*, we will restrict our attention to the fermionic setting. This allows us to avoid some further analytic difficulty since our rigorous definitions in the bosonic case (in which the Fock space is infinite-dimensional) relied on particle number conservation. It also eases the notational burden to keep track of  $\zeta$  to distinguish the bosonic and fermionic systems. In order to simply illustrate the points that are novel to this setting, we further restrict our attention to the zero-temperature equilibrium setting.

Now consider a self-adjoint Hamiltonian  $\hat{H}$  on the fermionic Fock space  $\mathcal{F}_{-1,d}$ , and we write  $\hat{H}$  of the form

$$\hat{H} = \hat{H}_0 + \hat{U},$$

where

$$\hat{H}_0 := \hat{H}_{\text{NA}} + \hat{H}_{\text{A}} + \hat{H}_{\text{A}}^\dagger$$

is the single-particle part of the Hamiltonian (no longer particle-number-conserving), specified by its non-anomalous and anomalous parts

$$\hat{H}_{\text{NA}} := \sum_{i,j=1}^d h_{ij} a_i^\dagger a_j, \quad \hat{H}_{\text{A}} := \frac{1}{2} \sum_{i,j=1}^d \Delta_{ij} a_i^\dagger a_j^\dagger.$$

Therefore, up to a scalar multiple of the identity operator,  $\hat{H}_0$  is given by

$$\begin{pmatrix} a \\ a^\dagger \end{pmatrix}^\dagger \begin{pmatrix} h & \Delta \\ -\bar{\Delta} & -\bar{h} \end{pmatrix} \begin{pmatrix} a \\ a^\dagger \end{pmatrix},$$

where we have abused notation slightly by using  $a$  to indicate both a row and a column vector of operators. Without loss of generality we assume that  $\Delta = (\Delta_{ij})$  is a complex antisymmetric matrix. (Note that then  $-\bar{\Delta} = \Delta^\dagger$ , and since  $h$  is Hermitian,  $-\bar{h} = -h^T$ .) Meanwhile, the interacting part  $\hat{U}$  is itself a self-adjoint operator on  $\mathcal{F}_{-1,d}$ , and we demand that it can be written as an *even* polynomial of the creation and annihilation operators, which includes the particle-number-conserving  $\hat{U}$  as a sub-case. In the case that  $\hat{U}$  can be written as a polynomial of the  $a_i^\dagger, a_i$  for  $i = 1, \dots, p$ , we say that  $\hat{H}$  is an *anomalous impurity Hamiltonian*, with a fragment specified by the indices  $1, \dots, p$ . As in earlier settings, the rest of the indices correspond to the environment. Note that the evenness of the polynomial specifying  $\hat{U}$  guarantees that  $\hat{U}$  commutes with  $a_j$  and  $a_j^\dagger$  for  $j > p$ .

To determine the expectations computed at the end of the last section, it suffices to determine the following Green's functions, which are themselves desirable to know:

$$\begin{aligned} G_{ij}^{\text{hp}}(z) &:= G_{ij}^{\text{hp},+}(z) + G_{ij}^{\text{hp},-}(z) \\ &:= \langle \Phi_0 | a_i \frac{1}{z - (\hat{H} - E_0)} a_j^\dagger | \Phi_0 \rangle + \langle \Phi_0 | a_j^\dagger \frac{1}{z + (\hat{H} - E_0)} a_i | \Phi_0 \rangle \end{aligned}$$

$$\begin{aligned} G_{ij}^{\text{pp}}(z) &:= G_{ij}^{\text{pp},+}(z) + G_{ij}^{\text{pp},-}(z) \\ &:= \langle \Phi_0 | a_i^\dagger \frac{1}{z - (\hat{H} - E_0)} a_j^\dagger | \Phi_0 \rangle + \langle \Phi_0 | a_j^\dagger \frac{1}{z + (\hat{H} - E_0)} a_i^\dagger | \Phi_0 \rangle \end{aligned}$$

$$\begin{aligned} G_{ij}^{\text{hh}}(z) &:= G_{ij}^{\text{hh},+}(z) + G_{ij}^{\text{hh},-}(z) \\ &:= \langle \Phi_0 | a_i \frac{1}{z - (\hat{H} - E_0)} a_j | \Phi_0 \rangle + \langle \Phi_0 | a_j \frac{1}{z + (\hat{H} - E_0)} a_i | \Phi_0 \rangle \end{aligned}$$

$$G_{ij}^{\text{ph}}(z) := G_{ij}^{\text{ph},+}(z) + G_{ij}^{\text{ph},-}(z)$$

$$:= \langle \Phi_0 | a_i^\dagger \frac{1}{z - (\hat{H} - E_0)} a_j | \Phi_0 \rangle + \langle \Phi_0 | a_j \frac{1}{z + (\hat{H} - E_0)} a_i^\dagger | \Phi_0 \rangle,$$

where  $|\Phi_0\rangle$  is the ground state of  $\hat{H}$  and  $E_0$  is the ground-state energy. The superscripts  $p$  and  $h$  stands for ‘particle’ and ‘hole’, respectively [14], so  $G^{\text{hh}}$  is called the hole-hole Green’s function,  $G^{\text{ph}}$  is the particle-hole Green’s function, etc.

Notice that the last two Green’s functions are actually redundant because  $G^{\text{ph}}(z) = -[G^{\text{hp}}(-z)]^T$  and  $G^{\text{hh}}(z) = [G^{\text{pp}}(\bar{z})]^\dagger$ . We can further define the anomalous Green’s function by

$$\mathbf{G}(z) := \begin{pmatrix} G^{\text{hp}}(z) & G^{\text{hh}}(z) \\ G^{\text{pp}}(z) & G^{\text{ph}}(z) \end{pmatrix}$$

and the anomalous self-energy by

$$\Sigma(z) := z - \begin{pmatrix} h & \Delta \\ -\bar{\Delta} & -\bar{h} \end{pmatrix} - \mathbf{G}(z)^{-1}.$$

In fact we will show the following result:

**Theorem 12.** *Suppose that  $\hat{H}$  is an anomalous impurity Hamiltonian, with a fragment specified by the indices  $1, \dots, p$ . Then the anomalous self-energy  $\Sigma : \mathbb{C} \rightarrow \mathbb{C}^{d \times d}$  is (up to the resolution of removable discontinuities) of the form*

$$\Sigma(z) = \begin{pmatrix} \Sigma_p^{\text{hp}}(z) & 0 & \Sigma_p^{\text{hh}}(z) & 0 \\ 0 & 0 & 0 & 0 \\ \Sigma_p^{\text{pp}}(z) & 0 & \Sigma_p^{\text{ph}}(z) & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

*Remark 13.* Note that in the case  $p = 0$  we recover the formula

$$\mathbf{G}(z) = \left[ z - \begin{pmatrix} h & \Delta \\ -\bar{\Delta} & -\bar{h} \end{pmatrix} \right]^{-1}$$

for the non-interacting anomalous Green’s function. .

Recall from Lemma 2 that

$$[\hat{H}_{\text{NA}}, a_j^\dagger] = \sum_{\gamma} h_{kj} a_k^\dagger$$

and

$$[\hat{H}_{\text{NA}}, a_j] = - \sum_{\gamma} h_{jk} a_k^\dagger.$$

Before proceeding with the proof of Theorem 12, we supplement this result with a further simple lemma:

**Lemma 14.** Let  $\hat{H}_A = \frac{1}{2} \sum_{i,j=1}^d \Delta_{ij} a_i^\dagger a_j^\dagger$  with  $\Delta = (\Delta_{ij})$  being a complex antisymmetric matrix. Then

$$[\hat{H}_A, a_j^\dagger] = 0, \quad [\hat{H}_A^\dagger, a_j] = 0, \quad [\hat{H}_A, a_j] = \sum_k \Delta_{kj} a_k^\dagger, \quad [\hat{H}_A^\dagger, a_j^\dagger] = \sum_k \bar{\Delta}_{jk} a_k.$$

*Proof.* The first two identities are obvious, and the fourth follows from the third by taking Hermitian conjugates and using the antisymmetry of  $\Delta$ . To see the claimed third identity, simply compute

$$\begin{aligned} \hat{H}_A a_j &= \frac{1}{2} \sum_{ik} \Delta_{ik} a_i^\dagger a_k^\dagger a_j \\ &= \frac{1}{2} \sum_{ik} \Delta_{ik} a_i^\dagger \delta_{jk} - \frac{1}{2} \sum_{ik} \Delta_{ik} a_i^\dagger a_j a_k^\dagger \\ &= \frac{1}{2} \sum_i \Delta_{ij} a_i^\dagger - \frac{1}{2} \sum_{ik} \Delta_{ik} \delta_{ij} a_k^\dagger + \frac{1}{2} \sum_{ik} \Delta_{ik} a_j a_i^\dagger a_k^\dagger \\ &= \frac{1}{2} \sum_k \Delta_{kj} a_k^\dagger - \frac{1}{2} \sum_k \Delta_{jk} a_k^\dagger + a_j \hat{H}_A \\ &= \sum_k \Delta_{kj} a_k^\dagger + a_j \hat{H}_A. \end{aligned}$$

□

*Proof.* (Of Theorem 12.) Throughout we will often use  $\langle \cdot \rangle$  to indicate the expectation  $\langle \Phi_0 | \cdot | \Phi_0 \rangle$ .

Now it suffices to show the following sparsity pattern

$$\mathbf{G}(z)\boldsymbol{\Sigma}(z) = \begin{pmatrix} * & 0 & * & 0 \\ * & 0 & * & 0 \\ * & 0 & * & 0 \\ * & 0 & * & 0 \end{pmatrix}, \quad \boldsymbol{\Sigma}(z)\mathbf{G}(z) = \begin{pmatrix} * & * & * & * \\ 0 & 0 & 0 & 0 \\ * & * & * & * \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

We will only prove the first of these claims; the other follows by similar reasoning. Note that this first claim is equivalent to the fact that each of the following equalities holds *along the last  $d - p$  columns*:

$$\begin{aligned} G^{\text{hp}}[z - h] - G^{\text{hh}} \Delta^\dagger &= I_d, \\ -G^{\text{hp}} \Delta + G^{\text{hh}}[z + h^T] &= 0, \\ G^{\text{pp}}[z - h] - G^{\text{ph}} \Delta^\dagger &= 0, \\ -G^{\text{pp}} \Delta + G^{\text{ph}}[z + h^T] &= I_d. \end{aligned}$$

Now we begin the computations. In the following we assume that  $j > p$ . Since  $(\hat{H} - E_0)|\Phi_0\rangle = 0$ , we have

$$\begin{aligned}
zG_{ij}^{\text{hp},+}(z) &= \langle \Phi_0 | a_i \frac{1}{z - (\hat{H} - E_0)} a_j^\dagger (z - (\hat{H} - E_0)) | \Phi_0 \rangle \\
&= \langle a_i a_j^\dagger \rangle + \langle a_i \frac{1}{z - (\hat{H} - E_0)} [\hat{H}, a_j^\dagger] \rangle \\
&= \langle a_i a_j^\dagger \rangle + \langle a_i \frac{1}{z - (\hat{H} - E_0)} [\hat{H}_{\text{NA}}, a_j^\dagger] \rangle + \langle a_i \frac{1}{z - (\hat{H} - E_0)} [\hat{H}_A^\dagger, a_j^\dagger] \rangle \\
&= \langle a_i a_j^\dagger \rangle + \sum_k \langle a_i \frac{1}{z - (\hat{H} - E_0)} a_k^\dagger \rangle h_{kj} + \sum_k \langle a_i \frac{1}{z - (\hat{H} - E_0)} a_k \rangle \bar{\Delta}_{jk} \\
&= \langle a_i a_j^\dagger \rangle + [G^{\text{hp},+} h]_{ij} + [G^{\text{hh},+} \Delta^\dagger]_{ij}.
\end{aligned}$$

Similarly,

$$\begin{aligned}
zG_{ij}^{\text{hp},-}(z) &= \langle \Phi_0 | (z + (\hat{H} - E_0)) a_j^\dagger \frac{1}{z + (\hat{H} - E_0)} a_i | \Phi_0 \rangle \\
&= \langle a_j^\dagger a_i \rangle + \langle [\hat{H}, a_j^\dagger] \frac{1}{z + (\hat{H} - E_0)} a_i \rangle \\
&= \langle a_j^\dagger a_i \rangle + \langle [\hat{H}_{\text{NA}}, a_j^\dagger] \frac{1}{z + (\hat{H} - E_0)} a_i \rangle + \langle [\hat{H}_A^\dagger, a_j^\dagger] \frac{1}{z + (\hat{H} - E_0)} a_i \rangle \\
&= \langle a_j^\dagger a_i \rangle + \sum_k \langle a_k^\dagger \frac{1}{z + (\hat{H} - E_0)} a_i \rangle h_{kj} + \sum_k \langle a_k \frac{1}{z + (\hat{H} - E_0)} a_i \rangle \bar{\Delta}_{jk} \\
&= \langle a_j^\dagger a_i \rangle + [G^{\text{hp},-} h]_{ij} + [G^{\text{hh},-} \Delta^\dagger]_{ij}.
\end{aligned}$$

Therefore, adding our results and recognizing that  $\langle a_i a_j^\dagger \rangle + \langle a_j^\dagger a_i \rangle = \delta_{ij}$ , we obtain

$$zG_{ij}^{\text{hp}} = \delta_{ij} + [G^{\text{hp}} h]_{ij} + [G^{\text{hh}} \Delta^\dagger]_{ij}$$

for all  $j > p$ , which implies our first desired result.

Next compute

$$\begin{aligned}
zG_{ij}^{\text{hh},+}(z) &= \langle \Phi_0 | a_i \frac{1}{z - (\hat{H} - E_0)} a_j (z - (\hat{H} - E_0)) | \Phi_0 \rangle \\
&= \langle a_i a_j \rangle + \langle a_i \frac{1}{z - (\hat{H} - E_0)} [\hat{H}, a_j] \rangle \\
&= \langle a_i a_j \rangle + \langle a_i \frac{1}{z - (\hat{H} - E_0)} [\hat{H}_{\text{NA}}, a_j] \rangle + \langle a_i \frac{1}{z - (\hat{H} - E_0)} [\hat{H}_A, a_j] \rangle \\
&= \langle a_i a_j \rangle - \sum_k \langle a_i \frac{1}{z - (\hat{H} - E_0)} a_k \rangle h_{jk} + \sum_k \langle a_i \frac{1}{z - (\hat{H} - E_0)} a_k^\dagger \rangle \Delta_{kj}
\end{aligned}$$



$$= \langle a_i a_j \rangle - [G^{\text{hh},+} h^T]_{ij} + [G^{\text{hp},+} \Delta]_{ij},$$

and

$$\begin{aligned} zG_{ij}^{\text{hh},-}(z) &= \langle \Phi_0 | (z + (\hat{H} - E_0)) a_j \frac{1}{z + (\hat{H} - E_0)} a_i | \Phi_0 \rangle \\ &= \langle a_j a_i \rangle + \langle [\hat{H}, a_j] \frac{1}{z + (\hat{H} - E_0)} a_i \rangle \\ &= \langle a_j a_i \rangle + \langle [\hat{H}_{\text{NA}}, a_j] \frac{1}{z + (\hat{H} - E_0)} a_i \rangle + \langle [\hat{H}_{\text{A}}, a_j] \frac{1}{z + (\hat{H} - E_0)} a_i \rangle \\ &= \langle a_j a_i \rangle - \sum_k \langle a_k \frac{1}{z + (\hat{H} - E_0)} a_i \rangle h_{jk} + \sum_k \langle a_k^\dagger \frac{1}{z + (\hat{H} - E_0)} a_i \rangle \Delta_{kj} \\ &= \langle a_j a_i \rangle - [G^{\text{hh},-} h^T]_{ij} + [G^{\text{hp},-} \Delta]_{ij}. \end{aligned}$$

Adding our results and recognizing that  $\langle a_i a_j \rangle + \langle a_j a_i \rangle = 0$ , we obtain our second desired result.

Next compute

$$\begin{aligned} zG_{ij}^{\text{pp},+}(z) &= \langle \Phi_0 | a_i^\dagger \frac{1}{z - (\hat{H} - E_0)} a_j^\dagger (z - (\hat{H} - E_0)) | \Phi_0 \rangle \\ &= \langle a_i^\dagger a_j^\dagger \rangle + \langle a_i^\dagger \frac{1}{z - (\hat{H} - E_0)} [\hat{H}, a_j^\dagger] \rangle \\ &= \langle a_i^\dagger a_j^\dagger \rangle + \langle a_i^\dagger \frac{1}{z - (\hat{H} - E_0)} [\hat{H}_{\text{NA}}, a_j^\dagger] \rangle + \langle a_i^\dagger \frac{1}{z - (\hat{H} - E_0)} [\hat{H}_{\text{A}}, a_j^\dagger] \rangle \\ &= \langle a_i^\dagger a_j^\dagger \rangle + \sum_k \langle a_i^\dagger \frac{1}{z - (\hat{H} - E_0)} a_k^\dagger \rangle h_{kj} + \sum_k \langle a_i^\dagger \frac{1}{z - (\hat{H} - E_0)} a_k \rangle \bar{\Delta}_{jk} \\ &= \langle a_i^\dagger a_j^\dagger \rangle + [G^{\text{pp},+} h]_{ij} + [G^{\text{ph},+} \Delta^\dagger]_{ij}, \end{aligned}$$

and

$$\begin{aligned} zG_{ij}^{\text{pp},-}(z) &= \langle \Phi_0 | (z + (\hat{H} - E_0)) a_j^\dagger \frac{1}{z + (\hat{H} - E_0)} a_i^\dagger | \Phi_0 \rangle \\ &= \langle a_j^\dagger a_i^\dagger \rangle + \langle [\hat{H}, a_j^\dagger] \frac{1}{z + (\hat{H} - E_0)} a_i^\dagger \rangle \\ &= \langle a_j^\dagger a_i^\dagger \rangle + \langle [\hat{H}_{\text{NA}}, a_j^\dagger] \frac{1}{z + (\hat{H} - E_0)} a_i^\dagger \rangle + \langle [\hat{H}_{\text{A}}, a_j^\dagger] \frac{1}{z + (\hat{H} - E_0)} a_i^\dagger \rangle \\ &= \langle a_j^\dagger a_i^\dagger \rangle + \sum_k \langle a_k^\dagger \frac{1}{z + (\hat{H} - E_0)} a_i^\dagger \rangle h_{kj} + \sum_k \langle a_k \frac{1}{z + (\hat{H} - E_0)} a_i^\dagger \rangle \bar{\Delta}_{jk} \\ &= \langle a_j^\dagger a_i^\dagger \rangle + [G^{\text{pp},-} h]_{ij} + [G^{\text{ph},-} \Delta^\dagger]_{ij}, \end{aligned}$$

yielding our third desired result.

Finally, compute

$$\begin{aligned}
zG_{ij}^{\text{ph},+}(z) &= \langle \Phi_0 | a_i^\dagger \frac{1}{z - (\hat{H} - E_0)} a_j (z - (\hat{H} - E_0)) | \Phi_0 \rangle \\
&= \langle a_i^\dagger a_j \rangle + \langle a_i^\dagger \frac{1}{z - (\hat{H} - E_0)} [\hat{H}, a_j] \rangle \\
&= \langle a_i^\dagger a_j \rangle + \langle a_i^\dagger \frac{1}{z - (\hat{H} - E_0)} [\hat{H}_{\text{NA}}, a_j] \rangle + \langle a_i^\dagger \frac{1}{z - (\hat{H} - E_0)} [\hat{H}_{\text{A}}, a_j] \rangle \\
&= \langle a_i^\dagger a_j \rangle - \sum_k \langle a_i^\dagger \frac{1}{z - (\hat{H} - E_0)} a_k \rangle h_{jk} + \sum_k \langle a_i^\dagger \frac{1}{z - (\hat{H} - E_0)} a_k^\dagger \rangle \Delta_{kj} \\
&= \langle a_i^\dagger a_j \rangle - [G^{\text{ph},+} h^T]_{ij} + [G^{\text{pp},+} \Delta]_{ij},
\end{aligned}$$

and

$$\begin{aligned}
zG_{ij}^{\text{ph},-}(z) &= \langle \Phi_0 | (z + (\hat{H} - E_0)) a_j \frac{1}{z + (\hat{H} - E_0)} a_i^\dagger | \Phi_0 \rangle \\
&= \langle a_j a_i \rangle + \langle [\hat{H}, a_j] \frac{1}{z + (\hat{H} - E_0)} a_i^\dagger \rangle \\
&= \langle a_j a_i \rangle + \langle [\hat{H}_{\text{NA}}, a_j] \frac{1}{z + (\hat{H} - E_0)} a_i^\dagger \rangle + \langle [\hat{H}_{\text{A}}, a_j] \frac{1}{z + (\hat{H} - E_0)} a_i^\dagger \rangle \\
&= \langle a_j a_i \rangle - \sum_k \langle a_k \frac{1}{z + (\hat{H} - E_0)} a_i^\dagger \rangle h_{jk} + \sum_k \langle a_k^\dagger \frac{1}{z + (\hat{H} - E_0)} a_i^\dagger \rangle \Delta_{kj} \\
&= \langle a_j a_i \rangle - [G^{\text{ph},-} h^T]_{ij} + [G^{\text{pp},-} \Delta]_{ij},
\end{aligned}$$

yielding the last desired result. □

## Part VII

# The dynamical mean-field theory

## 1 Introduction

Green's function embedding refers to an extremely flexible framework for the embedding of field theories arising from the Green's function formalism. Many Green's function embedding theories can be viewed as specific ansatzes for the Luttinger-Ward functional. Here we will focus on the dynamical mean field theory (DMFT) [37, 53, 106, 58], or rather its extension to fragments consisting of several sites, known as cluster DMFT (though we shall conflate the two under the umbrella of 'DMFT'). DMFT relies on the solution of impurity problems (as considered in Part VI), which specify the embedding of an interacting system into a non-interacting bath. The system is divided into several fragments or clusters, and to each of these is associated an impurity problem determined via a self-consistency condition that couples the fragments on a global level. The method can be introduced in the classical setting (of Euclidean field theory), though in this setting 'DMFT' is really a misnomer because the mean fields are not dynamical. However, it is nonetheless the formal analog of the original (fermionic) DMFT, so we retain the nomenclature. DMFT can be extended to consider overlapping fragments and to include diagrammatic corrections.

### 1.1 Outline

As a warm-up, we describe in section 2 the framework of DMFT (and several extensions) in the setting of classical field theory. Then in section 3 we introduce the fermionic DMFT. In section 4, we identify the key mathematical structures in the algorithmic loop for solving the fermionic DMFT, and in section 5 we describe some applications of this mathematical framework, e.g., to the well-posedness of the algorithmic DMFT loop. Please note that sections 3, 4, and 5 are based on joint work in preparation with Lin Lin and Reinhold Schneider.

## 2 Warm-up: DMFT in Euclidean field theory

### 2.1 Formulation of cluster DMFT

In this section we adopt the notation of Part III. Let  $G \in \mathbb{R}^{M \times M}$  be the exact Green's function

$$G_{ij} = \langle \phi_i \phi_j \rangle.$$

Suppose that  $M = RL$ . Define diagonal  $L \times L$  blocks of  $G$  via

$$G = \begin{pmatrix} G_1 & * & \cdots & * \\ * & G_2 & \cdots & * \\ \vdots & \vdots & \ddots & * \\ * & * & * & G_R \end{pmatrix}.$$

We will index the blocks ('clusters') by the index  $c$ . Let  $H_{1,c} = H_{1,c}(\phi_c)$  be the contribution to  $H_1$  that depends only on  $\phi_c := (\phi_{(c-1)L+1}, \dots, \phi_{cL})$ .

Then consider the *functionals*  $\Omega_c, \mathcal{F}_c, \Phi_c, \Sigma_c = \nabla \Phi_c$ , defined as previously but now for the  $L \times L$  problem with interaction  $H_{1,c}$ .

We define the (*exact*) *hybridization*  $\Delta_c$  via  $A_c + \Delta_c - G_c^{-1} = \Sigma_c[G_c]$ , i.e.,

$$\Delta_c = \Sigma_c[G_c] - A_c + G_c^{-1}.$$

In other words,  $\Delta_c$  is the  $L \times L$  matrix such that  $A_c + \Delta_c \mapsto G_c$  under the Legendre correspondence for the  $c$ -th cluster.

Since  $G_c = \nabla \Omega_c(A_c + \Delta_c)$ , we have

$$\Delta_c = \Sigma'_c[A_c + \Delta_c] - A_c + (\pi_c G \pi_c^T)^{-1},$$

where  $\Sigma'_c := \Sigma \circ \nabla \Omega_c$  and  $\pi_c : \mathbb{R}^M \rightarrow \mathbb{R}^L$  is the projection onto the  $c$ -th cluster, so  $G_c = \pi_c^T G \pi_c$ .

Recall that

$$G = (A - \Sigma)^{-1},$$

where  $\Sigma$  is the exact self-energy, so in fact

$$\Delta_c = \Sigma'_c[A_c + \Delta_c] - A_c + (\pi_c (A - \Sigma)^{-1} \pi_c^T)^{-1}.$$

We will introduce an ansatz for  $\Sigma$  in terms of the hybridizations  $\Delta_c$ , and it is this approximation that defines (C)DMFT:

$$\Sigma \leftarrow \bigoplus_{c'=1}^R \Sigma'_{c'} [A_{c'} + \Gamma_{c'}].$$

This yields a self-consistent set of equations for the  $\Gamma_c$ :

$$\Gamma_c = \Sigma'_c[A_c + \Delta_c] - A_c + \left[ \pi_c \left( A - \bigoplus_{c'=1}^R \Sigma'_{c'} [A_{c'} + \Delta_{c'}] \right)^{-1} \pi_c^T \right]^{-1}. \quad (2.1)$$

The (approximate) self-energy and Green's function can then be retrieved from the  $\Delta_c$  that solve these equations.

## 2.2 Algorithmic approach

One iterative approach to solving the equations (2.1) consists of setting

$$\Gamma_c^{(k+1)} = \Sigma'_c [A_c + \Delta_c^{(k)}] - A_c + \left[ \pi_c \left( A - \bigoplus_{c'=1}^R \Sigma'_{c'} [A_{c'} + \Delta_{c'}^{(k)}] \right)^{-1} \pi_c^T \right]^{-1}.$$

This is equivalent to the approach that is outlined in [96]. This is a fixed-point iteration, which can be mixed and/or accelerated by standard methods.

## 2.3 Luttinger-Ward Perspective

There is a much simpler way to view DMFT as an approximation of the Luttinger-Ward functional (and consequently the self-energy functional).

Suppose that we have solved (2.1) for the  $\Delta_c$ , and define

$$\Sigma := \bigoplus_{c=1}^R \Sigma'_c [A_c + \Delta_c], \quad G := (A - \Sigma)^{-1}, \quad G_c := \pi_c G \pi_c^T.$$

Then (2.1) reads

$$\Delta_c = \Sigma'_c [A_c + \Gamma_c] - A_c + G_c^{-1}, \quad (2.2)$$

But we also know (from stationarity for the  $c$ -th cluster) that

$$A_c + \Delta_c - \Sigma'_c [A_c + \Delta_c] = [\nabla \Omega_c (A_c + \Delta_c)]^{-1}.$$

Therefore, combining with (2.2) yields

$$G_c = \nabla \Omega_c (A_c + \Delta_c).$$

Thus

$$\Sigma'_{c'} [A_{c'} + \Delta_{c'}] = \Sigma_{c'} [G_{c'}],$$

and

$$\Sigma = \bigoplus_{c=1}^R \Sigma_c [G_c].$$

Now define a *functional* via

$$\Sigma^{\text{CDMFT}}[G] := \bigoplus_{c=1}^R \Sigma_c [\pi_c G \pi_c^T].$$

Then evidently the  $G$  retrieved from CDMFT satisfies

$$A - G^{-1} = \Sigma^{\text{DMFT}}[G].$$

In other words, DMFT is equivalent to replacing the universal functional  $\Sigma[\cdot]$  with  $\Sigma^{\text{DMFT}}[\cdot]$  in the Dyson equation.

Moreover, this self-energy is  $\Phi$ -derivable in that it can be viewed as the gradient of an appropriate (replacement for) the Luttinger-Ward functional, to wit:

$$\Phi^{\text{DMFT}}[G] := \sum_{c=1}^R \Phi_c [\pi_c G \pi_c^T].$$

## 2.4 Diagrammatic corrections to DMFT

We now outline a framework for considering diagrammatic corrections to cluster DMFT. In some sense, DMFT is good at treating local interactions, though a diagrammatic method could help take into account small long-range effects. More specifically, the DMFT self-energy should take into account all diagrams in which *all* the factors of  $G_{ij}$  have indices living within a *single* cluster. When we add diagrammatic corrections, we do not want to double-count diagrams. Therefore we define

$$\Sigma^{\text{D+DMFT}}[G] := \bigoplus_{c=1}^R (\Sigma_c[G_c] - \Sigma_c^{\text{D}}[G_c]) + \Sigma^{\text{D}}[G],$$

where  $\Sigma^{\text{D}}[\cdot]$  is the self-energy functional of the diagrammatic method with which we are combining DMFT, and  $\Sigma_c^{\text{D}}[\cdot]$  is the functional that takes into account all diagrams in which *all* the factors of  $G_{ij}$  have indices living in the  $c$ -th cluster. For most methods,  $\Sigma_c^{\text{D}}[\cdot]$  would simply be the self-energy functional for the lower-dimensional  $L \times L$  problem with interaction  $H_{1,c}$ .

Moreover, if the diagrammatic method (along with its lower-dimensional versions) is  $\Phi$ -derivable, then we have the following ansatz form for the approximate Luttinger-Ward functional:

$$\Phi^{\text{D+CDMFT}}[G] = \sum_{c=1}^R (\Phi_c [\pi_c G \pi_c^T] - \Phi_c^{\text{D}} [\pi_c G \pi_c^T]) + \Phi^{\text{D}}[G].$$

## 2.5 Overlapping DMFT

The block-diagonal ansatz for the self-energy seems to leave something to be desired. For instance, no matter how much we increase the block size, off-block-diagonal entries of  $G$  *never* appear in any of the diagrams contributing to  $\Sigma^{\text{DMFT}}[G]$ . In particular, at the juncture between two clusters, we have entries adjacent to the diagonal that do not contribute!

For now we will imagine the entries of  $\phi$  as corresponding to sites in a one-dimensional periodic system, but the following construction could be generalized to an arbitrary lattice in any dimension.

As before, we fix a block size  $L$ , though we no longer stipulate that  $L$  divide the total number of sites  $M$ . However, in the periodic case, for simplicity we insist that  $2L+1 \leq M$ . This implies that any two clusters of size  $L$  intersect only in a ‘connected’ patch. (If this were not the case, we would have to subtract off certain contributions due to more complicated intersections later on.) This is not a restrictive assumption because in applications we have that the cluster size is much smaller than the total number of sites.

For a given self-energy diagram, we say that the *range* of this diagram is the minimal size of a cluster  $C \subset \mathbb{Z}/M\mathbb{Z}$  (i.e., a contiguous patch of sites) such that  $i, j \in C$  for every factor of  $G_{ij}$  that appears in this diagram. We want to take into account every diagram of range at most  $L$ , and we will call the resulting approximate self-energy functional the *short-range self-energy (SRSE)*, denoted  $\Sigma^{\text{SR}}$ .

Let  $i = 1, \dots, n$  denote a ‘cluster shift,’ and let  $\pi_{i,L} : \mathbb{R}^M \rightarrow \mathbb{R}^L$  denote the projection onto the entries  $i, \dots, i+L-1 \pmod{M}$ . We will use this indexing convention for other objects (arising from the  $L \times L$  problem with interaction  $U_{i,L}$ ) without comment. Then define

$$\Sigma^{(L)}[G] = \sum_{i=1}^M \pi_{i,L}^T \left( \Sigma_{i,L} \left[ \pi_{i,L} G \pi_{i,L}^T \right] \right) \pi_{i,L}.$$

Notice that in this functional, every diagram of range  $L$  is counted once, every diagram of range  $L-1$  is counted twice, etc., and every diagram of range 1 is counted  $L$  times.

Therefore, more generally, for  $q = 1, \dots, L$  define  $\pi_{i,q} : \mathbb{R}^M \rightarrow \mathbb{R}^q$  to be the projection onto the entries  $i, \dots, i+q-1 \pmod{M}$ , and define

$$\Sigma^{(q)}[G] = \sum_{i=1}^M \pi_{i,q}^T \left( \Sigma_{i,q} \left[ \pi_{i,q} G \pi_{i,q}^T \right] \right) \pi_{i,q},$$

so the functional  $\Sigma^{(q)}$  counts every diagram of range  $q$  once, every diagram of range  $q-1$  twice, etc. Then define

$$\Sigma^{\text{SR}} := \Sigma^{(L)} - \Sigma^{(L-1)}.$$

Evidently this functional counts every diagram of any range up to  $p$  exactly once.

One can imagine a similar construction for the nonperiodic case. Indeed, define

$$\Sigma_{\text{nonper}}^{(q)}[G] = \sum_{i=1}^{M-q+1} \pi_{i,q}^T \left( \Sigma_{i,q} \left[ \pi_{i,q} G \pi_{i,q}^T \right] \right) \pi_{i,q},$$

and set

$$\Sigma_{\text{nonper}}^{\text{SR}}[G] := \Sigma_{\text{nonper}}^{(L)} - \Sigma_{\text{nonper}}^{(L-1)}.$$

Then  $\Sigma_{\text{nonper}}^{\text{SR}}$  counts any diagram of ‘nonperiodic range’ at most  $L$ , where the nonperiodic range of a diagram is the minimal size of a cluster  $C \subset \{1, \dots, M\}$  such that  $i, j \in C$  for every factor of  $G_{ij}$  that appears in this diagram. Note that for this construction, we do not need to assume  $2L+1 \leq M$ .

## 2.6 Diagrammatic corrections to overlapping DMFT

We can consider diagrammatic corrections to the SRSE in the same way as we did before. Take the periodic one-dimensional case for specificity. Define

$$\Sigma_q^{\text{D}}[G] = \sum_{i=1}^M \pi_{i,q}^T \left( \Sigma_{i,q} \left[ \pi_{i,q} G \pi_{i,q}^T \right] - \Sigma_{i,q}^{\text{D}} \left[ \pi_{i,q} G \pi_{i,q}^T \right] \right) \pi_{i,q},$$

and set

$$\Sigma^{\text{D+SR}} := \Sigma_L^{\text{D}} - \Sigma_{L-1}^{\text{D}}$$

Lastly, observe that the remarks about the  $\Phi$ -derivability of (D+)DMFT carry over to (D+)SRFT with the obvious modifications. In other words,  $\Sigma^{\text{SR}}$  can be viewed as the gradient of the appropriate *short-range Luttinger-Ward functional (SRLW)*  $\Phi^{\text{SR}}$ , and, if the diagrammatic method D is  $\Phi$ -derivable, we can say the same about  $\Sigma^{\text{D+SR}}$ .

## 3 Fermionic DMFT

### 3.1 Basic definitions

For simplicity and mathematical convenience we confine our attention to the case of fermionic many-body systems described by a finite-dimensional Fock space  $\mathcal{F}$ , i.e., a Fock space with finitely many creation operators  $a_1^\dagger, \dots, a_M^\dagger$ . This setting can describe lattice models such as the Hubbard model as well as tight-binding approximations of continuum systems, and we view the set  $\{1, \dots, M\}$  as indexing ‘sites’ in our model. For convenience we shall let  $a = (a_1, \dots, a_N)$  denote the vector of annihilation operators for the sites in our model. Recall the canonical anticommutation relations

$$a_i a_j + a_j a_i = 0, \quad a_i^\dagger a_j^\dagger + a_j^\dagger a_i^\dagger = 0, \quad a_i^\dagger a_j + a_j a_i^\dagger = \delta_{ij}.$$

For further details on second quantization, refer to section 4 of Part I.

We remark that, more generally, one can reduce a continuum problem to this setting via the choice of an orbital basis. To wit, in the electronic structure problem one replaces the single-particle wavefunction space  $\mathcal{V} := H^1(\mathbb{R}^3, \pm\frac{1}{2}) \subset L_2(\mathbb{R}^3, \pm\frac{1}{2})$  with a finite-dimensional subspace  $\mathcal{V}_{\mathcal{D}}$  spanned by an orthonormal single-particle basis  $\mathcal{D} := \{\varphi_1, \dots, \varphi_M\}$ . Then one defines the  $N$ -particle space  $\mathcal{F}_{\mathcal{D}}^N := \bigwedge_{i=1}^N \mathcal{V}_{\mathcal{D}}$  and considers the discrete Fock space  $\mathcal{F} = \mathcal{F}_{\mathcal{D}} = \bigoplus_{N=0}^{\infty} \mathcal{F}_{\mathcal{D}}^N$ . The Fock space  $\mathcal{F}_{\mathcal{D}}$  can be viewed as being induced by creation operators  $a_i^\dagger$  corresponding to the orbital basis functions  $\varphi_i$ .

We consider a Hamiltonian  $\hat{H}$  on the Fock space  $\mathcal{F}$ , namely a Hermitian linear operator  $\hat{H} : \mathcal{F} \rightarrow \mathcal{F}$ . We write the Hamiltonian as

$$\hat{H} = \hat{H}_0 + \hat{V},$$



where

$$\hat{H}_0 := a^\dagger h a = \sum_{i,j=1}^N h_{pq} a_p^\dagger a_q$$

is the *single-particle* (or *noninteracting*) part of the Hamiltonian, specified by a Hermitian  $N \times N$  matrix  $h$ , and  $\hat{V}$  is the *interacting* part of the Hamiltonian, which is an even polynomial in  $a^\dagger, a$ . Though we need not define  $\hat{V}$  explicitly for most of the developments of this Part, we have in mind the two-body interaction

$$\hat{V} = \frac{1}{2} \sum_{pqrs} (pq|V|rs) a_p^\dagger a_q^\dagger a_s a_r.$$

In the setting of electronic structure,  $h_{pq}$  and  $(pq|V|rs)$  are the matrix and tensor elements for the one- and two-body interactions, respectively; section 4 of Part I for more details. Nonetheless, for our present purposes we need only assume that  $\hat{V}$  commutes with the total number operator  $\hat{N} := \sum_{i=1}^M a_i^\dagger a_i$ . In this case we say that  $\hat{H}$  conserves particle number.

When discussing quantum many-body problems in this Part, we take the perspective a chemical potential  $\mu \in \mathbb{R}$  is specified in advance. In practical DMFT, the chemical potential may be adjusted (concurrently with the solution of self-consistency conditions for the method) to guarantee that a given total particle number constraint is satisfied. In this Part, we ignore the details of chemical potential fitting and simply assume that the desired chemical potential is given.

Now in the zero-temperature setting, we let  $P_0$  be the orthogonal projector onto the lowest eigenspace of  $\hat{H} - \mu \hat{N}$ , and we denote the corresponding eigenvalue (the ground-state energy) by  $E_0$ . In the case in which the lowest eigenspace is simple, we have  $P_0 = |\Psi_0\rangle\langle\Psi_0|$ , where  $|\Psi_0\rangle$  is the unique ground state. If  $\hat{N}|\Psi_0\rangle = N|\Psi_0\rangle$ , i.e., if the ground state is of definite particle number  $N$ , then we will sometimes write  $|\Psi_0^N\rangle$  to emphasize this point. Then one defines the Green's function (cf. 7 of Part I) as follows.

**Definition 1.** With notation and conventions as in the preceding,<sup>14</sup> we define the zero-temperature Green's function  $G : \mathbb{C} \rightarrow \mathbb{C}^{M \times M}$  by<sup>15</sup>

$$G(z) = G^{(+)}(z) + G^{(-)}(z),$$

where

$$G_{ij}^{(+)}(z) := \frac{1}{Z} \text{Tr} \left[ a_i (z - [\hat{H} - E_0])^{-1} a_j^\dagger P_0 \right], \quad G_{ij}^{(-)}(z) := \frac{1}{Z} \text{Tr} \left[ a_j^\dagger (z + [\hat{H} - E_0])^{-1} a_i P_0 \right].$$

<sup>14</sup>In order to reduce notational burden, all definitions and results should be understood as being stated with this caveat.

<sup>15</sup>In fact  $G$  is only rational, hence undefined at certain points in  $\mathbb{C}$ , but we nonetheless shall write the domain of such functions as  $\mathbb{C}$ .

Here  $Z = \text{Tr}[P_0]$ . If the ground state  $|\Psi_0^N\rangle$  of  $\hat{H} - \mu\hat{N}$  is nondegenerate, we have

$$G_{ij}^{(+)}(z) := \langle \Psi_0^N | a_i(z - [\hat{H} - E_0])^{-1} a_j^\dagger | \Psi_0^N \rangle, \quad G_{ij}^{(-)}(z) := \langle \Psi_0^N | a_j^\dagger(z + [\hat{H} - E_0])^{-1} a_i | \Psi_0^N \rangle.$$

We call these two terms, respectively, the  $(N \pm 1)$ -particle parts of the Green's functions.

*Remark 2.* Note carefully that we do not include any infinitesimal offset  $i\eta$  for the poles in our definition of the Green's function; the choice of such offsets distinguishes the advanced, retarded, and time-ordered Green's functions from one another (see, e.g., [77]). However, this additional structure is *irrelevant* to the DMFT loop, i.e., it is sufficient for our purposes to think of Green's functions merely as *rational functions* without any such additional data. As such, the reader familiar with many-body Green's functions may think of our definition as conflating the advanced, retarded, and time-ordered Green's functions.

Next we consider the finite temperature setting, where we let  $\beta \in (0, \infty)$  denote the inverse temperature.

**Definition 3.** The Green's function  $G : \mathbb{C} \rightarrow \mathbb{C}^{M \times M}$  for inverse temperature  $\beta$  is defined via

$$G(z) = G^{(+)}(z) + G^{(-)}(z),$$

where now

$$G_{ij}^{(+)}(z) := \frac{1}{Z} \sum_{\lambda \in \sigma(\hat{H})} e^{-\beta\lambda} \text{Tr} \left[ a_i(z - [\hat{H} - \lambda])^{-1} a_j^\dagger P_\lambda \right]$$

and

$$G_{ij}^{(-)}(z) := \frac{1}{Z} \sum_{\lambda \in \sigma(\hat{H})} e^{-\beta\lambda} \text{Tr} \left[ a_j^\dagger(z + [\hat{H} - \lambda])^{-1} a_i P_\lambda \right].$$

Here  $\sigma(\hat{H})$  denotes the spectrum of  $\hat{H}$ ,  $P_\lambda$  is the orthogonal projector onto the  $\lambda$ -eigenspace of  $\hat{H}$ , and  $Z = \sum_{\lambda \in \sigma(\hat{H})} e^{-\beta\lambda} \text{Tr}[P_\lambda]$ .

One sees that the zero-temperature definition is recovered in the limit  $\beta \rightarrow \infty$ . Notice that the adjustment via chemical potential corresponds to the subtraction of the scalar matrix  $\mu I_N$  from  $h$ . Throughout we will simply assume that the given  $h$  has already incorporated any such chemical potential.

In the case of a noninteracting Hamiltonian, i.e., the case  $\hat{V} = 0$ , the corresponding (noninteracting) Green's function is denoted  $G_0(z)$ , and we have the following result:

**Lemma 4.** *In arbitrary temperature,  $G_0(z) = (zI_M - h)^{-1}$ , i.e.,  $G_0$  is the resolvent of  $h$ .*

*Proof.* This result follows from elementary manipulations (made easier by reducing to the case of diagonal  $h$  via a suitable unitary transformation of the orbitals), but see Remark (7) below for a quick proof from the point of view of impurity problems.  $\square$

This motivates the definition of the *self-energy* as the difference of the pointwise matrix inverses of the interacting and noninteracting Green's functions:

**Definition 5.** The self-energy  $\Sigma : \mathbb{C} \rightarrow \mathbb{C}^{M \times M}$  is defined (in arbitrary temperature) by

$$\Sigma(z) = zI_M - h - G^{-1}(z). \quad (3.1)$$

We shall sometimes abbreviate  $z = zI_M$ , but sometimes it is useful to emphasize the dimension of the identity matrix.

Note that (3.1) is equivalent to

$$G(z) = (z - h - \Sigma(z))^{-1},$$

hence  $h + \Sigma(z)$  can be interpreted as specifying the effective (frequency-dependent) single-particle Hamiltonian yielding  $G(z)$  as its *noninteracting* Green's function.

Observe, moreover, that (3.1) is equivalent to the *Dyson equation*

$$G(z) = G_0(z) + G_0(z)\Sigma(z)G(z),$$

which can therefore be alternatively taken as the defining property for the self-energy.

## 3.2 The impurity problem and the hybridization

To define the impurity problem we partition our sites  $\{1, \dots, N\}$  into two pieces: a fragment or impurity (thought of as being small) and an environment. Let  $L$  denote the number of sites in the fragment, indexed  $\mathcal{A} = \{1, \dots, L\}$  without loss of generality. Let  $M - L$  is the number of sites in the environment, indexed  $\mathcal{B} = \{L+1, \dots, M\}$ . Let  $\mathcal{F}_{\mathcal{A}} \simeq \mathbb{C}^{2^L}$  and  $\mathcal{F}_{\mathcal{B}} \simeq \mathbb{C}^{2^{M-L}}$  denote the Fock spaces for the fragment and environment, respectively, and let  $\mathcal{F} = \mathcal{F}_{\mathcal{A}} \otimes \mathcal{F}_{\mathcal{B}} \simeq \mathbb{C}^{2^M}$  denote the total Fock space.

We define an *impurity Hamiltonian* (relative to this partition) as an operator  $\hat{H} = a^\dagger h a + \hat{V}$  as above satisfying  $\hat{V} = \hat{V}_{\mathcal{A}} \otimes \text{Id}_{\mathcal{F}_{\mathcal{B}}}$ , where  $\hat{V}_{\mathcal{A}}$  is an operator on  $\mathcal{F}_{\mathcal{A}}$ .

Define blocks of  $h$  relative to our partition by

$$h = \begin{pmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{pmatrix},$$

where  $h_{11}$  is  $L \times L$ , etc.

The self-energy has an special sparsity pattern in the case of impurity problems, proved in Part VI, which we recapitulate here:

**Proposition 6.** *For an impurity Hamiltonian as defined above, the self-energy (in both zero and finite temperature) has the sparsity pattern*

$$\Sigma(z) = \begin{pmatrix} * & 0 \\ 0 & 0 \end{pmatrix},$$

where the upper-left block is  $L \times L$ .

*Remark 7.* In fact Lemma 4 follows in particular from Proposition 6 by considering the case of a fragment of size zero. In this case,  $\Sigma \equiv 0$ , so by (3.1),  $G(z) = (z - h)^{-1}$ .

We now turn to discuss the hybridization function, or hybridization for short:

**Definition 8.** The hybridization  $\Delta : \mathbb{C} \rightarrow \mathbb{C}^{L \times L}$  associated to an impurity problem is defined by

$$\Delta(z) = h_{12}(z - h_{22})^{-1}h_{21}.$$

Thus in fact the hybridization is a Schur complement derived from the matrix  $z - h$ . The significance of the hybridization can be explained as follows. Recall that the self-energy is given by

$$\Sigma(z) = z - h - G^{-1}(z),$$

so, left-multiplying by  $G(z)$  and right-multiplying by  $(zI_M - h)^{-1}$ , we obtain

$$G(z)\Sigma(z)(z - h)^{-1} = G(z) - (z - h)^{-1}.$$

We use the subscript indices 1 and 2 to denote appropriate blocks as done earlier for the matrix  $h$ . By the sparsity result Proposition, taking the upper-left block of the preceding equality yields

$$G_{11}(z)\Sigma_{11}(z) [(zI_N - h)^{-1}]_{11} = G_{11}(z) - [(zI_N - h)^{-1}]_{11}.$$

By the Schur complement theorem,  $[(zI_N - h)^{-1}]_{11} = [z - h_{11} - \Delta(z)]^{-1}$ , so left-multiplying by  $G_{11}(z)$  and right-multiplying by  $[z - h_{11} - \Delta(z)]$  yields

$$\Sigma_{11}(z) = z - h_{11} - \Delta(z) - G_{11}(z)^{-1}, \quad (3.2)$$

so the hybridization is an error term that measures the failure of the upper-left blocks of the matrices  $\Sigma$ ,  $h$ , and  $G$  to satisfy the appropriate  $L \times L$  Dyson equation. In other words, we can interpret

$$h_{11} + \Delta(z) + \Sigma_{11}(z)$$

as specifying the effective single-particle Hamiltonian on the fragment that yields  $G_{11}(z)$  as its noninteracting Green's function.

For concreteness, suppose that we have an eigenvalue decomposition of  $h_{22}$ , i.e.,  $h_{22}\Phi_2 = \Phi_2\Lambda_2$ , where  $\Lambda_2 = \text{diag}(\lambda_1, \dots, \lambda_{M-L})$  is diagonal. Then we can write

$$\Delta(z) = h_{12}\Phi_2(z - \Lambda_2)^{-1}\Phi_2^*h_{21} = T(z - \Lambda_2)^{-1}T^*,$$

where  $T := h_{12}\Phi_2$ . Elementwise we have

$$\Delta_{ij}(z) = \sum_{l=1}^{M-L} \frac{T_{il}\overline{T_{jl}}}{z - \lambda_l}.$$

We have exhibited an *exact* rational expression for the hybridization. Note that we can alternatively write

$$\Delta(z) = \sum_{l=1}^{M-L} \frac{T_l T_l^*}{z - \lambda_l} = \sum_{l=1}^{M-L} \frac{Y_l}{z - \lambda_l},$$

where  $T_l$  is the  $l$ -th column of  $T$  and  $Y_l := T_l T_l^*$  is positive semidefinite of rank one.

### 3.2.1 Bath fitting

One can approximate the influence of the bath on the fragment by replacing  $\Delta(z)$  with an approximation  $\Delta_{\text{approx}}(z)$  of the form

$$\Delta_{\text{approx}}(z) = \sum_{k=1}^K \frac{X_k}{z - \varepsilon_k}, \quad (3.3)$$

where  $X_k \succeq 0$  (i.e.,  $X_k$  is Hermitian positive semidefinite) and  $\varepsilon_k \in \mathbb{R}$ . (This form defines the class  $\mathcal{S}_+^L$  to be studied in detail below.) Note that a function of this form can be interpreted as specifying the coupling of a fragment of size  $L$  to an effective bath of size  $M' = \sum_{k=1}^K \text{rank}(X_k)$ . Therefore one may attempt to reduce the difficulty of solving a given impurity problem with hybridization  $\Delta$  by approximating  $\Delta_{\text{approx}} \approx \Delta$  with  $M'$  relatively small, i.e., small enough so that the size  $M + M'$  of the composite fragment/effective-bath system can be approached with solvers of high accuracy, or even solved directly by exact diagonalization.

The sense in which one attempts to approximate  $\Delta(z)$  is not the focus of this work, but we will highlight the method of [73], which is naturally motivated by the analytic framework. The method fixes a number  $K$  of poles and then decomposes the optimization over approximate hybridizations of the form (3.3) into an outer optimization over poles  $\varepsilon = (\varepsilon_1, \dots, \varepsilon_K) \in \mathbb{R}^K$  and an inner optimization, for fixed poles  $\varepsilon$ , over residues  $X_k \succeq 0$ . To wit, one defines a loss for fixed poles  $\varepsilon$  as

$$F(\varepsilon) := \inf \left\{ \sum_l \left\| \sum_{k=1}^K \frac{X_k}{z_l - \varepsilon_k} - \Delta(z_l) \right\|_{\mathbb{F}}^2 \mid X_1, \dots, X_k \succeq 0 \right\},$$

where  $\{z_l\}$  is a sampling of points in the complex plane, in practice chosen to be a discretization of part of the imaginary axis. Then the quantity  $F(\varepsilon)$ , as well as  $\nabla_{\varepsilon} F(\varepsilon)$ , can be computed via convex optimization over the  $X_k$ . The gradients are used to implement an outer optimization over the poles  $\varepsilon$ .

### 3.2.2 Chain Hamiltonian perspective

There is another perspective on approximating the influence of the bath on the fragment. Again consider an impurity Hamiltonian of the form  $\hat{H} = a^\dagger h a + \hat{V}$ , where

$$h = \begin{pmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{pmatrix}.$$

Suppose, for simplicity that  $h$  is of size  $(nL) \times (nL)$ . Then via, e.g., the block Lanczos algorithm, one can compute a unique unitary matrix

$$U = \begin{pmatrix} I_L & 0 \\ 0 & V \end{pmatrix},$$

where  $V \in \mathbb{C}^{(M-L) \times (M-L)}$  is unitary, such that  $U^*hU$  is tridiagonal, i.e.,

$$U^*hU = \begin{pmatrix} h_{11} & D_0 & & & & \\ D_0^* & E_0 & D_1 & & & \\ & \ddots & \ddots & \ddots & & \\ & & D_{n-2}^* & E_{n-2} & D_{n-1} & \\ & & & D_{n-1}^* & E_{n-1} & \end{pmatrix}.$$

After a suitable canonical transformation of our creation and annihilation operators, it follows that we can assume that  $h$  is of tridiagonal form.

The benefit of this perspective is that our Hamiltonian assumes the form of a *local one-dimensional* Hamiltonian, or a ‘chain Hamiltonian,’ if we identify our Fock space with the  $n$ -fold tensor product  $\bigotimes^n (\mathbb{C}^{2^L})$ . Hence tensor methods such as DMRG can be fruitfully applied, even when the chain is relatively long (i.e., even when the dimension  $2^{nL}$  of the total composite Fock space is much too large to permit exact diagonalization).

Moreover, this perspective suggests a physically intuitive framework for truncating the size of the bath, i.e., simply truncating the tridiagonal matrix at a certain depth  $n' < n$ . Since ultimately one cares to compute only the fragment part of the Green’s function (which determines the self-energy via (3.2)), the hope is that the influence of the tail of the chain becomes negligible at a certain distance. More rigorous and quantitative justification of such truncation remains an open question.

### 3.3 The (cluster) dynamical mean field theory

Consider a partition of the set  $\mathcal{C} = \{1, \dots, M\}$  of site indices into  $P$  disjoint subsets  $\mathcal{C}^{(1)}, \dots, \mathcal{C}^{(P)}$ , which may be referred to as the ‘clusters.’<sup>16</sup> We write  $L^{(p)} := |\mathcal{C}^{(p)}|$ . Throughout we shall write  $\bigoplus_{p=1}^P A^{(p)}$  to indicate a block-diagonal matrix, with diagonal block corresponding to the indices  $\mathcal{C}^{(p)}$  given by  $A^{(p)}$ . For an  $M \times M$  matrix  $A$ , we let  $A_{\mathcal{C}^{(p)}, \mathcal{C}^{(q)}}$  denote the appropriate submatrix of  $A$ . More generally we may consider submatrices  $A_{\mathcal{I}, \mathcal{J}}$  for any  $\mathcal{I}, \mathcal{J} \subset \mathcal{C}$ , and we always denote  $A_{\mathcal{I}} := A_{\mathcal{I}, \mathcal{I}}$ .

Before proceeding, we comment that DMFT is applicable to the setting in which our global Hamiltonian is specified by

$$\hat{H} = a^\dagger h a + \sum_{p=1}^P \hat{V}^{(p)},$$

---

<sup>16</sup>When we think the cluster as the impurity part of an impurity problem, we shall use the terminology ‘fragment’ as above.

where  $\hat{V}^{(p)}$  is an interaction on  $\mathcal{F}^{(p)}$ , the Fock-space for  $p$ -th cluster, i.e., a Hermitian particle-number-conserving operator that is written as a polynomial in  $a_p^\dagger, a_p$  for  $p \in \mathcal{C}^{(p)}$ . In particular, this form recovers the Hubbard model and several related Hamiltonians as special cases. While DMFT cannot ‘see’ interaction terms that do not live entirely within a single cluster, we shall discuss in section 5.3 the extension HF+DMFT, which can include the effect of longer-range interaction.

### 3.3.1 The self-consistency condition

Now the dynamical mean field theory (DMFT)<sup>17</sup> amounts to a block-diagonal ansatz for the global self-energy:

$$\Sigma_{\text{DMFT}}(z) = \bigoplus_{p=1}^P \Sigma^{(p)}(z), \quad (3.4)$$

where each  $\Sigma^{(p)}$  is determined in a manner to be described shortly. First, however, observe that an ansatz for the global self-energy determines a global Green’s function via the Dyson equation as

$$G^{\text{glob}}(z) = (zI_M - h - \Sigma_{\text{DMFT}}(z))^{-1}. \quad (3.5)$$

Now for each cluster  $p$ , we consider an impurity problem specified by a frequency-dependent single-particle matrix  $h^{(p)} + \Delta^{(p)}(z)$ , where  $h^{(p)} = h_{\mathcal{C}^{(p)}}$ , and by the impurity interaction  $\hat{V}^{(p)}$ , in the sense of section 3.2 above. If we view each of these impurity problems as a many-body problem on a composite fragment/bath Fock space, then we let  $\Sigma^{(p)}$  denote the fragment block of the self-energy. (Recall that all other blocks of the fragment-bath self-energy are zero by Proposition (6).) The hybridizations  $\Delta^{(p)}$  are the unknowns of the method, and they are determined via the stipulation, inspired by (3.2), that

$$\Sigma^{(p)}(z) = z - h^{(p)} - \Delta^{(p)}(z) - G^{(p)}(z)^{-1}, \quad (3.6)$$

where  $G^{(p)} := [G^{\text{glob}}]_{\mathcal{C}^{(p)}}$  is the  $p$ -th cluster block of the global Green’s function. Note that these equations must be solved self-consistently (and moreover are coupled across  $p = 1, \dots, P$ ) because the Green’s function blocks  $\{G^{(p)}\}$  are determined by the hybridizations  $\{\Delta^{(p)}\}$  via the progression

$$\begin{array}{ccc} \{\Delta^{(p)}\} & \longmapsto & \{\Sigma^{(p)}\} \\ \text{impurity problems} & & \text{ansatz (3.4)} \\ & \longmapsto & \Sigma_{\text{DMFT}} \\ & \longmapsto & G^{\text{glob}} \\ \text{Dyson (3.5)} & & \text{extract blocks} \\ & & \{G^{(p)}\}. \end{array}$$

---

<sup>17</sup>In fact, when the clusters contain more than a single site, what we describe is often referred to as the cluster DMFT.

We shall call this map from  $G^{(p)} [\{\Delta^{(q)}\}]$ , and moreover we denote the map  $\Delta^{(p)} \mapsto \Sigma^{(p)}[\Delta^{(p)}]$  defined by the solution the appropriate impurity problem by  $\Sigma^{(p)}[\Delta^{(p)}]$ . Hence we may write our self-consistency equations somewhat more painstakingly as

$$\Delta^{(p)}(z) := z - h^{(p)} - \Sigma^{(p)}[\Delta^{(p)}](z) - (G^{(p)}[\{\Delta^{(q)}\}](z))^{-1}. \quad (3.7)$$

Recall that to be a ‘physical’ hybridization, i.e., a hybridization that corresponds to a closed fragment+bath system, we must solve for  $\Delta^{(p)} \in \mathcal{S}_+^{L^{(p)}}$ , i.e.,  $\Delta^{(p)}$  must be of the form (3.3).

We remark that the self-consistency equations (3.6) can be viewed as formally equivalent to the ansatz

$$\Phi_{\text{DMFT}}[G] = \sum_{p=1}^P \Phi^{(p)}[G_{\mathcal{C}^{(p)}}; \hat{V}^{(p)}]$$

for the Luttinger-Ward functional,  $G$  is the global Green’s function and where the quantity  $\Phi^{(p)}[G_{\mathcal{C}^{(p)}}; \hat{V}^{(p)}]$  is the exact Luttinger-Ward functional for the cluster with interaction  $\hat{V}^{(p)}$ , or equivalently, the ansatz

$$\Sigma_{\text{DMFT}}[G] = \bigoplus_{p=1}^P \Sigma^{(p)}[G_{\mathcal{C}^{(p)}}; \hat{V}^{(p)}]$$

for the self-energy functional, where  $\Sigma^{(p)}$  is the exact self-energy functional for the cluster with interaction  $\hat{V}^{(p)}$ . The well-posedness of this formulation, however, requires a rigorous construction of the fermionic Luttinger-Ward functional, which is a topic of current debate; see [54, 32, 103, 42]. In the simpler setting of the Euclidean lattice field theory, see Part III for a rigorous construction of the Luttinger-Ward functional, which in turn suggests a formal analog of DMFT in that setting, as described in section 2 above.

### 3.3.2 The DMFT loop

Now we outline the ‘DMFT loop,’ an iterative algorithm, which, if convergent, provides a solution to the self-consistency equations (3.6).

As previously mentioned, the hybridizations  $\Delta^{(p)}$  are the unknowns, and an iteration of our loop amounts to a map from our current iterative guesses  $\{\Delta^{(p)}\}$  to a new set of guesses  $\{\Delta_{\text{new}}^{(p)}\}$ . Inductively assuming that  $\Delta^{(p)} \in \mathcal{S}_+^{L^{(p)}}$  for  $p = 1, \dots, P$ , we need only describe how to compute the the subsequent iteration’s hybridizations  $\Delta_{\text{new}}^{(p)}$ . Motivated by (3.7), we simply define

$$\Delta_{\text{new}}^{(p)}(z) := z - h^{(p)} - \Sigma^{(p)}[\Delta^{(p)}](z) - (G^{(p)}[\{\Delta^{(q)}\}](z))^{-1}. \quad (3.8)$$



If necessary for convergence, a mixing scheme may be employed across iterations. For example, in simple mixing we would instead define

$$\Delta_{\text{new}}^{(p)}(z) := (1 - \alpha)\Delta^{(p)}(z) + \alpha \left[ z - h^{(p)} - \Sigma^{(p)}(z) - (G^{(p)}[\{\Delta^{(q)}\}](z))^{-1} \right].$$

For simplicity we will stick to the case  $\alpha = 1$ , though generalization of our discussion to arbitrary  $\alpha$  is straightforward.

Now it is trivial by construction that if  $\Delta_{\text{new}}^{(p)} = \Delta^{(p)}$  for all  $p$  (i.e., if the loop has converged), then we have solved the self-consistency equations (3.7). But it is then nontrivial that  $\Delta_{\text{new}}^{(p)} \in \mathcal{S}_+^{L(p)}$  for all  $p$ , as is required to implement the next iteration via the identification of each impurity problem with a composite fragment/bath system. It turns out that in fact it *is* true (Corollary 20 below) that the  $\Delta_{\text{new}}^{(p)}$  defined as in (3.8) lie in  $\mathcal{S}_+^{L(p)}$ , and we call this point the ‘well-posedness of the DMFT loop.’ See [52] for related work.

### 3.3.3 The DMFT loop with bath fitting

Although  $\Delta_{\text{new}}^{(p)}$  defined as in (3.8) does indeed lie in  $\mathcal{S}_+^{L(p)}$ , the effective bath size induced by these hybridization iterates may overwhelm the solvers used in practice for the resulting composite fragment/bath systems. The reader should note that Monte Carlo schemes (such as in [41]) can avoid explicit construction of any fragment/bath system, but for so-called Hamiltonian-based DMFT schemes, in which the fragment/bath system is constructed explicitly, it is necessary to introduce an additional step into the iteration. That is, one approximates  $\Delta_{\text{new}}^{(p)}$  with  $\Delta_{\text{approx}}^{(p)}$  of reduced effective bath size as in section (3.2.1). This requires one to compute samples of  $\Delta_{\text{new}}^{(p)}$  at several points on the complex plane.

### 3.3.4 The DMFT loop with chain truncation

An alternative approach makes use of the chain Hamiltonian perspective of section 3.2.2. Note that although it is obvious how to compute samples of  $\Delta_{\text{new}}^{(p)}(z)$  in terms of the values  $\Sigma^{(p)}(z)$  computed directly from the impurity problems, it is *not* immediately obvious how to compute the quantities  $X_k \succeq 0$  and  $\varepsilon_k \in \mathbb{R}$  appearing in the exact representation

$$\Delta_{\text{new}}^{(p)} = \sum_k \frac{X_k^{(p)}}{z - \varepsilon_k^{(p)}}.$$

We shall see that a function of this form admits expression as a matrix continued fraction as in

$$\Delta_{\text{new}}^{(p)}(z) = D_0^{(p)} \frac{1}{z - E_0^{(p)} - D_1^{(p)} \frac{1}{z - E_1^{(p)} - D_2^{(p)} \frac{1}{z - E_2^{(p)} - \dots}} D_1^{(p)*}} D_0^{(p)*},$$

where the continued fraction terminates at some finite, though potentially large, depth  $n$ , i.e., terminates in the expression  $z - E_{n-1}^{(p)}$ . Note that by repeated application of the Schur complement theorem, we have

$$[h^{(p)} + \Delta_{\text{new}}^{(p)}(z)]^{-1} = \left[ \begin{pmatrix} h^{(p)} & D_0^{(p)} & & & & \\ D_0^{(p)*} & E_0^{(p)} & D_1^{(p)} & & & \\ & \ddots & \ddots & \ddots & & \\ & & D_{n-2}^{(p)*} & E_{n-2}^{(p)} & D_{n-1}^{(p)} & \\ & & & D_{n-1}^{(p)*} & E_{n-1}^{(p)} & \end{pmatrix}^{-1} \right]_{11},$$

hence  $\Delta^{(p)}(z)$  is the Hybridization of an impurity problem specified by the single-particle matrix

$$\tilde{h}^{(p)} := \begin{pmatrix} h^{(p)} & D_0^{(p)} & & & & \\ D_0^{(p)*} & E_0^{(p)} & D_1^{(p)} & & & \\ & \ddots & \ddots & \ddots & & \\ & & D_{n-2}^{(p)*} & E_{n-2}^{(p)} & D_{n-1}^{(p)} & \\ & & & D_{n-1}^{(p)*} & E_{n-1}^{(p)} & \end{pmatrix}$$

for the composite fragment/bath system and fragment interaction  $\hat{V}^{(p)}$ . Hence these data specify a chain Hamiltonian for the impurity problem in the sense of section 3.2.

Likewise, if we can compute the continued fraction matrices  $D_k^{(p)}, E_k^{(p)}$  only up to a given depth  $n' < n$ , then these data *exactly* determine the truncation of the chain Hamiltonian at depth  $n'$ . Truncating these chain Hamiltonians amounts to taking the hybridization approximation  $\Delta_{\text{approx}}^{(p)}$  that exactly matches the continued fraction of  $\Delta_{\text{new}}^{(p)}$  up to its depth  $n'$ .

## 4 Mathematical structures

### 4.1 Function spaces of interest

To begin we set the notations  $\mathbf{H}^L$ ,  $\mathbf{H}_+^L$ , and  $\mathbf{H}_{++}^L$  for  $L \times L$  Hermitian, Hermitian positive semidefinite, and Hermitian positive definite matrices, respectively. We also write  $A \succeq 0$  and  $A \succ 0$  for  $A \in \mathbf{H}_+^L$  and  $A \in \mathbf{H}_{++}^L$ , respectively.

Now the discussion of section 3.2 motivates the consideration of the class  $\mathcal{S}_+^L$  of functions  $\mathbb{C} \rightarrow \mathbb{C}^{L \times L}$  which are *Stieltjes transforms of discrete positive-operator-valued measures (POVM) on the real line*; more precisely

$$\mathcal{S}_+^L = \left\{ f : \mathbb{C} \rightarrow \mathbb{C}^{L \times L} \mid f(z) = \sum_{k=1}^K \frac{X_k}{z - \varepsilon_k}, X_k \succeq 0, \varepsilon_k \in \mathbb{R} \right\}.$$

In the following the dependence on  $L$  will often be omitted for simplicity, i.e., we shall simply write  $\mathcal{S}_+^L$  when the meaning is clear from context. Note that  $\mathcal{S}_+^L$  is the class of functions that can be realized as impurity hybridizations and is thus worthy of our attention. Note that  $\mathcal{S}_+^L$  is closed under addition and nonnegative scalar multiplication. (Hence  $\mathcal{S}_+^L$  is a convex cone.) Furthermore  $\mathcal{S}_+^L$  is closed under congruence transformations, i.e., transformations of the form  $f \mapsto AfA^*$  for fixed  $A$ . Moreover, if  $f \in \mathcal{S}_+^L$ , then  $z \mapsto \overline{f(\bar{z})}$  is also in  $\mathcal{S}_+^L$ .

To motivate the italicized long-form name for  $\mathcal{S}_+^L$ , recall that POVMs on  $\mathbb{R}$  generalize ordinary measures to the operator-valued case, where the space of operators is determined by a choice of Hilbert space (here simply  $\mathbb{C}^L$ ). Heuristically POVMs can be thought of as positive-semidefinite-valued densities  $d\mathbf{M}(\varepsilon) = X(\varepsilon) d\varepsilon$ , with the discrete case recovered as  $\mathbf{M} = \sum_k X_k \delta(\cdot - \varepsilon_k)$ , where  $X_k \succeq 0$  and  $\delta$  indicates the Dirac delta. As we shall see, it is natural to consider the subclass  $\mathcal{S}_1^L$  induced by discrete ‘probability’ POVMs, i.e., POVMs satisfying  $\int_{\mathbb{R}} d\mathbf{M}(\varepsilon) = \text{Id}$ , defined precisely by

$$\mathcal{S}_1^L := \left\{ f : \mathbb{C} \rightarrow \mathbb{C}^{L \times L} \mid f(z) = \sum_{k=1}^K \frac{X_k}{z - \varepsilon_k}, X_k \succeq 0, \sum_{k=1}^K X_k = I_L, \varepsilon_k \in \mathbb{R} \right\}.$$

Note that the closure of  $\mathcal{S}_1^L$  under congruence transformations by invertible matrices  $A$  (the matrix analog of scaling) is the class

$$\mathcal{S}_{++}^L := \left\{ f : \mathbb{C} \rightarrow \mathbb{C}^{L \times L} \mid f(z) = \sum_{k=1}^K \frac{X_k}{z - \varepsilon_k}, X_k \succeq 0, \sum_{k=1}^K X_k \succ 0, \varepsilon_k \in \mathbb{R} \right\}.$$

Motivated by the appearance in section 3.2 above of the expression  $z - h - \Delta(z)$ , we also consider the classes

$$\mathcal{N}_+^L := \{g : \mathbb{C} \rightarrow \mathbb{C}^{L \times L} : g(z) = zB - C - f(z), B \in \mathbf{H}_+^L, C \in \mathbf{H}^L, f \in \mathcal{S}_+^L\},$$

$$\mathcal{N}_1^L := \{g : \mathbb{C} \rightarrow \mathbb{C}^{L \times L} : g(z) = z - C - f(z), A \in \mathbf{H}_+^L, B \in \mathbf{H}^L, f \in \mathcal{S}_+^L\},$$

$$\mathcal{N}_{++}^L := \{g : \mathbb{C} \rightarrow \mathbb{C}^{L \times L} : g(z) = zB - C - f(z), B \in \mathbf{H}_{++}^L, C \in \mathbf{H}^L, f \in \mathcal{S}_+^L\}.$$

We shall prove below (Lemma 16) that  $f \in \mathcal{S}_{++}^L$  if and only if the pointwise inverse  $f^{-1} \in \mathcal{N}_{++}^L$ , so we write  $(\mathcal{S}_{++}^L)^{-1} = \mathcal{N}_{++}^L$ . Similarly,  $(\mathcal{S}_1^L)^{-1} = \mathcal{N}_1^L$ . Note moreover that  $\mathcal{S}_+^L \subset -\mathcal{N}_+^L$ .

Moreover, as we shall see confirm below (Lemma 12), for any  $f \in \mathcal{S}_+$ ,  $\Im[f(z)] \leq 0$  for all  $z \in \mathbb{H}_+$ , where  $\mathbb{H}_+$  is the (strict) upper complex half-plane and  $\Im$  indicates the imaginary part of a matrix, i.e.,  $\Im[A] = \frac{1}{2i}(A - A^\dagger)$ . It follows easily then that for  $g \in \mathcal{N}_+$ ,  $\Im[g(z)] \geq 0$  for all  $z \in \mathbb{H}_+$ . In fact, this fact can roughly be seen as a defining property for these spaces, in a sense that we shall describe shortly. In light of this observation, together with the elementary fact that  $\Im[A] \geq 0$  if and only if  $\Im[A^{-1}] \leq 0$  (for all invertible  $A$ ), the pointwise inversion relation is quite fundamental.

Now for broader context, the class of ( $\mathbb{C}^{L \times L}$ -valued) *Nevanlinna functions* is defined by

$$\overline{\mathcal{N}}_+^L := \left\{ g : \mathbb{H}_+ \rightarrow \mathbb{C}^{L \times L} \text{ analytic} \mid \Im[g(z)] \succeq 0 \text{ for all } z \in \mathbb{H}_+ \right\}.$$

Note that this notation is not standard but rather is chosen to be suggestive in the current context. Indeed, it can be shown [31] that any  $g \in \overline{\mathcal{N}}_+^L$  can be written

$$g(z) = Bz + C - \int_{\mathbb{R}} \left( \frac{1}{z - \varepsilon} + \frac{\varepsilon}{1 + \varepsilon^2} \right) d\mathbf{M}(\varepsilon),$$

where  $B \in \mathbf{H}_+^L$ ,  $C \in \mathbf{H}^L$ , and  $\mathbf{M}$  is a POVM on  $\mathbb{R}$  such that  $\int_{\mathbb{R}} \frac{d\mathbf{M}(\varepsilon)}{1 + \varepsilon^2}$  is convergent. Hence  $\overline{\mathcal{N}}_+^L$  can in fact be thought of as the closure (taken in a suitable sense) of  $\mathcal{N}_+^L$ .

Meanwhile, one can recover the closure  $\overline{\mathcal{S}}_+^L \subset \overline{\mathcal{N}}_+^L$  via

$$\overline{\mathcal{S}}_+^L := \left\{ -g \mid g \in \overline{\mathcal{N}}_+^L, \lim_{|b| \rightarrow \infty} f(ib) = 0 \right\},$$

and any  $f \in \overline{\mathcal{S}}_+^L$  admits the representation

$$f(z) = \int_{\mathbb{R}} \left( \frac{1}{z - \varepsilon} + \frac{\varepsilon}{1 + \varepsilon^2} \right) d\mathbf{M}(\varepsilon),$$

for  $\mathbf{M}$  a POVM on  $\mathbb{R}$  such that  $\int_{\mathbb{R}} \frac{d\mathbf{M}(\varepsilon)}{1 + \varepsilon^2}$  is convergent.

In this work, we will only need to work with functions in  $\mathcal{S}_+^L$  and  $\mathcal{N}_+^L$  specified by discrete POVMs; however, the number of poles needed to specify such functions may be large, so it is natural to think of them as approximating functions in the closure.

## 4.2 Structure of Green's functions

Though  $\mathcal{S}$  was introduced as the class of physical hybridizations, in fact Green's functions, as defined in section 3.1, all lie in  $\mathcal{S}_1 \subset \mathcal{S}$ , as we shall now see. This is essentially a recapitulation of the Lehmann representation (see, e.g., [77]).

**Lemma 9** (Lehmann). *If  $G$  is a Green's function in the sense of Definition 1 (zero temperature) or 3 (finite temperature), then  $G \in \mathcal{S}_1$ .*

*Remark 10.* In the zero-temperature case, assuming for simplicity a nondegenerate ground state  $|\Psi_0^N\rangle$ , we derive

$$G(z) = \sum_{\varepsilon \in \sigma(\hat{H}_{N-1})} \frac{X_\varepsilon^{(-)}}{z + (\varepsilon - E_0)} + \sum_{\varepsilon \in \sigma(\hat{H}_{N+1})} \frac{X_\varepsilon^{(+)}}{z - (\varepsilon - E_0)},$$

where

$$X_{\varepsilon,ij}^{(-)} = \langle \Psi_0^N | a_j^\dagger P_\varepsilon a_i | \Psi_0^N \rangle, \quad X_{\varepsilon,ij}^{(+)} = \langle \Psi_0^N | a_i P_\varepsilon a_j^\dagger | \Psi_0^N \rangle$$

In the finite-temperature case, we derive

$$G(z) : = \sum_{\varepsilon, \lambda \in \sigma(\hat{H})} \frac{X_{\varepsilon\lambda}^{(-)}}{z - (\varepsilon - \lambda)} + \sum_{\varepsilon, \lambda \in \sigma(\hat{H})} \frac{X_{\varepsilon\lambda}^{(+)}}{z + (\varepsilon - \lambda)},$$

where

$$X_{\varepsilon\lambda, ij}^{(-)} = \frac{1}{Z} e^{-\beta\lambda} \text{Tr} \left[ a_i P_\varepsilon a_j^\dagger P_\lambda \right], \quad X_{\varepsilon\lambda, ij}^{(+)} = \frac{1}{Z} e^{-\beta\lambda} \text{Tr} \left[ a_j^\dagger P_\varepsilon a_i P_\lambda \right].$$

*Proof.* Consider an arbitrary many-body Hamiltonian  $\hat{H}$ , i.e., an operator on  $\mathcal{F}$ , and recall from Definition 1 that the zero-temperature Green's function  $G : \mathbb{C} \rightarrow \mathbb{C}^{M \times M}$  is given by

$$G_{ij}(z) := \langle \Psi_0^N | a_j^\dagger (z + [\hat{H} - E_0])^{-1} a_i | \Psi_0^N \rangle + \langle \Psi_0^N | a_i (z - [\hat{H} - E_0])^{-1} a_j^\dagger | \Psi_0^N \rangle,$$

where for simplicity of presentation we assume a nondegenerate ground state  $|\Psi_0^N\rangle$ .

Index the eigenvalues of  $H$  by  $\varepsilon$ , and let  $P_\varepsilon$  be the  $\varepsilon$ -eigenspace of  $H$ . Then we can decompose

$$\hat{H} = \sum_{\varepsilon \in \sigma(\hat{H})} \varepsilon P_\varepsilon,$$

where  $\sigma(\hat{H})$  is the spectrum of  $H$ . Moreover,

$$(z \pm [\hat{H} - E_0])^{-1} = \sum_{\varepsilon \in \sigma(\hat{H})} \frac{1}{z \pm (\varepsilon - E_0)} P_\varepsilon,$$

Evidently the poles of  $G$  are given by  $\varepsilon \in \sigma(\hat{H}_{N\pm 1}) \subset \sigma(\hat{H})$ , where  $\hat{H}_{N\pm 1}$  is the restriction of  $H$  to the  $N \pm 1$ -particle space.

Then we can write

$$\begin{aligned} G_{ij}(z) &= \sum_{\varepsilon \in \sigma(\hat{H}_{N-1})} \frac{1}{z + (\varepsilon - E_0)} \langle \Psi_0^N | a_j^\dagger P_\varepsilon a_i | \Psi_0^N \rangle \\ &\quad + \sum_{\varepsilon \in \sigma(\hat{H}_{N+1})} \frac{1}{z - (\varepsilon - E_0)} \langle \Psi_0^N | a_i P_\varepsilon a_j^\dagger | \Psi_0^N \rangle, \end{aligned}$$

or

$$G(z) = \sum_{\varepsilon \in \sigma(\hat{H}_{N-1})} \frac{X_\varepsilon^{(-)}}{z + (\varepsilon - E_0)} + \sum_{\varepsilon \in \sigma(\hat{H}_{N+1})} \frac{X_\varepsilon^{(+)}}{z - (\varepsilon - E_0)},$$

where

$$X_{\varepsilon, ij}^{(-)} = \langle \Psi_0^N | a_j^\dagger P_\varepsilon a_i | \Psi_0^N \rangle, \quad X_{\varepsilon, ij}^{(+)} = \langle \Psi_0^N | a_i P_\varepsilon a_j^\dagger | \Psi_0^N \rangle.$$

Now

$$\begin{aligned}
\langle \Psi_0^N | a_j^\dagger P_\varepsilon a_i | \Psi_0^N \rangle &= \text{Tr} \left[ |\Psi_0^N\rangle \langle \Psi_0^N | a_j^\dagger P_\varepsilon a_i \right] \\
&= \overline{\text{Tr} \left[ \left( |\Psi_0^N\rangle \langle \Psi_0^N | a_j^\dagger P_\varepsilon a_i \right)^\dagger \right]} \\
&= \overline{\text{Tr} \left[ a_i^\dagger P_\varepsilon a_j | \Psi_0^N \rangle \langle \Psi_0^N | \right]} \\
&= \overline{\langle \Psi_0^N | a_i^\dagger P_\varepsilon a_j | \Psi_0^N \rangle},
\end{aligned}$$

i.e.,

$$X_{\varepsilon,ij}^{(-)} = \overline{X_{\varepsilon,ji}^{(-)}},$$

and by similar reasoning  $X_{\varepsilon,ij}^{(+)} = \overline{X_{\varepsilon,ji}^{(+)}}$ , i.e.,  $X_\varepsilon^{(\pm)}$  is Hermitian.

In fact,  $X_\varepsilon^{(\pm)} \succeq 0$ . Indeed, let  $v \in \mathbb{C}^N$ . Then

$$\begin{aligned}
v^* X_\varepsilon^{(-)} v &= \sum_{ij} \langle \Psi_0^N | (v_j a_j)^\dagger P_\varepsilon (v_i a_i) | \Psi_0^N \rangle \\
&= \langle \Psi_0^N | c^\dagger P_\varepsilon c | \Psi_0^N \rangle,
\end{aligned}$$

where  $c := \sum_j v_j a_j$ . But  $P_\varepsilon \succeq 0$ , so  $c^\dagger P_\varepsilon c \succeq 0$ , and  $v^* X_\varepsilon^{(-)} v \geq 0$ . Similarly,

$$\begin{aligned}
v^* X_\varepsilon^{(+)} v &= \sum_{ij} \langle \Psi_0^N | (v_i a_i^\dagger)^\dagger P_\varepsilon (v_j a_j^\dagger) | \Psi_0^N \rangle \\
&= \langle \Psi_0^N | d^\dagger P_\varepsilon d | \Psi_0^N \rangle,
\end{aligned}$$

where  $d := \sum_j v_j a_j^\dagger$ . Thus  $X_\varepsilon^{(\pm)} \succeq 0$ , as claimed. It follows that  $G \in \mathcal{S}_+$ .

Moreover, the sum of the residues of  $G$  is equal to the identity. Indeed, this sum is given by

$$\begin{aligned}
&\sum_{\varepsilon \in \sigma(\hat{H}_{N-1})} X_{\varepsilon,ij}^{(-)} + \sum_{\varepsilon \in \sigma(\hat{H}_{N+1})} X_{\varepsilon,ij}^{(+)} \\
&= \sum_{\varepsilon \in \sigma(\hat{H}_{N-1})} \langle \Psi_0^N | a_j^\dagger P_\varepsilon a_i | \Psi_0^N \rangle + \sum_{\varepsilon \in \sigma(\hat{H}_{N+1})} \langle \Psi_0^N | a_i P_\varepsilon a_j^\dagger | \Psi_0^N \rangle \\
&= \sum_{\varepsilon \in \sigma(\hat{H})} \langle \Psi_0^N | a_j^\dagger P_\varepsilon a_i | \Psi_0^N \rangle + \sum_{\varepsilon \in \sigma(\hat{H})} \langle \Psi_0^N | a_i P_\varepsilon a_j^\dagger | \Psi_0^N \rangle \\
&= \left\langle \Psi_0^N \left| a_j^\dagger \left( \sum_{\varepsilon \in \sigma(\hat{H})} P_\varepsilon \right) a_i \right| \Psi_0^N \right\rangle + \left\langle \Psi_0^N \left| a_i \left( \sum_{\varepsilon \in \sigma(\hat{H})} P_\varepsilon \right) a_j^\dagger \right| \Psi_0^N \right\rangle \\
&= \langle \Psi_0^N | a_j^\dagger a_i | \Psi_0^N \rangle + \langle \Psi_0^N | a_i a_j^\dagger | \Psi_0^N \rangle \\
&= \langle \Psi_0^N | a_j^\dagger a_i + a_i a_j^\dagger | \Psi_0^N \rangle
\end{aligned}$$

$$\begin{aligned}
&= \langle \Psi_0^N | \delta_{ij} | \Psi_0^N \rangle \\
&= \delta_{ij}.
\end{aligned}$$

Similarly, in the finite temperature case, we have

$$G(z) := \sum_{\varepsilon, \lambda \in \sigma(\hat{H})} \frac{X_{\varepsilon\lambda}^{(-)}}{z + (\varepsilon - \lambda)} + \sum_{\varepsilon, \lambda \in \sigma(\hat{H})} \frac{X_{\varepsilon\lambda}^{(+)}}{z - (\varepsilon - \lambda)},$$

where

$$X_{\varepsilon\lambda, ij}^{(-)} = \frac{1}{Z} e^{-\beta\lambda} \text{Tr} \left[ a_j^\dagger P_\varepsilon a_i P_\lambda \right], \quad X_{\varepsilon\lambda, ij}^{(+)} = \frac{1}{Z} e^{-\beta\lambda} \text{Tr} \left[ a_i P_\varepsilon a_j^\dagger P_\lambda \right].$$

One can likewise show that  $X_{\varepsilon\lambda}^{(\pm)} \succeq 0$ . □

### 4.3 Pointwise inversion relations

We would like to study the pointwise inverses of functions in  $\mathcal{S}_+^L$ . For  $f \in \mathcal{S}_+^L$ , write

$$f(z) = \sum_{k=1}^K \frac{X_k}{z - \varepsilon_k}$$

and consider

$$f^{-1}(z) := [f(z)]^{-1}.$$

Notice that if  $\bigcap_{k=1}^K \ker(X_k)$  is nontrivial (which holds in particular if  $\sum_k \text{rank}(X_k) < L$ ), then  $f^{-1}$  exists nowhere. Thus we assume that  $\bigcap_{k=1}^K \ker(X_k) = \{0\}$ . Equivalently, we assume that  $\sum_k X_k \succ 0$ , i.e., we restrict our attention to  $f \in \mathcal{S}_{++}^L$ .

Now observe that  $f^{-1}$  is a rational matrix-valued function. Thus the holomorphic part  $h$  (i.e., whatever is left after subtracting off the poles) is polynomial. Since

$$f(z) \sim \frac{1}{z} \sum_{k=1}^K X_k$$

as  $|z| \rightarrow \infty$ , it follows that  $f^{-1}(z) = O(z)$  as  $z \rightarrow \infty$ . Thus  $h$  is affine, and in fact we have

$$h(z) = z \left( \sum_{k=1}^K X_k \right)^{-1} - C,$$

where  $C \in \mathbb{C}^{N \times N}$  is a constant matrix. In fact  $C$  is Hermitian (because  $f$  — and hence  $f^{-1}$  as well — is Hermitian on real axis), and we will provide an explicit formula later on. Now it will further turn out that  $h - f^{-1} \in \mathcal{S}_{++}^L$ . The first step in establishing this is the following:

**Lemma 11.** *Suppose that  $f \in \mathcal{S}_{++}^L$ . Then all of the poles of  $f^{-1}(z)$  lie on the real axis.*

*Proof.* Also notice that for  $z = x + iy \in \mathbb{C}$ ,  $\Im[z^{-1}] = -\frac{y}{x^2+y^2}$ , so  $\Im[z^{-1}] = 0$  if and only if  $z \in \mathbb{R}$ , and moreover,  $\Im[z^{-1}] < 0$  for  $z \in \mathbb{H}_+$  (the upper half-plane) and  $\Im[z^{-1}] > 0$  for  $z \in \mathbb{H}_-$ .

Suppose  $z \in \mathbb{H}_+$ , and let  $v \in \mathbb{C}^L$  be nonzero. Then

$$v^* f(z) v = \sum_{k=1}^K \frac{v^* X_k v}{z - \varepsilon_k} = \sum_{k=1}^K \frac{a_k}{z - \varepsilon_k},$$

where  $a_k \geq 0$ . Now

$$\Im[v^* f(z) v] = \sum_{k=1}^K a_k \Im[(z - \varepsilon_k)^{-1}].$$

Now  $z \in \mathbb{H}_+$  implies that  $z - \varepsilon_k \in \mathbb{H}_+$ , so  $\Im[(z - \varepsilon_k)^{-1}] < 0$ , and

$$\Im[v^* f(z) v] \leq 0$$

with equality if and only if  $a_k = 0$  for all  $k$ . Since  $\bigcap_{k=1}^K \ker(X_k) = \{0\}$  and  $v \neq 0$ , we must have  $a_k = v^* X_k v > 0$  for at least one  $k$ . This implies in particular that  $v^* f(z) v \neq 0$  for all  $z \in \mathbb{H}_+$  and all nonzero  $v \in \mathbb{C}^L$ , i.e., that  $f(z)$  is invertible for  $z \in \mathbb{H}_+$ . Similar reasoning shows that  $f(z)$  is invertible for  $z \in \mathbb{H}_-$ . Thus the only poles of  $f^{-1}$  are on the real axis.  $\square$

Note that from the proof of Lemma 11, we can further conclude:

**Lemma 12.** *Suppose that  $f \in \mathcal{S}_+^L$ . Then  $\Im[f(z)] \preceq 0$  for  $z \in \mathbb{H}_+$  and  $\Im[f(z)] \succeq 0$  for  $z \in \mathbb{H}_-$ . If  $f \in \mathcal{S}_{++}^L$ , then we have  $\Im[f(z)] \prec 0$  and  $\Im[f(z)] \succ 0$  in these cases, respectively.*

We go on to show that the poles of  $f^{-1}$  have residues of the correct sign:

**Lemma 13.** *Suppose that  $f \in \mathcal{S}_{++}^L$ . The poles of  $f^{-1}$  are simple with negative semidefinite residues.*

*Proof.* Let  $v \in \mathbb{C}^L$  be nonzero. Then

$$v^* f'(z) v = - \sum_{k=1}^K \frac{v^* X_k v}{(z - \varepsilon_k)^2},$$

so — by the fact that at least one the  $v^* X_k v$  is strictly positive — we have  $v^* f'(z) v < 0$  everywhere on the real axis, with infinite value at the poles. (Recall from the previous lemma that all of the poles of  $f^{-1}$  lie only on the real axis.) This means in particular that  $f'$  is negative definite on the real axis (except the poles). (Note that, on the real axis,  $f$  and  $f'$  take Hermitian values.)

Then consider  $z_0 \in \mathbb{R}$  which is a pole, i.e., for which  $f(z_0)$  has a zero eigenvalue. (Generically one expects  $f(z_0)$  to have only a single zero eigenvalue at such points,



but we need not assume this.) By shifting if necessary we can assume for simplicity that  $z_0 = 0$ , and by changing basis if necessary, we can assume that  $f(0)$  has the block structure

$$f(0) = \begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix},$$

where  $D$  is diagonal with nonzero diagonal entries. Further write

$$f'(0) =: A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix},$$

so  $A \prec 0$ . Then Taylor's theorem gives

$$f(z) = f(0) + zA + O(|z|^2) = \begin{pmatrix} D + zA_{11} & zA_{12} \\ zA_{21} & zA_{22} \end{pmatrix} + O(|z|^2).$$

Since  $A \prec 0$ ,  $f(0) + zA$  — and hence also  $f(z)$  — is nonsingular for all  $z \neq 0$  sufficiently small. Then for such  $z$ , by the Schur complement theorem we can see that

$$f^{-1}(z) = \begin{pmatrix} D^{-1} & -D^{-1}A_{12}A_{22}^{-1} \\ -A_{22}^{-1}A_{21}D^{-1} & z^{-1}A_{22}^{-1} + A_{22}^{-1}A_{21}D^{-1}A_{12}A_{22}^{-1} \end{pmatrix} + O(|z|).$$

In particular, there exists a holomorphic function  $g(z)$  such that

$$f^{-1}(z) - g(z) = \frac{1}{z} \begin{pmatrix} 0 & 0 \\ 0 & A_{22}^{-1} \end{pmatrix}.$$

Thus the pole at 0 is simple with negative semidefinite residue. (Generically one expects the rank of the residue to be 1.)  $\square$

We collect our results into the following proposition:

**Lemma 14.** *Let  $f \in \mathcal{S}_{++}^L$ , written*

$$f(z) = \sum_{k=1}^K \frac{X_k}{z - \varepsilon_k}.$$

*Then the holomorphic part of  $f^{-1}$  is given by  $h(z) = zB - C$ , where*

$$B = \left( \sum_{k=1}^K X_k \right)^{-1}, \quad C = B \left( \sum_{k=1}^K \varepsilon_k X_k \right) B,$$

*and moreover  $h - f^{-1} \in \mathcal{S}_+^L$ , i.e.,  $f^{-1} \in \mathcal{N}_{++}^L$ .*

*Proof.* The only part of the proposition that we have not already established is the explicit formula for the constant matrix  $C$ .

Since the holomorphic part of  $f^{-1}$  is given by  $zB - C$ , note that

$$\frac{d}{dz}f^{-1}(z) = B + O(|z|^{-2})$$

as  $|z| \rightarrow \infty$ . (Note that here all limits are restricted to  $z \in \mathbb{R}$ .) Therefore

$$zB + O(|z|^{-1}) = z \frac{d}{dz}f^{-1}(z) = -zf^{-1}(z)f'(z)f^{-1}(z).$$

Now

$$f^{-1}(z) = zB - C + O(|z|^{-1}),$$

so

$$\begin{aligned} C &= - \lim_{|z| \rightarrow \infty} [f^{-1}(z) + zf^{-1}(z)f'(z)f^{-1}(z)] \\ &= - \lim_{|z| \rightarrow \infty} [f^{-1}(z) [f(z) + zf'(z)] f^{-1}(z)]. \end{aligned}$$

Now

$$f(z) + zf'(z) = \sum_{k=1}^K \left( \frac{1}{z - \varepsilon_k} - \frac{z}{(z - \varepsilon_k)^2} \right) X_k = - \sum_{k=1}^K \frac{\varepsilon_k X_k}{(z - \varepsilon_k)^2}.$$

Then we can rewrite

$$C = - \lim_{|z| \rightarrow \infty} \left[ \frac{f^{-1}(z)}{z} \cdot z^2 [f(z) + zf'(z)] \frac{f^{-1}(z)}{z} \right]$$

and observe that

$$\frac{f^{-1}(z)}{z} \rightarrow B, \quad z^2 [f(z) + zf'(z)] \rightarrow - \sum_{k=1}^K \varepsilon X_k$$

as  $|z| \rightarrow \infty$ . Therefore  $C = B \left( \sum_{k=1}^K \varepsilon X_k \right) B$ , as was to be shown.  $\square$

Lastly, we show a converse:

**Lemma 15.** *Let  $g(z) = zB - C - f(z)$ , where  $f \in \mathcal{S}_+^L$ ,  $B \succ 0$ , and  $C \in \mathbf{H}^L$ . Then  $g^{-1} \in \mathcal{S}_{++}^L$ .*

*Proof.* First we claim that  $g$  is nonsingular outside of the real axis. By Lemma 12,  $\Im[f] \preceq 0$  on  $\mathbb{H}_+$  and  $\Im[f] \succeq 0$  on  $\mathbb{H}_-$ . Since  $C$  is Hermitian and  $B \succeq 0$ , it follows that  $\Im[g] \prec 0$  and  $\Im[g] \succ 0$  in these cases, respectively. In particular,  $g$  is nonsingular outside of the real axis, as claimed, and the poles of  $g^{-1}$  lie only on the real axis.

Next notice that since  $f(z) \rightarrow 0$  as  $|z| \rightarrow \infty$  and  $|z|B \rightarrow +\infty$  as  $|z| \rightarrow \infty$ , it follows that  $g^{-1}(z) \rightarrow 0$  as  $|z| \rightarrow \infty$ . Therefore the holomorphic part of the rational function  $g^{-1}$  must be zero.

It remains only to show that the poles of  $g^{-1}$  are simple with positive semidefinite residue. But by the reasoning of Lemma 13, it suffices to show that  $g' \succ 0$  (away from its poles). But of course  $g' = B - f'$ , and by the reasoning of Lemma 13,  $-f' \succ 0$  away from the poles. This completes the proof.  $\square$

In particular, Lemmas 14 and 15 imply:

**Lemma 16.**  $f \in \mathcal{S}_{++}^L$  if and only if  $f^{-1} \in \mathcal{N}_{++}^L$ . Consequently  $f \in \mathcal{S}_1^L$  if and only if  $f^{-1} \in \mathcal{N}_1^L$ .

#### 4.4 Structure of the self-energy and the Hartree-Fock contribution

We wish to understand the structure of the self-energy a little more carefully. Recall that the self-energy is defined by

$$\Sigma(z) = z - h - G^{-1}(z).$$

For simplicity we discuss only the zero-temperature case and assume a nondegenerate ground state  $|\Phi_0^N\rangle$ . Since  $G \in \mathcal{S}_1$ , it follows from Lemma 14 and Remark 10 that  $G^{-1}(z) = z - C - f(z)$ , where  $f \in \mathcal{S}_+$  and

$$\begin{aligned} C_{ij} &= \sum_{\varepsilon \in \sigma(\hat{H}_{N+1})} (\varepsilon - E_0) \langle \Psi_0^N | a_i P_\varepsilon a_j^\dagger | \Psi_0^N \rangle - \sum_{\varepsilon \in \sigma(\hat{H}_{N-1})} (\varepsilon - E_0) \langle \Psi_0^N | a_j^\dagger P_\varepsilon a_i | \Psi_0^N \rangle \\ &= \sum_{\varepsilon \in \sigma(\hat{H})} (\varepsilon - E_0) \langle \Psi_0^N | a_i P_\varepsilon a_j^\dagger | \Psi_0^N \rangle - \sum_{\varepsilon \in \sigma(\hat{H})} (\varepsilon - E_0) \langle \Psi_0^N | a_j^\dagger P_\varepsilon a_i | \Psi_0^N \rangle \\ &= \left\langle \Psi_0^N \left| a_i \left( \sum_{\varepsilon \in \sigma(\hat{H})} (\varepsilon - E_0) P_\varepsilon \right) a_j^\dagger \right| \Psi_0^N \right\rangle \\ &\quad - \left\langle \Psi_0^N \left| a_j^\dagger \left( \sum_{\varepsilon \in \sigma(\hat{H})} (\varepsilon - E_0) P_\varepsilon \right) a_i \right| \Psi_0^N \right\rangle. \end{aligned}$$

Since  $\hat{H} - E_0 = \sum_{\varepsilon \in \sigma(\hat{H})} (\varepsilon - E_0) P_\varepsilon$ , we obtain

$$C_{ij} = \langle \Psi_0^N | a_i (\hat{H} - E_0) a_j^\dagger | \Psi_0^N \rangle - \langle \Psi_0^N | a_j^\dagger (\hat{H} - E_0) a_i | \Psi_0^N \rangle. \quad (4.1)$$

Now

$$\Sigma(z) = C - h + f(z),$$

so  $\Sigma$  is the sum of a function in  $\mathcal{S}_+$  and a constant Hermitian matrix, namely  $C-h$ . We call this the *static* part of the self-energy, and we suggestively define it  $\Sigma^{\text{HF}} := C-h$ . It is natural to attempt to derive a formula for this contribution, and indeed it will turn out to be the Hartree-Fock contribution, which is the first-order contribution (and in fact the only static contribution) to the self-energy in bold diagrammatic perturbation theory (see Part II). Under the assumption that the interacting part of the Hamiltonian is given by a two-body interaction, we will now recover the standard formula for this term.

To wit, we assume that the interaction is given by

$$\hat{V} = \frac{1}{2} \sum_{pqrs} (pq|V|rs) a_p^\dagger a_q^\dagger a_s a_r,$$

and  $(pq|V|rs)$  are the tensor elements for the two-body interaction, given in the setting of electronic structure by suitable two-electron integrals. By defining antisymmetrized matrix elements

$$\{pq|V|rs\} := \frac{1}{2} [(pq|V|rs) - (pq|V|sr) - (qp|V|rs) + (qp|V|sr)],$$

we can also write

$$V = \frac{1}{4} \sum_{pqrs} \{pq|V|rs\} a_p^\dagger a_q^\dagger a_s a_r.$$

Note that

$$\{pq|V|rs\} = -\{qp|V|rs\} = -\{pq|V|sr\} = \{qp|V|sr\}.$$

**Lemma 17.** *The self-energy can be written  $\Sigma(z) = \Sigma_{\text{HF}} + f(z)$ , where  $f \in \mathcal{S}_+$ , and the static part of the self-energy, namely the Hartree-Fock contribution, is given by*

$$[\Sigma_{\text{HF}}]_{ij} = \sum_{qr} \{iq|V|jr\} D_{qr},$$

where the 1-RDM  $D = (D_{ij})$  is defined by  $D = \langle \Psi_0^N | a_i^\dagger a_j | \Psi_0^N \rangle$ .

*Remark 18.* In the case  $(pq|V|rs) = v_{pq} \delta_{pr} \delta_{qs}$  where  $v$  is real-symmetric and  $D$  is real (hence symmetric), we have

$$\{pq|V|rs\} = v_{pq} \delta_{pr} \delta_{qs} - v_{pq} \delta_{ps} \delta_{qr},$$

hence recover the perhaps more familiar expression

$$[\Sigma_{\text{HF}}]_{ij} = \delta_{ij} \sum_q v_{ip} D_{pp} - v_{ij} D_{ij}.$$

In this last expression the first term is the diagonal Hartree potential, and the second term is the Fock exchange operator.

*Proof.* To simplify our expression for  $C$ , our first goal is to commute  $\hat{H} - E_0$  past  $a_j^\dagger$  and  $a_i$  in the first and second terms of (4.1), respectively. In so doing, it will be useful to recall that standard fact

$$[\hat{H}_0, a_j^\dagger] = [a^\dagger h a, a_j^\dagger] = \sum_k h_{kj} a_k^\dagger,$$

which may be recovered easily via the canonical anticommutation relations.

In order to derive the commutator  $[\hat{V}, a_j^\dagger]$ , we likewise compute

$$\begin{aligned} a_j^\dagger \hat{V} &= \frac{1}{4} a_j^\dagger \sum_{pqrs} \{pq|V|rs\} a_p^\dagger a_q^\dagger a_s a_r \\ &= \frac{1}{4} \sum_{pqrs} \{pq|V|rs\} a_p^\dagger a_q^\dagger a_j^\dagger a_s a_r \\ &= \frac{1}{4} \sum_{pqrs} \{pq|V|rs\} a_p^\dagger a_q^\dagger (\delta_{js} a_r - a_s a_j^\dagger a_r) \\ &= \frac{1}{4} \sum_{pqrs} \{pq|V|rs\} a_p^\dagger a_q^\dagger (\delta_{js} a_r - a_s (\delta_{jr} - a_r a_j^\dagger)) \\ &= \hat{V} a_j^\dagger + \frac{1}{4} \sum_{pqrs} \{pq|V|rs\} a_p^\dagger a_q^\dagger (\delta_{js} a_r - \delta_{jr} a_s) \\ &= \hat{V} a_j^\dagger + \frac{1}{4} \sum_{pqr} \{pq|V|rj\} a_p^\dagger a_q^\dagger a_r - \sum_{pqs} \{pq|V|js\} a_p^\dagger a_q^\dagger a_s \\ &= \hat{V} a_j^\dagger + \frac{1}{4} \sum_{pqr} [\{pq|V|rj\} - \{pq|V|jr\}] a_p^\dagger a_q^\dagger a_r \\ &= \hat{V} a_j^\dagger - \frac{1}{2} \sum_{pqr} \{pq|V|jr\} a_p^\dagger a_q^\dagger a_r. \end{aligned}$$

Therefore

$$[\hat{V}, a_j^\dagger] = \frac{1}{2} \sum_{pqr} \{pq|V|jr\} a_p^\dagger a_q^\dagger a_r.$$

It follows that

$$[\hat{H}, a_j^\dagger] = \sum_k h_{kj} a_k^\dagger + \frac{1}{2} \sum_{pqr} \{pq|V|jr\} a_p^\dagger a_q^\dagger a_r.$$

Then, using the fact that  $(\hat{H} - E_0)|\Psi_0^N\rangle = 0$ , we can easily compute:

$$\begin{aligned} \langle \Psi_0^N | a_i (\hat{H} - E_0) a_j^\dagger | \Psi_0^N \rangle &= \langle \Psi_0^N | a_i a_j^\dagger (\hat{H} - E_0) | \Psi_0^N \rangle + \langle \Psi_0^N | a_i [\hat{H}, a_j^\dagger] | \Psi_0^N \rangle \\ &= \langle \Psi_0^N | a_i [\hat{H}, a_j^\dagger] | \Psi_0^N \rangle \\ &= \sum_k \langle \Psi_0^N | a_i a_k^\dagger | \Psi_0^N \rangle h_{kj} \end{aligned}$$

$$+ \frac{1}{2} \sum_{pqr} \{pq|V|jr\} \langle \Psi_0^N | a_i a_p^\dagger a_q^\dagger a_r | \Psi_0^N \rangle,$$

and

$$\begin{aligned} \langle \Psi_0^N | a_j^\dagger (\hat{H} - E_0) a_i | \Psi_0^N \rangle &= -\langle \Psi_0^N | [\hat{H}, a_j^\dagger] a_i | \Psi_0^N \rangle \\ &= -\sum_k \langle \Psi_0^N | a_k^\dagger a_i | \Psi_0^N \rangle h_{kj} \\ &\quad - \frac{1}{2} \sum_{pqr} \{pq|V|jr\} \langle \Psi_0^N | a_p^\dagger a_q^\dagger a_r a_i | \Psi_0^N \rangle. \end{aligned}$$

Then, by (4.1),

$$C_{ij} = \sum_k \delta_{ik} h_{kj} + \frac{1}{2} \sum_{pqr} \{pq|V|jr\} \langle \Psi_0^N | a_i a_p^\dagger a_q^\dagger a_r + a_p^\dagger a_q^\dagger a_r a_i | \Psi_0^N \rangle.$$

Now

$$\begin{aligned} a_i a_p^\dagger a_q^\dagger a_r &= \delta_{ip} a_q^\dagger a_r - a_p^\dagger a_i a_q^\dagger a_r \\ &= \delta_{ip} a_q^\dagger a_r - a_p^\dagger (\delta_{iq} - a_q^\dagger a_i) a_r \\ &= \delta_{ip} a_q^\dagger a_r - \delta_{iq} a_p^\dagger a_r - a_p^\dagger a_q^\dagger a_r a_i, \end{aligned}$$

so

$$a_i a_p^\dagger a_q^\dagger a_r + a_p^\dagger a_q^\dagger a_r a_i = \delta_{ip} a_q^\dagger a_r - \delta_{iq} a_p^\dagger a_r,$$

and

$$\begin{aligned} C_{ij} &= h_{ij} + \frac{1}{2} \sum_{pqr} \{pq|V|jr\} \delta_{ip} \langle \Psi_0^N | a_q^\dagger a_r | \Psi_0^N \rangle - \frac{1}{2} \sum_{pqr} \{pq|V|jr\} \delta_{iq} \langle \Psi_0^N | a_p^\dagger a_r | \Psi_0^N \rangle \\ &= h_{ij} + \frac{1}{2} \sum_{qr} \{iq|V|jr\} \langle \Psi_0^N | a_q^\dagger a_r | \Psi_0^N \rangle - \frac{1}{2} \sum_{pr} \{pi|V|jr\} \langle \Psi_0^N | a_p^\dagger a_r | \Psi_0^N \rangle \\ &= h_{ij} + \frac{1}{2} \sum_{qr} [\{iq|V|jr\} - \{qi|V|jr\}] \langle \Psi_0^N | a_q^\dagger a_r | \Psi_0^N \rangle \\ &= h_{ij} + \sum_{qr} \{iq|V|jr\} \langle \Psi_0^N | a_q^\dagger a_r | \Psi_0^N \rangle. \end{aligned}$$

Define the one-particle reduced density matrix (1-RDM) by

$$D_{qr} = \langle \Psi_0^N | a_q^\dagger a_r | \Psi_0^N \rangle.$$

Then

$$C_{ij} = h_{ij} + \sum_{qr} \{iq|V|jr\} D_{qr}.$$

□

## 5 Application to DMFT

### 5.1 Well-posedness of DMFT loop

We assume the notation of section (3.3) and let  $\Delta^{(p)} \in \mathcal{S}_+^{L^{(p)}}$  for  $p = 1, \dots, P$ . Then let  $\Sigma^{(p)} = \Sigma^{(p)}[\Delta^{(p)}]$  be the fragment block of the self-energy of the composite fragment/bath system specified by the single-particle matrix  $h^{(p)} + \Delta^{(p)}$ . By Lemma (17), we have in particular that  $-\Sigma^{(p)} \in \mathcal{N}_+^{L^{(p)}}$ .

Recall that

$$G^{\text{glob}} := (z - h - \Sigma_{\text{DMFT}}(z))^{-1},$$

where  $\Sigma_{\text{DMFT}}(z) = \bigoplus_{p=1}^P \Sigma^{(p)}(z)$ . Recall further that  $G^{(p)} := [G^{\text{glob}}]_{\mathcal{C}^{(p)}}$ , so by the Schur complement theorem we have

$$\begin{aligned} [G^{(p)}]^{-1} &= [(G^{\text{glob}})^{-1}]_{\mathcal{C}^{(p)}} \\ &\quad - [(G^{\text{glob}})^{-1}]_{\mathcal{C}^{(p)}, \mathcal{C} \setminus \mathcal{C}^{(p)}} \left( [(G^{\text{glob}})^{-1}]_{\mathcal{C} \setminus \mathcal{C}^{(p)}} \right)^{-1} [(G^{\text{glob}})^{-1}]_{\mathcal{C} \setminus \mathcal{C}^{(p)}, \mathcal{C}^{(p)}}. \end{aligned}$$

Now on the off-block-diagonal  $(G^{\text{glob}})^{-1}$  simply coincides with  $-h$ , so we have

$$G^{(p)} = \left( z - h^{(p)} - \Sigma^{(p)} - h_{\mathcal{C}^{(p)}, \mathcal{C} \setminus \mathcal{C}^{(p)}} \left( [z - h - \Sigma_{\text{DMFT}}(z)]_{\mathcal{C} \setminus \mathcal{C}^{(p)}} \right)^{-1} h_{\mathcal{C} \setminus \mathcal{C}^{(p)}, \mathcal{C}^{(p)}} \right)^{-1}.$$

But recall from (3.8) that

$$\Delta_{\text{new}}^{(p)}(z) := z - h^{(p)} - \Sigma^{(p)} - G^{(p)}(z)^{-1},$$

so we have shown that

**Lemma 19.** *The DMFT iteration (3.8) can alternatively be written*

$$\Delta_{\text{new}}^{(p)}(z) = h_{\mathcal{C}^{(p)}, \mathcal{C} \setminus \mathcal{C}^{(p)}} \left( [z - h - \Sigma_{\text{DMFT}}(z)]_{\mathcal{C} \setminus \mathcal{C}^{(p)}} \right)^{-1} h_{\mathcal{C} \setminus \mathcal{C}^{(p)}, \mathcal{C}^{(p)}}.$$

**Corollary 20** (Well-posedness of DMFT loop). *The  $\Delta_{\text{new}}^{(p)}$  as defined in (3.8) lie in  $\mathcal{S}_+^{L^{(p)}}$ .*

*Proof.* Now since  $-\Sigma^{(p)} \in \mathcal{N}_+^{L^{(p)}}$ , it follows that the map  $z \mapsto z - h - \Sigma_{\text{DMFT}}(z)$  is in  $\mathcal{N}_1^{L^{(p)}}$ . Hence the pointwise inverse

$$z \mapsto \left( [z - h - \Sigma_{\text{DMFT}}(z)]_{\mathcal{C} \setminus \mathcal{C}^{(p)}} \right)^{-1}$$

is in  $\mathcal{S}_1^{L^{(p)}}$ . The congruence transformation by  $h_{\mathcal{C}^{(p)}, \mathcal{C} \setminus \mathcal{C}^{(p)}} = h_{\mathcal{C} \setminus \mathcal{C}^{(p)}, \mathcal{C}^{(p)}}^*$ , which yields  $\Delta_{\text{new}}^{(p)}$  by Lemma 19, lies within  $\mathcal{S}_+^{L^{(p)}}$ , as desired.  $\square$

## 5.2 Sparsity pattern of the hybridization for local models

The formula in the statement of Lemma 19 allows us to easily recover a constraint on the hybridization, namely that the only nonzero block of the hybridization is the block corresponding to the boundary of the cluster, in a certain sense. For hybridization fitting in the sense of 3.2.1, this constraint allows us to fit hybridizations in a smaller space.

Now we can think of the hopping matrix  $h$  as defining a graph on the vertex set  $\mathcal{C} = \{1, \dots, M\}$ , in which two vertices  $i, j$  are connected if and only if  $h_{ij} \neq 0$ . Then the boundary  $\partial\mathcal{I}$  of a subset  $\mathcal{I} \subset \mathcal{C}$  is the subset of vertices in  $\mathcal{I}$  that share an edge with some vertex in  $\mathcal{C} \setminus \mathcal{I}$ . We define  $\mathcal{I}^\circ := \mathcal{I} \setminus \partial\mathcal{I}$ . Then we can state:

**Corollary 21.** *Consider  $\Delta_{\text{new}}^{(p)}$  as a  $2 \times 2$  block matrix via the partition of the indices  $\mathcal{C}^{(p)}$  into  $[\mathcal{C}^{(p)}]^\circ$  and  $\partial\mathcal{C}^{(p)}$ . Then the only nonzero block is the diagonal block  $[\Delta_{\text{new}}^{(p)}]_{\partial\mathcal{C}^{(p)}}$ , which is given by*

$$[\Delta_{\text{new}}^{(p)}(z)]_{\partial\mathcal{C}^{(p)}} = h_{\partial\mathcal{C}^{(p)}, \mathcal{C} \setminus \mathcal{C}^{(p)}} \left( [z - h - \Sigma_{\text{DMFT}}(z)]_{\mathcal{C} \setminus \mathcal{C}^{(p)}} \right)^{-1} h_{\mathcal{C} \setminus \mathcal{C}^{(p)}, \partial\mathcal{C}^{(p)}}.$$

*Proof.* The proof follows from Lemma 19, together with the observation that, due to the construction of the graph induced by  $h$ ,  $h_{[\mathcal{C}^{(p)}]^\circ, \mathcal{C} \setminus \mathcal{C}^{(p)}} = 0$ .  $\square$

## 5.3 Definition and well-posedness of HF+DMFT loop

Suppose that our Hamiltonian is of the more general form

$$\hat{H} = a^\dagger h a + \sum_{p=1}^P \hat{V}^{(p)} + \hat{V}^{\text{LR}},$$

where  $\hat{V}_{\text{LR}}$  indicates the long-range part of the interaction. For simplicity we let all interaction terms be of two-body form, in which case one can assume without loss of generality that

$$\hat{V}^{\text{LR}} = \frac{1}{4} \sum_{ijkl} \{ij|V^{\text{LR}}|kl\} a_i^\dagger a_j^\dagger a_l a_k,$$

where  $\{ij|V^{\text{LR}}|kl\} = 0$  if  $i, j, k, l \in \mathcal{C}^{(p)}$  for some  $p$  and moreover

$$\{ij|V^{\text{LR}}|kl\} = -\{ji|V^{\text{LR}}|kl\} = -\{ij|V^{\text{LR}}|lk\} = \{ji|V^{\text{LR}}|lk\}$$

for all  $i, j, k, l$ .

Then, motivated by section 4.4, the Hartree-Fock plus DMFT (HF+DMFT) method is specified by replacing  $\Sigma_{\text{DMFT}}$  in the DMFT loop with the quantity

$$\Sigma_{\text{HF+DMFT}} := \Sigma_{\text{HF-LR}} + \Sigma_{\text{DMFT}},$$



where  $\Sigma_{\text{DMFT}}$  is defined as before and

$$[\Sigma_{\text{HF-LR}}]_{ij} := \sum_{qr} \{iq|V^{\text{LR}}|jr\} D_{qr}^{\text{glob}}$$

for  $D^{\text{glob}} = D^{\text{glob}}[G^{\text{glob}}]$  is defined in terms of  $G^{\text{glob}}(z) = \int_{\mathbb{R}} \frac{d\mathbf{M}(\varepsilon)}{z-\varepsilon}$  via  $D^{\text{glob}} := \int_{-\infty}^{\mu} d\mathbf{M}(\varepsilon)$ , where  $\mu$  is the chemical potential. Here  $G^{\text{glob}}$  is the global Green's function retained from the previous iteration.

Then it is immediate from repeated arguments that

$$\Delta_{\text{new}}^{(p)}(z) = \tilde{h}_{\mathcal{C}^{(p)}, \mathcal{C} \setminus \mathcal{C}^{(p)}} \left( [z - h - \Sigma_{\text{DMFT}}(z)]_{\mathcal{C} \setminus \mathcal{C}^{(p)}} \right)^{-1} \tilde{h}_{\mathcal{C} \setminus \mathcal{C}^{(p)}, \mathcal{C}^{(p)}},$$

where  $\tilde{h} := h + \Sigma_{\text{HF-LR}}$ . Hence in particular the hybridization iterates stay within  $\mathcal{S}_+$ .

## Part VIII

# Strictly correlated electrons in second quantization

## 1 Introduction

In this Part, we consider a convex relaxation of the fermionic Gibbs variational principle (also considered in Part V). In fact we consider a succession of two relaxations. The first of these relaxations (note: we shall come to the second later on) is motivated by its asymptotic tightness in the limit of infinitely strong Coulomb repulsion, called the limit of *strictly correlated electrons (SCE)*. The formalism of SCE that we consider can be seen as a second-quantized analog of the first-quantized formalism heretofore considered in the literature [95, 94, 20, 66, 26], which has been motivated largely by its relevance to the widely-used Kohn-Sham density functional theory (DFT) [44, 50].

For ground-state electronic structure calculations, the success of DFT hinges on the accuracy of the approximate exchange-correlation functionals. Although tremendous progress has been made in the construction of approximate functionals [82, 8, 55, 81], these approximations are mostly derived by fitting known results for the uniform electron gas, single atoms, small molecules, and perfect crystal systems. Such functionals often perform well when the underlying quantum systems are ‘weakly correlated,’ i.e., when the single-particle energy is significantly more important than the electron-electron interaction energy. In order to extend the capability of DFT to the treatment of strongly correlated quantum systems, one recent direction of functional development considers the limit in which the electron-electron interaction energy is infinitely large compared to other components of the total energy, i.e., the SCE limit, which provides an alternative route to derive exchange-correlation energy functionals. The study of Kohn-Sham DFT with SCE-based functionals is still in its infancy, but such approaches have already been used to treat strongly correlated model systems and simple chemical systems (see e.g. [75, 22, 40]).

Here we recapitulate some relevant background from Part I. A system of  $N$  interacting electrons in a  $d$ -dimensional space can be described using either the first-quantized or the second-quantized representation. In the first-quantized representation, the number of electrons  $N$  is fixed, and the electronic wavefunction is an anti-symmetric function in  $\bigwedge^N L^2(\mathbb{R}^d; \mathbb{C}^2)$ , which is a subset of the tensor product space  $\bigotimes^N L^2(\mathbb{R}^d; \mathbb{C}^2)$ . Here  $\mathbb{C}^2$  corresponds to the spin degree of freedom. In first quantization, the anti-symmetry condition needs to be treated explicitly. By contrast, in the second-quantized formalism, one chooses a basis for a subspace of  $L^2(\mathbb{R}^d; \mathbb{C}^2)$ . In practice, the basis is of some finite size  $L$ , corresponding to a discretized model with  $L$  sites that encode both spatial and spin degrees of freedom. The electronic

wavefunction is an element of the Fock space  $\mathcal{F} \cong \mathbb{C}^{2^L}$ . The Fock space contains wavefunctions of all possible electron numbers, and finding wavefunctions of the desired electron number is achieved by constraining to a subspace of the Fock space. In the second-quantized representation, the anti-symmetry constraint is in some sense baked into the Hamiltonian operator instead of the wavefunction, and this perspective often simplifies book-keeping efforts. Due to the inherent computational difficulty of studying strongly correlated systems such as high-temperature superconductors, it is often necessary to introduce simplified Hamiltonians such as in Hubbard-type models. These model problems are formulated directly in the second-quantized formalism via specification of an appropriate Hamiltonian.

To the extent of our knowledge, all existing works on SCE treat electrons in the first-quantized representation with (essentially) a real space basis. In this Part we aim at studying the SCE limit in the second-quantized setting. Note that generally Kohn-Sham-type theories in the second-quantized representation are known as ‘site occupation functional theory’ (SOFT) or ‘lattice density functional theory’ in the physics literature [93, 60, 21, 97, 23]. A crucial assumption of this Part is that the electron-electron interaction takes the form  $\sum_{p,q=1}^L v_{pq} \hat{n}_p \hat{n}_q$ , which we call the generalized Coulomb interaction. (The meaning of the symbols will be explained in Section 2.) We remark that the form of the generalized Coulomb interaction is more restrictive than the general form  $\sum_{p,q,r,s=1}^L v_{pqrs} \hat{a}_p^\dagger \hat{a}_q^\dagger \hat{a}_s \hat{a}_r$  appearing in the quantum chemistry literature, to which our formulation does not yet apply. Assuming a generalized Coulomb interaction, we demonstrate that the corresponding SCE problem can be formulated as a multi-marginal optimal transport (MMOT) problem over classical probability measures on the binary hypercube  $\{0, 1\}^L$ . The cost function in this problem is of pairwise form. Hence the objective function in the Kantorovich formulation of the MMOT can be written in terms of only the 2-marginals of the probability measure. In order to solve the MMOT problem directly, even the storage cost of the exact solution scales as  $2^L$ , and the computational cost also scales exponentially with respect to  $L$ . Thus a direct approach becomes impractical even when the number of sites becomes moderately large.

## 1.1 Contributions

Please note that this Part is based on [47] (joint work with Yuehaw Khoo, Lin Lin, and Lexing Ying). Based on the recent work of Khoo and Ying [48], in this Part we propose a convex relaxation approach for MMOT that imposes only certain necessary constraints satisfied by the 2-marginals. This can be considered as the *second* relaxation in our succession of relaxations of the fermionic variational principle. The relaxed problem can be solved efficiently via semidefinite programming (SDP). While the 2-marginal formulation provides a lower bound to the optimal cost of the MMOT problem, we also propose a tighter lower bound obtained via an SDP involving the 3-marginals. The computational cost for solving these relaxed problems is polynomial

with respect to  $L$ , and, in particular, the semidefinite constraint is only enforced on a matrix of size  $2L \times 2L$ . Numerical results for spinless and spinful Hubbard-type systems demonstrate that the 2-marginal and 3-marginal relaxation schemes are already quite tight, especially when compared to the modeling error due to the Kohn-Sham SCE formulation itself.

By solving the dual problems for our SDPs, we can obtain the Kantorovich dual potentials, which yield the SCE potential needed for carrying out the self-consistent field iteration (SCF) in the Kohn-Sham SCE formalism. To this end we need to show that the dual problem satisfies strong duality and moreover that the dual optimizer is actually attained. We show that a straightforward formulation of the primal SDP does not have any strictly feasible point, and hence Slater’s condition cannot be directly applied to establish strong duality (see, e.g., [17]). By a careful study of the structure of the dual problem, we prove that the strong duality and dual attainment conditions are indeed satisfied. We also explain how the SDP relaxations introduced in this Part can be applied to arbitrary MMOT problems with pairwise cost functions. We comment that the justification of the strong duality and dual attainment conditions holds in this more general setting as well.

## 1.2 Related work

In the first-quantized formulation, for a fixed real-space discretization the computational cost of the direct solution of the SCE problem scales exponentially with respect to the number of electrons  $N$ . This curse of dimensionality is a serious obstacle for SCE-based approaches to the quantum many-body problem. Notable exceptions to the unfavorable computational scaling are the cases of strictly one-dimensional systems (i.e.,  $d = 1$ ) and spherically symmetric systems (for any  $d$ ) [94], for which semi-analytic solutions exist.

In [9], the Sinkhorn scaling approach is applied to an entropically regularized MMOT problem. This method requires the marginalization of a probability measure on a product space of size that is exponential in the number of electrons  $N$ . Thus the complexity of this method also scales exponentially with respect to  $N$ . Meanwhile, a method based on the Kantorovich dual of the MMOT problem was proposed in [20, 74]. However, there are exponentially many constraints in the dual problem. Furthermore, [20] assumes a Monge solution to the MMOT problem, but it is unknown whether the MMOT problem with pairwise Coulomb cost has a Monge solution for  $d = 2, 3$ . Moreover, if it exists, the Monge solution is hard to evaluate in the context of the Coulomb cost.

Recently, Khoo and Ying proposed a semidefinite relaxation-based approach to the MMOT problem arising from SCE in the first-quantized setting [48]. This is the first approximation method for the general SCE problem with polynomial complexity with respect to the system size. The relaxation avoids exponential scaling by directly handling only the 2-marginal distributions (known as the pair densities in the physics

literature), which are subjected to certain necessary joint representability constraints. In particular, the method provides a lower bound to the SCE energy. Furthermore, by proper treatment of the 3-marginal distributions, an upper bound to the SCE energy is recovered as well. Numerical results indicate that both the lower and upper bounds are rather tight approximations to the SCE energy.

In the second-quantized setting, our semidefinite relaxation-based approach for finding a lower bound to the SCE energy is also related to the two-particle reduced density matrix (2-RDM) theories in quantum chemistry [24, 71, 69, 70]. However, the MMOT problem in SCE only requires the knowledge of the pair density instead of the entire 2-RDM. The number of constraints in our formulation is also considerably smaller than the number of constraints in 2-RDM theories, thanks to the generalized Coulomb form of the interaction.

### 1.3 Outline

In Section 2, we describe the Hamiltonians under consideration and derive an appropriate formulation of Kohn-Sham DFT based on the SCE functional, which is in turn defined in terms of a MMOT problem. In Section 3, we solve the MMOT problem by introducing a convex relaxation of the set of representable 2-marginals, and we prove strong duality for the relaxed problem. In Section 4, a tighter lower bound is obtained by considering a convex relaxation of the set of representable 3-marginals. In Section 5, we comment on how a general MMOT problem with pairwise cost can be solved by directly applying the methods introduced in Sections 3 and 4. We demonstrate the effectiveness of the proposed methods through numerical experiments in Section 6, and we discuss conclusions and future directions in Section 7.

## 2 Preliminaries

### 2.1 Density functional theory in second quantization

Our goal is to compute the ground-state energy of a fermionic system with  $L$  states. With some abuse of terminology, we will refer to fermions simply as electrons. Also for simplicity we use a single index for all of the states, as opposed to using separate site and spin indices in the case of spinful systems. Double indexing for spinful fermionic systems can be recovered simply by rearranging indices, e.g., by associating odd state indices with spin-up components and even state indices with spin-down components.

In second quantization, the state space is called the Fock space, denoted by  $\mathcal{F}$ . We shall now recapitulate some of the relevant background on second quantization from section 4 of Part I and meanwhile set some notation for this Part.

The occupation number basis set for the Fock space is

$$\{|s_1, \dots, s_L\rangle\}_{s_i \in \{0,1\}, i=1, \dots, L},$$

which is an orthonormal basis set satisfying

$$\langle s_{i_1}, \dots, s_{i_L} | s_{j_1}, \dots, s_{j_L} \rangle = \delta_{i_1 j_1} \cdots \delta_{i_L j_L}. \quad (2.1)$$

A state  $|\psi\rangle \in \mathcal{F}$  will be written as a linear combination of occupation number basis elements as follows:

$$|\psi\rangle = \sum_{s_1, \dots, s_L \in \{0,1\}} \psi(s_1, \dots, s_L) |s_1, \dots, s_L\rangle, \quad \psi(s_1, \dots, s_L) \in \mathbb{C}. \quad (2.2)$$

Hence the state vector  $|\psi\rangle$  can be identified with a vector  $\psi \in \mathbb{C}^{2^L}$ , and  $\mathcal{F}$  is isomorphic to  $\mathbb{C}^{2^L}$ . We call  $|\psi\rangle$  normalized if the following condition is satisfied:

$$\langle \psi | \psi \rangle = \sum_{s_1, \dots, s_L \in \{0,1\}} |\psi(s_1, \dots, s_L)|^2 = 1. \quad (2.3)$$

We also refer to  $|0\rangle = |0, \dots, 0\rangle$  as the vacuum state.

The fermionic creation and annihilation operators are respectively defined as

$$\begin{aligned} \hat{a}_p^\dagger |s_1, \dots, s_L\rangle &= (-1)^{\sum_{q=1}^{p-1} s_q} (1 - s_p) |s_1, \dots, 1 - s_p, \dots, s_L\rangle, \\ \hat{a}_p |s_1, \dots, s_L\rangle &= (-1)^{\sum_{q=1}^{p-1} s_q} s_p |s_1, \dots, 1 - s_p, \dots, s_L\rangle, \quad p = 1, \dots, L. \end{aligned} \quad (2.4)$$

The number operator defined as  $\hat{n}_p := \hat{a}_p^\dagger \hat{a}_p$  satisfies

$$\hat{n}_p |s_1, \dots, s_L\rangle = s_p |s_1, \dots, s_L\rangle, \quad p = 1, \dots, L. \quad (2.5)$$

The Hamiltonian operator is assumed to take the following form:

$$\hat{H} = \sum_{p,q=1}^L t_{pq} \hat{a}_p^\dagger \hat{a}_q + \sum_{p=1}^L w_p \hat{n}_p + \sum_{p,q=1}^L v_{pq} \hat{n}_p \hat{n}_q. \quad (2.6)$$

Here  $t \in \mathbb{C}^{L \times L}$  is a Hermitian matrix, which is often interpreted as the ‘hopping’ term arising from the kinetic energy contribution to the Hamiltonian.  $w$  is an on-site term, which can be viewed as an external potential.  $v \in \mathbb{C}^{L \times L}$  is also a Hermitian matrix, which may be viewed as representing the electron-electron Coulomb interaction. Note that  $\hat{n}_p = \hat{a}_p^\dagger \hat{a}_p = \hat{n}_p \hat{n}_p$ , hence without loss of generality we can assume the diagonal entries  $t_{pp} = v_{pp} = 0$  by absorbing, if necessary, such contributions into the on-site potential  $w$ . Following the spirit of Kohn-Sham DFT, one could think of  $t, v$  as fixed matrices, and of the external potential  $w$  as a contribution that may change depending on the system (in the context of DFT,  $w$  represents the electron-nuclei interaction and is therefore ‘external’ to the electrons). We remark that the restriction of the form of the two-body interaction  $\sum_{p,q=1}^L v_{pq} \hat{n}_p \hat{n}_q$  is crucial for the purpose of this Part. In particular, we do not consider the more general form  $\sum_{p,q,r,s=1}^L v_{pqrs} \hat{a}_p^\dagger \hat{a}_q^\dagger \hat{a}_s \hat{a}_r$  as is done in the quantum chemistry literature when a general basis set (such as the

Gaussian basis set) is used to discretize a quantum many-body Hamiltonian in the continuous space. In the discussion below, for simplicity we will omit the index range of our sums as long as the meaning is clear.

The exact ground state energy  $E_0$  can be obtained by the following minimization problem:

$$E_0 = \inf_{|\psi\rangle \in \mathcal{F}: \langle \psi | \psi \rangle = 1} \langle \psi | \hat{H} - \mu \hat{N} | \psi \rangle. \quad (2.7)$$

Here the minimizer  $|\psi\rangle$  is the many-body ground state wavefunction, and  $\hat{N} := \sum_p \hat{n}_p$  is the total number operator.  $\mu$ , which is called the chemical potential, is a Lagrange multiplier chosen so that the ground state wavefunction  $|\psi\rangle$  has a number of electrons equal to a pre-specified integer  $N \in \{0, 1, \dots, L\}$ , i.e., such that

$$\langle \psi | \hat{N} | \psi \rangle = N. \quad (2.8)$$

It is clear that  $\mu \hat{N}$  is an on-site potential, and without loss of generality we absorb  $\mu$  into  $w$ , and hence write  $\hat{H} - \mu \hat{N}$  as  $\hat{H}$  in the discussion below.

The electron density  $\rho \in \mathbb{R}^L$  is defined as

$$\rho_p = \langle \psi | \hat{n}_p | \psi \rangle = \sum_{s_1, \dots, s_L} |\psi(s_1, \dots, s_L)|^2 s_p, \quad p = 1, \dots, L, \quad (2.9)$$

which satisfies  $\sum_p \rho_p = N$ . Note that

$$\langle \psi | \sum_p w_p \hat{n}_p | \psi \rangle = \sum_p w_p \rho_p =: W[\rho]. \quad (2.10)$$

Then we follow the Levy-Lieb constrained minimization approach [56, 59] and rewrite the ground state minimization problem (2.7) as follows:

$$\begin{aligned} E_0 &= \inf_{\rho \in \mathcal{J}_N} \left\{ \sum_p \rho_p w_p + \left( \inf_{|\psi\rangle \mapsto \rho, |\psi\rangle \in \mathcal{F}} \langle \psi | \sum_{pq} t_{pq} \hat{a}_p^\dagger \hat{a}_q + \sum_{pq} v_{pq} \hat{n}_p \hat{n}_q | \psi \rangle \right) \right\} \\ &= \inf_{\rho \in \mathcal{J}_N} \{W[\rho] + F_{\text{LL}}[\rho]\}, \end{aligned} \quad (2.11)$$

where

$$F_{\text{LL}}[\rho] := \inf_{|\psi\rangle \mapsto \rho, |\psi\rangle \in \mathcal{F}} \langle \psi | \sum_{pq} t_{pq} \hat{a}_p^\dagger \hat{a}_q + \sum_{pq} v_{pq} \hat{n}_p \hat{n}_q | \psi \rangle. \quad (2.12)$$

Here the notation  $\psi \mapsto \rho$  indicates that the corresponding infimum is taken over states  $|\psi\rangle$  that yield the density  $\rho$  in the sense of Eq. (2.9), and the domain  $\mathcal{J}_N$  of  $\rho$  is defined by

$$\mathcal{J}_N := \left\{ \rho \in \mathbb{R}^L \mid \rho \geq 0, \sum_p \rho_p = N \right\}. \quad (2.13)$$

Note that the external potential  $w$  is only coupled with  $\rho$  and is singled out in the constrained minimization. It is easy to see that for any  $\rho \in \mathcal{J}_N$ , the set  $\{|\psi\rangle \in \mathcal{F} : |\psi\rangle \mapsto \rho\}$  is non-empty, as we may simply choose

$$|\psi\rangle = \sum_p \sqrt{\rho_p} \left| s_1^{(p)}, \dots, s_L^{(p)} \right\rangle, \quad s_q^{(p)} = \delta_{pq}.$$

Therefore the constrained minimization problem (2.11) is in fact defined over a nonempty set for all  $\rho \in \mathcal{J}_N$ .

The functional  $F_{\text{LL}}[\rho]$ , which is called the Levy-Lieb functional, is a universal functional in the sense that it depends only on the hopping term  $t$  and the interaction term  $v$ , hence in particular is independent of the potential  $w$ . Once the functional  $F_{\text{LL}}[\rho]$  is known,  $E_0$  can be obtained by minimization with respect to a single vector  $\rho$  using standard optimization algorithms, or via the self-consistent field (SCF) iteration to be detailed below. The construction above is called the ‘site occupation functional theory’ (SOFT) or ‘lattice density functional theory’ in the physics literature [93, 60, 21, 97, 23]. To our knowledge, SOFT or lattice DFT often imposes an additional sparsity pattern on the  $v$  matrix for the electron-electron interaction, so that the Hamiltonian becomes a Hubbard-type model.

## 2.2 Strictly correlated electron limit

Using the fact that the infimum of a sum is greater than the sum of infimums, we can lower-bound the ground state energy in the following way:

$$F_{\text{LL}}[\rho] \geq \inf_{|\psi\rangle \mapsto \rho} \langle \psi | \sum_{pq} t_{pq} \hat{a}_p^\dagger \hat{a}_q | \psi \rangle + \inf_{|\psi\rangle \mapsto \rho} \langle \psi | \sum_{pq} v_{pq} \hat{n}_p \hat{n}_q | \psi \rangle =: T[\rho] + E_{\text{sce}}[\rho], \quad (2.14)$$

where the functionals  $T[\rho]$  and  $E_{\text{sce}}[\rho]$  are defined via the last equality in the manner suggested by the notation. The first of these quantities is called the kinetic energy, and the second the strictly correlated electron (SCE) energy. The SCE approximation is obtained by treating  $T[\rho] + E_{\text{sce}}[\rho]$  as an approximation for the Levy-Lieb functional. Though in general it is only a lower-bound for the Levy-Lieb functional, this bound is expected to become tight in the limit of infinitely strong interaction. We do not prove this fact in this Part (though we demonstrate it numerically below), but we nonetheless refer to this approximation as the SCE limit by analogy to the literature on SCE in first quantization [95, 94].

Due to the inequality in Eq. (2.14), we have in general the following lower bound for the total energy, which we shall call the Kohn-Sham SCE energy:

$$E_0 \geq E_{\text{KS-SCE}} := \inf_{\rho \in \mathcal{J}_N} \{W[\rho] + T[\rho] + E_{\text{sce}}[\rho]\}. \quad (2.15)$$

The advantage of the preceding manipulations is that now each term in this infimum can be computed. Specifically,  $W[\rho]$  is trivial to compute,  $T[\rho]$  is defined in terms of a



non-interacting many-body problem (i.e., a problem with Hamiltonian only quadratic in the creation and annihilation operators), for which an exact solution can be obtained via the diagonalization of  $t$  [77]. Finally, as we shall see below the SCE term (and its gradient) can be computed in terms of a MMOT problem (and its dual). Thus in principle, it would be possible to take gradient descent approach for computing the infimum in the definition (2.15) of  $E_{\text{KS-SCE}}$ .

### 2.2.1 The Kohn-Sham SCE equations

In practice, to compute the Kohn-Sham SCE energy we will instead adopt the self-consistent field (SCF) iteration as is common practice in Kohn-Sham DFT. It can be readily checked that  $E_{\text{sce}}[\rho]$  is convex with respect to  $\rho$ . By the convexity of  $W[\rho]$ ,  $T[\rho]$ , and  $E_{\text{sce}}[\rho]$ , the expression in Eq. (2.15) admits a minimizer, which is unique unless the functional fails to be strictly convex. We assume that the solution is unique and  $E_{\text{sce}}[\rho]$  is differentiable for simplicity, and we derive nonlinear fixed-point equations satisfied by the minimizer as follows.

For suitable  $\rho$ , define the SCE potential via

$$v_{\text{sce}}[\rho] = \nabla_{\rho} E_{\text{sce}}[\rho], \quad (2.16)$$

and we will discuss how to compute this gradient later. Now assume that the (unique) infimum in Eq. (2.15) is obtained at  $\rho^*$ , which is then in particular a critical point of the expression

$$W[\rho] + T[\rho] + E_{\text{sce}}[\rho]. \quad (2.17)$$

But then  $\rho^*$  is also a critical point of the expression obtained by replacing  $E_{\text{sce}}[\rho]$  with its expansion up to first order about  $\rho^*$ , which is (modulo a constant term that does not affect criticality)

$$G[\rho] := W[\rho] + T[\rho] + v_{\text{sce}}[\rho^*] \cdot \rho = T[\rho] + (w + v_{\text{sce}}[\rho^*]) \cdot \rho. \quad (2.18)$$

Hence  $\cdot$  means the inner product, and we are motivated to try to minimize  $G[\rho]$  over  $\rho \in \mathcal{J}_N$ . But we can write

$$G[\rho] = \inf_{|\psi\rangle \mapsto \rho} \langle \psi | \sum_{pq} h_{pq}[\rho^*] \hat{a}_p^\dagger \hat{a}_q | \psi \rangle,$$

where

$$h[\rho] := t + \text{diag}(w + v_{\text{sce}}[\rho]).$$

Here  $\text{diag}(\cdot)$  is a diagonal matrix. Then

$$\inf_{\rho \in \mathcal{J}_N} G[\rho] = \inf_{|\psi\rangle \in \mathcal{F} : \langle \psi | \psi \rangle = 1, \langle \psi | \hat{N} | \psi \rangle = N} \langle \psi | \sum_{pq} h_{pq}[\rho^*] \hat{a}_p^\dagger \hat{a}_q | \psi \rangle.$$

The latter infimum is a ground-state problem for a non-interacting Hamiltonian and is obtained [77] at a so-called Slater determinant of the form

$$|\psi\rangle = \hat{c}_1^\dagger \cdots \hat{c}_N^\dagger |0\rangle. \quad (2.19)$$

Here the  $\hat{c}_k^\dagger$  are ‘canonically transformed’ creation operators defined by

$$\hat{c}_k^\dagger = \sum_p \hat{a}_p^\dagger \varphi_{pk}, \quad (2.20)$$

where  $\Phi = [\varphi_1 \cdots \varphi_N] = [\varphi_{pk}] \in \mathbb{C}^{L \times N}$  is a matrix whose columns are the  $N$  lowest eigenvectors of  $h[\rho^*]$ . We assume the eigenvectors form an orthonormal set, i.e.  $\Phi^* \Phi = I_N$ .

Moreover, one may directly compute that the electron density of  $|\psi\rangle$  as defined in Eq. (2.19) is given by

$$\rho_p = \langle \psi | \hat{n}_p | \psi \rangle = \sum_{k=1}^N |\varphi_{pk}|^2, \quad (2.21)$$

i.e.,  $\rho = \text{diag}(\Phi \Phi^*)$ . Hence the optimizer  $\rho^*$  of Eq. (2.15) solves the Kohn-Sham SCE equations:

$$\begin{aligned} (t + \text{diag}(w + v_{\text{sce}}[\rho])) \varphi_k &= \varepsilon_k \varphi_k, \quad k = 1, \dots, N. \\ \rho &= \text{diag}(\Phi \Phi^*). \end{aligned} \quad (2.22)$$

Here  $(\varepsilon_k, \varphi_k)$  are understood to be the  $N$  lowest (orthonormal) eigenpairs of the matrix in the first line of Eq. (2.22).

Eq. (2.22) is a nonlinear eigenvalue problem and should be solved self-consistently. The standard iterative procedure for this task works as follows. (1) For the  $k$ -th iterate  $\rho^{(k)}$ , form the matrix  $h[\rho^{(k)}]$ , and compute  $\Phi^{(k)}$  by solving the corresponding eigenproblem. (2) Define  $\rho^{(k+1)} := \text{diag}(\Phi^{(k)} \Phi^{(k)*})$ . (3) Iterate until convergence, possibly using mixing schemes [3, 87, 64] to ensure or accelerate convergence.

Once self-consistency is reached, the total energy can be recovered by the relation

$$E_{\text{KS-SCE}} = \sum_{k=1}^N \varepsilon_k - v_{\text{sce}}[\rho^*] \cdot \rho^* + E_{\text{sce}}[\rho^*], \quad (2.23)$$

as can be observed by adding back to  $G[\rho^*]$  the constant term discarded between equations (2.17) and (2.18).

### 2.2.2 The SCE energy and potential

The problem is then reduced to the computation of  $E_{\text{sce}}[\rho]$  and its gradient  $v_{\text{sce}}[\rho]$ . To this end, let us rewrite

$$\begin{aligned} E_{\text{sce}}[\rho] &= \inf_{|\psi\rangle \mapsto \rho} \langle \psi | \sum_{pq} v_{pq} \hat{n}_p \hat{n}_q | \psi \rangle \\ &= \inf_{|\psi\rangle \mapsto \rho} \sum_{s_1, \dots, s_L} \sum_{pq} v_{pq} s_p s_q |\psi(s_1, \dots, s_L)|^2 \\ &= \inf_{\mu \in \Pi(\rho)} \sum_{s_1, \dots, s_L} \sum_{pq} v_{pq} s_p s_q \mu(s_1, \dots, s_L), \end{aligned} \quad (2.24)$$

where  $\Pi(\rho)$  is the space of joint probability mass functions on  $\{0, 1\}^L$ . The 1-marginals  $\mu_p^{(1)}$  are defined in terms of  $\mu$  via

$$\mu_p^{(1)}(s_p) := \sum_{s_1, \dots, s_L \setminus \{s_p\}} \mu(s_1, \dots, s_L), \quad (2.25)$$

and they satisfy

$$\mu_p^{(1)}(s) = (1 - \rho_p) \delta_{s0} + \rho_p \delta_{s1}, \quad s = 0, 1. \quad (2.26)$$

Considering the  $\mu_p^{(1)}$  alternately as vectors, we also write (by some abuse of notation)

$$\mu_p^{(1)} = [1 - \rho_p, \rho_p]^\top. \quad (2.27)$$

Note that the last line of Eq. (2.24) is obtained by considering  $|\psi(s_1, \dots, s_L)|^2$  as a classical probability density  $\mu(s_1, \dots, s_L) \in \Pi(\rho)$ . (The marginal condition derives from the condition  $|\Psi\rangle \mapsto \rho$ .)

Define the cost function  $C : \{0, 1\}^L \rightarrow \mathbb{R}$  by

$$C(s_1, \dots, s_L) := \sum_{pq} v_{pq} s_p s_q. \quad (2.28)$$

Then our SCE energy may be written

$$E_{\text{sce}}[\rho] = \inf_{\mu \in \Pi(\rho)} \sum_{s_1, \dots, s_L} C(s_1, \dots, s_L) \mu(s_1, \dots, s_L) = \inf_{\mu \in \Pi(\rho)} \langle C, \mu \rangle, \quad (2.29)$$

where the angle bracket notation is introduced to indicate the suggested inner product, i.e., the inner product of  $L^2(\{0, 1\}^L)$ . This is precisely the form of a MMOT problem, namely, minimization of a linear functional of a joint probability measure subject to constraints on all of the marginals of the measure [80]. Note that the dimension of the feasible space for this problem is exponential in  $L$ , rendering infeasible any direct approach based on the formulation as a general MMOT, at least for  $L$  of moderate size.

Nonetheless, we remark that in this exact formulation,  $\nabla_{\rho} E_{\text{sce}}[\rho]$  is the derivative of the optimal value of a convex optimization problem (in particular, a linear program) with respect to a variation of its constraints. This quantity can be obtained in terms of the variables *dual* to the varied constraints [17]. In the setting of MMOT, these dual variables are known as the Kantorovich potentials [104]. We will discuss the duality theory of our SDP relaxations in detail later on.

Despite the fact that it is possible to formulate our problem as a general MMOT problem, doing so loses the important structure of our *pairwise* cost. To wit, recall that the diagonal entries of  $v$  are set to zero,  $C$  can be written

$$C(s_1, \dots, s_L) = \sum_{p \neq q} v_{pq} s_p s_q =: \sum_{p \neq q} C_{pq}(s_p, s_q).$$

Hence the sum can be taken over  $p \neq q$ . Accordingly, the objective function of (2.24) can be written as

$$E_{\text{sce}}[\rho] = \inf_{\mu \in \Pi(\rho)} \sum_{p \neq q} \langle C_{pq}, \mu_{pq}^{(2)} \rangle, \quad (2.30)$$

where angle brackets are now used to indicate the suggested inner product, i.e., that of  $L^2(\{0, 1\}^2)$ , and where the 2-marginals  $\mu_{pq}^{(2)}$  are defined implicitly in terms of  $\mu$  by marginalizing out all components other than  $p, q$ , i.e., by

$$\mu_{pq}^{(2)}(s_p, s_q) := \sum_{s_1, \dots, s_L \setminus \{s_p, s_q\}} \mu(s_1, \dots, s_L). \quad (2.31)$$

Later we also identify  $\mu_{pq}^{(2)}$  with the  $2 \times 2$  matrix

$$\mu_{pq}^{(2)} = \begin{bmatrix} \mu_{pq}^{(2)}(0, 0) & \mu_{pq}^{(2)}(0, 1) \\ \mu_{pq}^{(2)}(1, 0) & \mu_{pq}^{(2)}(1, 1) \end{bmatrix}, \quad (2.32)$$

and we do likewise for  $C_{pq}$ . Using the matrix notation (and the symmetry of  $C_{pq}$ ), it follows that

$$E_{\text{sce}}[\rho] = \inf_{\mu \in \Pi(\rho)} \sum_{p \neq q} \text{Tr}[C_{pq} \mu_{pq}^{(2)}], \quad (2.33)$$

where ‘Tr’ indicates the matrix trace.

At first glance, it might seem that one may achieve a significant reduction of complexity by directly changing the optimization variable in Eq. (2.30) from  $\mu$  to  $\{\mu_{pq}^{(2)}\}_{p, q=1}^L$ . However, extra constraints would then need to be enforced in order to relate the different 2-marginals; i.e., the two-marginals must be jointly *representable* in the sense that all of them could simultaneously be yielded from a single joint probability measure on  $\{0, 1\}^L$ .

### 3 Convex relaxation

In this section, we show that a relaxation of the representability condition implicit in Eq. (2.30) allows us to formulate a tractable optimization problem in terms of the  $\{\mu_{pq}^{(2)}\}_{p,q=1}^L$  alone. In fact, this optimization problem will be a semidefinite program (SDP).

#### 3.1 Primal problem

We now derive certain necessary constraints satisfied by 2-marginals  $\{\mu_{pq}^{(2)}\}_{p,q=1}^L$  that are obtained from a probability measure  $\mu$  on  $\{0, 1\}^L$ . In the following we adopt the notation

$$\mathbf{s} = (s_1, \dots, s_L) \in \{0, 1\}^L.$$

Then for any such  $\mathbf{s}$ , let  $e_{\mathbf{s}} : \{0, 1\}^L \rightarrow \mathbb{R}$  be the Dirac probability mass function on  $\{0, 1\}^L$  localized at  $\mathbf{s}$ , i.e.,

$$e_{\mathbf{s}}(\mathbf{s}') = \delta_{\mathbf{s}, \mathbf{s}'}$$

Note that we can also write  $e_{\mathbf{s}}$  as an  $L$ -tensor, i.e., an element of  $\mathbb{R}^{2 \times 2 \times \dots \times 2}$ , via

$$e_{\mathbf{s}} = e_{s_1} \otimes \dots \otimes e_{s_L},$$

where we adopt the (zero-indexing) convention  $e_0 = [1, 0]^\top$ ,  $e_1 = [0, 1]^\top$ .

Any probability measure on  $\{0, 1\}^L$  can be written as a convex combination of the  $e_{\mathbf{s}}$  since they are the extreme points of the set of probability measures; in particular we can write a probability density  $\mu \in \Pi(\rho)$  as

$$\mu = \sum_{\mathbf{s}} a_{\mathbf{s}} e_{\mathbf{s}}, \quad \text{where} \quad \sum_{\mathbf{s}} a_{\mathbf{s}} = 1, \quad a_{\mathbf{s}} \geq 0. \quad (3.1)$$

From the definitions of the 1- and 2-marginals (2.25), (2.31), it follows that

$$\mu_p^{(1)} = \sum_{\mathbf{s}} a_{\mathbf{s}} e_{s_p}, \quad \mu_{pq}^{(2)} = \sum_{\mathbf{s}} a_{\mathbf{s}} e_{s_p} \otimes e_{s_q} = \sum_{\mathbf{s}} a_{\mathbf{s}} e_{s_p} e_{s_q}^\top. \quad (3.2)$$

Now define

$$M = M(\{a_{\mathbf{s}}\}) = \sum_{\mathbf{s}} a_{\mathbf{s}} \begin{bmatrix} e_{s_1} \\ \vdots \\ e_{s_L} \end{bmatrix} [e_{s_1}^\top \cdots e_{s_L}^\top], \quad (3.3)$$

Then by Eq. (3.2),  $M$  is the matrix of  $2 \times 2$  blocks  $M_{pq}$  given by

$$M_{pq} = \begin{cases} \text{diag}(\mu_p^{(1)}), & p = q, \\ \mu_{pq}^{(2)}, & p \neq q. \end{cases} \quad (3.4)$$

Accordingly we write  $M = (M_{pq}) \in \mathbb{R}^{(2L) \times (2L)}$ . Then let  $C = (C_{pq}) \in \mathbb{R}^{(2L) \times (2L)}$  be the matrix of the  $2 \times 2$  blocks  $C_{pq}$  defined above, which specifies the pairwise cost on each pair of marginals<sup>18</sup>. Observe that the value of the objective function of Eq. (2.33) can in fact be rewritten as

$$\sum_{p \neq q} \text{Tr}[C_{pq} \mu_{pq}^{(2)}] = \text{Tr}[CM].$$

Then the MMOT problem Eq. (2.33) can be equivalently rephrased as

$$\begin{aligned} E_{\text{sce}}[\rho] = & \underset{M \in \mathbb{R}^{(2L) \times (2L)}, \{a_{\mathbf{s}}\}_{\mathbf{s} \in \{0,1\}^L}}{\text{minimize}} && \text{Tr}(CM) \\ & \text{subject to} && M = \sum_{\mathbf{s}} a_{\mathbf{s}} \begin{bmatrix} e_{\mathbf{s}_1} \\ \vdots \\ e_{\mathbf{s}_L} \end{bmatrix} [e_{\mathbf{s}_1}^\top \cdots e_{\mathbf{s}_L}^\top], && (3.5) \\ & && M_{pp} = \text{diag}(\mu_p^{(1)}) \text{ for all } p = 1, \dots, L, \\ & && \sum_{\mathbf{s}} a_{\mathbf{s}} = 1, \quad a_{\mathbf{s}} \geq 0 \text{ for all } \mathbf{s} \in \{0,1\}^L. \end{aligned}$$

Note that in our application to SCE, we have fixed

$$\mu_p^{(1)} = \begin{bmatrix} 1 - \rho_p \\ \rho_p \end{bmatrix}$$

in advance, i.e.,  $\mu_p^{(1)}$  is *not* an optimization variable.

At this point, our reformulation of the problem has not alleviated its exponential complexity; indeed, note that  $\{a_{\mathbf{s}}\}_{\mathbf{s} \in \{0,1\}^L}$  is a vector of size  $2^L$ . However, the reformulation does suggest a way to reduce the complexity by accepting some approximation. In fact, we will omit  $\{a_{\mathbf{s}}\}_{\mathbf{s} \in \{0,1\}^L}$  entirely from the optimization, retaining only  $M$  as an optimization variable and enforcing several necessary constraints on  $M$  that are satisfied by the solution of the exact problem.

First, note from the constraint (3.5) that  $M$  is both entry-wise nonnegative (written  $M \geq 0$ ) and positive semidefinite (written  $M \succeq 0$ ). Second, the fact that the 1-marginals can be written in terms of the 2-marginals imposes additional *local consistency* constraints on  $M$ . Indeed, with  $\mathbf{1}_2 \in \mathbb{R}^2$  denoting the vector of all ones, we can write

$$\mu_{pq}^{(2)} \mathbf{1}_2 = \mu_p^{(1)}, \quad p \neq q, \quad (3.6)$$

from which it follows that

$$M_{pq} \mathbf{1}_2 = \begin{bmatrix} 1 - \rho_p \\ \rho_p \end{bmatrix}, \quad p, q = 1, \dots, L. \quad (3.7)$$

---

<sup>18</sup>Without loss of generality, one can assume  $C_{pp} = 0$ .

Then we obtain the relaxation

$$\begin{aligned}
E_{\text{sce}}[\rho] \geq E_{\text{sce}}^{\text{sdp}}[\rho] := & \underset{M \in \mathbb{R}^{(2L) \times (2L)}}{\text{minimize}} && \text{Tr}(CM) && (3.8) \\
& \text{subject to} && M \succeq 0, \\
& && M_{pq} \geq 0 \text{ for all } p, q = 1, \dots, L \text{ (} p \neq q \text{)}, \\
& && M_{pq} \mathbf{1}_2 = \mu_p^{(1)} \text{ for all } p, q = 1, \dots, L \text{ (} p \neq q \text{)}, \\
& && M_{pp} = \text{diag}(\mu_p^{(1)}) \text{ for all } p = 1, \dots, L.
\end{aligned}$$

Again,  $\mu_p^{(1)}$  is *not* an optimization variable. It is actually helpful to reformulate the primal 2-marginal SDP (3.8) as

$$E_{\text{sce}}^{\text{sdp}}[\rho] = \underset{M \in \mathbb{R}^{(2L) \times (2L)}}{\text{minimize}} \text{Tr}(CM) \quad (3.9)$$

$$\text{subject to} \quad M \succeq 0, \quad (3.10)$$

$$M_{pq} \geq 0 \text{ for all } p, q = 1, \dots, L \text{ (} p < q \text{)}, \quad (3.11)$$

$$M_{pq} \mathbf{1}_2 = \mu_p^{(1)} \text{ for all } p, q = 1, \dots, L \text{ (} p < q \text{)}, \quad (3.12)$$

$$M_{pq}^\top \mathbf{1}_2 = \mu_q^{(1)} \text{ for all } p, q = 1, \dots, L \text{ (} p < q \text{)}, \quad (3.13)$$

$$M_{pp} = \text{diag}(\mu_p^{(1)}) \text{ for all } p = 1, \dots, L. \quad (3.14)$$

Note that this formulation is equivalent to (3.8), given the symmetry of  $M$  (implicit in the notation  $M \succeq 0$ ). However, the new formulation removes a few redundant constraints and will help us derive a more intuitive dual problem. The problem (3.9) will be referred to as the primal 2-marginal SDP, or the *primal problem* for short. Note that the optimal value of the primal problem is in fact attained because the constraints (3.10)-(3.14) define a compact feasible set.

Reflecting back on the derivation, we caution that replacing  $E_{\text{sce}}[\rho]$  with  $E_{\text{sce}}^{\text{sdp}}[\rho]$  comes at a price. Since we only enforce certain necessary conditions on  $M$ , the 2-marginals that we recover from  $M$  may not in fact be the 2-marginals of a joint probability measure on  $\{0, 1\}^L$ . Thus  $E_{\text{sce}}^{\text{sdp}}[\rho]$  should in general only be expected to be a lower-bound to  $E_{\text{sce}}[\rho]$ , though we will see that the error is often small in practice.

## 3.2 Dual problem

As detailed in Section 2.2.1, in order to implement the SCF for Kohn-Sham SCE it is necessary to compute  $\nabla_\rho E_{\text{sce}}[\rho]$ . After replacing the density functional  $E_{\text{sce}}[\rho]$  with the efficient approximation  $E_{\text{sce}}^{\text{sdp}}[\rho]$ , the same derivation motivates us to compute  $\nabla_\rho E_{\text{sce}}^{\text{sdp}}[\rho]$ . This quantity can be obtained by examining the convex duality of our primal 2-marginal SDP.

We let  $Y \succeq 0$  be the variable dual to the constraint (3.10),  $Z_{pq} \geq 0$  be dual to (3.11),  $\phi_{pq}$  be dual to (3.12),  $\psi_{pq}$  be dual to (3.13), and finally let  $X_p$  be dual to (3.14). Note that  $Z_{pq} \in \mathbb{R}^{2 \times 2}$  and  $\phi_{pq}, \psi_{pq} \in \mathbb{R}^2$  for each  $p < q$ , and  $X_p \in \mathbb{R}^{2 \times 2}$  for each  $p$ .

Then our formal Lagrangian is of the form

$$\mathcal{L}(M, Y, \{Z_{pq}, \phi_{pq}, \psi_{pq}\}_{p < q}, \{X_p\}),$$

where the domains of  $M$  is the set of symmetric  $2L \times 2L$  matrices (equivalently, it is convenient to think of  $M$  as depending only on its upper-block-triangular part), and the dual variables are as specified above (i.e., only  $Y \succeq 0$  and  $Z_{pq} \geq 0$  are constrained), and more specifically we have (omitting the arguments of  $\mathcal{L}$  from the notation)

$$\begin{aligned} \mathcal{L} &= \text{Tr}(CM) - \text{Tr}(YM) \\ &\quad - 2 \sum_{p < q} [\text{Tr}(Z_{pq}^\top M_{pq}) + \phi_{pq}^\top (M_{pq} \mathbf{1}_2 - \mu_p^{(1)}) + \psi_{pq}^\top (M_{pq}^\top \mathbf{1}_2 - \mu_q^{(1)})] \\ &\quad - \sum_p \text{Tr}(X_p^\top [M_{pp} - \text{diag}(\mu_p^{(1)})]). \end{aligned} \tag{3.15}$$

It is helpful to realize the identities

$$\phi_{pq}^\top M_{pq} \mathbf{1}_2 = \text{Tr}(M_{pq} [\mathbf{1}_2 \phi_{pq}^\top]), \quad \psi_{pq}^\top M_{pq}^\top \mathbf{1}_2 = \text{Tr}(M_{pq} [\psi_{pq}^\top \mathbf{1}_2^\top]).$$

Then, recognizing that  $C = C^\top$  and  $Y = Y^\top$  (so that  $C_{pq}^\top = C_{qp}$  and  $Y_{pq}^\top = Y_{qp}$ ), minimization over  $M$  of the Lagrangian (3.15) yields the dual problem

$$\begin{aligned} &\underset{Y, \{Z_{pq}, \phi_{pq}, \psi_{pq}\}_{p < q}, \{X_p\}}{\text{maximize}} && \sum_p \text{Tr}(X_p^\top \text{diag}(\mu_p^{(1)})) + 2 \sum_{p < q} (\phi_{pq}^\top \mu_p^{(1)} + \psi_{pq}^\top \mu_q^{(1)}) \\ &\text{subject to} && Y \succeq 0, \\ &&& Z_{pq} \geq 0 \text{ for } p < q, \tag{3.16} \\ &&& C_{pq} - Y_{pq} - Z_{pq} - \phi_{pq} \mathbf{1}_2^\top - \mathbf{1}_2 \psi_{pq}^\top = 0 \text{ for } p < q, \tag{3.17} \\ &&& C_{pp} - Y_{pp} - X_p^\top = 0. \tag{3.18} \end{aligned}$$

Observe that the variables  $Z_{pq}$  can be removed by combining constraints (3.16) and (3.17) to yield

$$C_{pq} - Y_{pq} - \phi_{pq} \mathbf{1}_2^\top - \mathbf{1}_2 \psi_{pq}^\top \geq 0.$$

Moreover,  $X_p$  can be removed simply by substituting  $X_p = -Y_{pp}$  into the objective function (recall that  $C_{pp} = 0$ ). These reductions yield

$$\underset{Y, \{\phi_{pq}, \psi_{pq}\}_{p < q}}{\text{maximize}} \quad 2 \sum_{p < q} (\phi_{pq} \cdot \mu_p^{(1)} + \psi_{pq} \cdot \mu_q^{(1)}) - \sum_{p, s} Y_{pp}(s, s) \mu_p^{(1)}(s) \tag{3.19}$$

$$\text{subject to} \quad Y \succeq 0, \tag{3.20}$$

$$\phi_{pq} \mathbf{1}_2^\top + \mathbf{1}_2 \psi_{pq}^\top \leq C_{pq} - Y_{pq} \text{ for } p < q. \tag{3.21}$$

Here we think of  $Y_{pp}(s, s)$  as the  $(s, s)$  entry of the  $2 \times 2$  matrix  $Y_{pp}$ , and likewise  $\mu_p^{(1)}(s)$  is the  $s$ -th entry of  $\mu_p^{(1)}$ .



The dual problem may be interpreted as follows. Observe that for  $Y$  fixed (e.g., fixed to its optimal value), the maximization problem decouples into a set of independent maximization problems for each pair of marginals. We think of  $\tilde{C}_{pq} := C_{pq} - Y_{pq}$  as defining an *effective* cost function for each pair of marginals. Then the decoupled problem for a pair  $p < q$  is *exactly* the Kantorovich dual problem in standard (i.e., not multi-marginal) optimal transport, specified by cost function  $\tilde{C}_{pq}$  and marginals  $\mu_p^{(1)}$ ,  $\mu_q^{(1)}$  [104]. In other words, after fixing  $Y$ , our problem decouples into independent *standard* optimal transport problems for each pair of marginals. Nonetheless, these problems are in turn themselves coupled via the optimization over  $Y \succeq 0$ .

Recall that we wanted to compute  $\nabla_\rho E_{\text{sce}}^{\text{sdp}}[\rho]$ . Assuming that strong duality holds, as shall be established later, the optimal value of the dual problem (3.19) is in fact equal to  $E_{\text{sce}}^{\text{sdp}}[\rho]$ . (Recall that here we think of the 1-marginals  $\mu_p^{(1)} = [1 - \rho_p, \rho_p]^\top$  as being defined in terms of  $\rho$ .) Hence we can compute derivatives by evaluating the gradient of the objective function (3.19) with respect to  $\rho$  at the *optimizer* ( $Y, \{\phi_{pq}, \psi_{pq}\}_{p \neq q}$ ). (If the optimizer is not unique, then in general we will get a subgradient [91].)

To carry out this program, first note that  $\frac{\partial}{\partial \rho_r} \mu_p^{(1)} = \delta_{pr} [-1, 1]^\top$ . Therefore the partial derivative of the objective function (3.19) with respect to  $\rho_r$  yields

$$\frac{\partial E_{\text{sce}}^{\text{sdp}}[\rho]}{\partial \rho_r} = 2 \sum_{q>r} [\phi_{rq}(1) - \phi_{rq}(0)] + 2 \sum_{p<r} [\psi_{pr}(1) - \psi_{pr}(0)] - [Y_{rr}(1, 1) - Y_{rr}(0, 0)].$$

If one extends the definition of  $\phi_{pq}, \psi_{pq}$  to  $p > q$  via the stipulation  $\phi_{pq} = \psi_{qp}$ , then one has

$$\frac{\partial E_{\text{sce}}^{\text{sdp}}[\rho]}{\partial \rho_r} = \sum_{p \neq r} [\phi_{rp}(1) - \phi_{rp}(0)] - [Y_{rr}(1, 1) - Y_{rr}(0, 0)].$$

### 3.3 Strong duality and dual attainment

In order to faithfully compute the SCE energy and potential via the dual problem (3.19), we need to verify that the dual problem satisfies strong duality, i.e., that the duality gap defined by the difference between the infimum of Eq. (3.9) and the supremum of Eq. (3.19) is zero. In fact, since the domain of the primal problem is compact, Sion's minimax theorem [51] immediately guarantees that the duality gap is zero. We state this result as a lemma:

**Lemma 1.** *The primal and dual problems (3.9) and (3.19), respectively, have the same (finite) optimal value.*

However, in order to compute the SCE potential, we actually require not only that the duality gap is zero, but also that the supremum in the dual problem is *attained*. One might hope to verify Slater's condition [17], which provides a standard method for verifying both strong duality and such 'dual attainment' simultaneously.

The trouble is that Slater's condition requires the existence of a feasible *interior* point  $M$ , i.e., a point  $M$  satisfying  $M \succ 0$  and  $M_{pq} > 0$  for all  $p \neq q$ . This scenario is in fact impossible since for example the vector

$$[\mathbf{1}_2^\top \quad -\mathbf{1}_2^\top \quad 0 \quad \cdots \quad 0]^\top \in \mathbb{R}^{2L} \quad (3.22)$$

lies in the null space of any feasible  $M$ , hence  $M \succ 0$  *never* holds for feasible  $M$ .

Instead of using Slater's condition, we will prove dual attainment via a very careful study of the structure of the dual problem.

**Theorem 2.** *The optimal value of the dual 2-marginal SDP (3.19) is attained. By Lemma 1, this optimal value is equal to the optimal value of the primal 2-marginal SDP (3.9).*

*Proof.* Without loss of generality we assume

$$0 < \rho_p < 1, \quad p = 1, \dots, L. \quad (3.23)$$

To see why this assumption can be made, observe that if  $\rho_p \in \{0, 1\}$  for some  $p$ , then attainment for the dual problem (3.19) can be reduced to attainment for a strictly smaller dual 2-marginal SDP. We leave further details of such a reduction to the reader.) Also, for later reference, we let  $F(Y, \{\phi_{pq}, \psi_{pq}\}_{p < q})$  denote the objective function (3.19), and we let  $\mathcal{D}$  denote the feasible domain defined by the constraints (3.20), (3.21).

Now to get started, observe that if we fix  $Y \succeq 0$  and view (3.19) as an optimization problem over  $\{\phi_{pq}, \psi_{pq}\}_{p < q}$  only, the resulting problem is in fact a linear program. Let us call this the  $Y$ -program, more specifically:

$$\begin{aligned} & \underset{\{\phi_{pq}, \psi_{pq}\}_{p < q}}{\text{maximize}} && 2 \sum_{p < q} (\phi_{pq} \cdot \mu_p^{(1)} + \psi_{pq} \cdot \mu_q^{(1)}) - \sum_{p, s} Y_{pp}(s, s) \mu_p^{(1)}(s) \\ & \text{subject to} && \phi_{pq} \mathbf{1}_2^\top + \mathbf{1}_2 \psi_{pq}^\top \leq C_{pq} - Y_{pq} \text{ for } p < q. \end{aligned}$$

In fact we may consider the  $Y$ -program for *any* matrix  $Y$ , and this will slightly simplify some discussion later. Observe that each  $Y$ -program is feasible, and the optimal values  $f(Y)$  of all  $Y$ -programs are finite. Since they are linear programs, this means that the optimal values of the  $Y$ -programs can be attained. Thus for each  $Y$ , there exist  $\phi_{pq}^*(Y)$ ,  $\psi_{pq}^*(Y)$  for  $p < q$  which optimize the  $Y$ -program, i.e., attain the value  $f(Y)$ . By construction  $f(Y)$  is concave, hence continuous, in  $Y$ .

Now let  $d_0 = f(0)$ , so  $d^* \geq d_0$ , where  $d^*$  is the optimal value of the dual problem (3.19). Hence the feasible set of (3.19) could be refined to  $S \cap \mathcal{D}$ , where

$$S := \{Y \succeq 0 : f(Y) \geq d_0\},$$

without altering the optimal value. Now if  $S$  were compact, then the lemma would follow. To see this, note that since  $d^* < \infty$  (which follows from weak duality), we

could take an optimizing sequence  $(Y^{(k)}, \{\phi_{pq}^{(k)}, \psi_{pq}^{(k)}\}_{p<q})$  for (3.19), where  $Y^{(k)} \in S \cap \mathcal{D}$ . Then by compactness we could find a subsequence of  $Y^{(k)}$  converging to some  $Y^*$ . By the continuity of  $f$ , then  $f(Y^*) = d^*$ . Then it would follow that the optimum is attained at the point  $(Y^*, \{\phi_{pq}^*(Y^*), \psi_{pq}^*(Y^*)\}_{p<q})$ .

Unfortunately,  $S$  is not compact, but we will find a further constraint that does yield a compact feasible set without altering the optimal value. Then the preceding argument will complete the proof.

To further constrain the feasible set, we will observe a transformation of  $Y$  that preserves the value of  $f(Y)$ , then ‘mod out’ by this transformation. To this end, first note that via the discussion of Kantorovich duality following (3.19) we can in fact write

$$f(Y) = - \sum_{p=1}^L \text{Tr} [Y_{pp} \text{diag}(\mu_p^{(1)})] + \sum_{p,q=1}^L \mathbf{OT}_{pq}(C_{pq} - Y_{pq}),$$

where  $\mathbf{OT}_{pq}(A)$  is the optimal cost of the *standard* optimal transport problem with cost matrix  $A$  and marginals  $\mu_p^{(1)}, \mu_q^{(1)}$ .

Then let  $P \in \mathbb{R}^{(2L) \times (L-1)}$  be defined by

$$P := \begin{bmatrix} \mathbf{1}_2 & & & & & \\ -\mathbf{1}_2 & \mathbf{1}_2 & & & & \\ & -\mathbf{1}_2 & \ddots & & & \\ & & \ddots & \mathbf{1}_2 & & \\ & & & & -\mathbf{1}_2 & \end{bmatrix}, \quad (3.24)$$

and let its columns be denoted  $P_i$  for  $i = 1, \dots, L-1$ . Then we claim that

$$f(Y) = f(Y + P_i v^\top + v P_i^\top) \quad (3.25)$$

for any  $Y, v \in \mathbb{R}^{2L}$ , and any  $i = 1, \dots, L-1$ . To prove this, write

$$v = [v_1^\top \cdots v_L^\top]^\top,$$

where  $v_q \in \mathbb{R}^2$  for  $q = 1, \dots, L$ . Then observe that, via the discussion of Kantorovich duality following the statement (3.19) of the dual problem, we can in fact write

$$f(Y) = - \sum_{p=1}^L \text{Tr} [Y_{pp} \text{diag}(\mu_p^{(1)})] + 2 \sum_{p<q} \mathbf{OT}_{pq}(C_{pq} - Y_{pq}),$$

where  $\mathbf{OT}_{pq}(A)$  is the optimal cost of the *standard* optimal transport problem with cost matrix  $A$  and marginals  $\mu_p^{(1)}, \mu_q^{(1)}$ .

Then compute

$$f(Y + P_i v^\top) = - \sum_{p=1}^L \text{Tr} [Y_{pp} \text{diag}(\mu_p^{(1)})] - \text{Tr} [\mathbf{1}_2 v_i^\top \text{diag}(\mu_i^{(1)})]$$

$$\begin{aligned}
& + \operatorname{Tr} \left[ \mathbf{1}_2 v_{i+1}^\top \operatorname{diag}(\mu_{i+1}^{(1)}) \right] + 2 \sum_{p < q, p \notin \{i, i+1\}} \mathbf{OT}_{pq}(C_{pq} - Y_{pq}) \\
& + 2 \sum_{q=i+1}^L \mathbf{OT}_{iq}(C_{iq} - Y_{iq} - \mathbf{1}_2 v_q^\top) \\
& + 2 \sum_{q=i+2}^L \mathbf{OT}_{i+1,q}(C_{i+1,q} - Y_{i+1,q} + \mathbf{1}_2 v_q^\top).
\end{aligned}$$

Now

$$\operatorname{Tr} \left[ \mathbf{1}_2 v_i^\top \operatorname{diag}(\mu_i^{(1)}) \right] = v_i \cdot \mu_i^{(1)}, \quad \operatorname{Tr} \left[ \mathbf{1}_2 v_{i+1}^\top \operatorname{diag}(\mu_{i+1}^{(1)}) \right] = v_{i+1} \cdot \mu_{i+1}^{(1)},$$

and moreover it is not hard to see that

$$\mathbf{OT}_{pq}(A + \mathbf{1}_2 x^\top) = \mathbf{OT}_{pq}(A) + x \cdot \mu_q^{(1)}$$

for any  $A \in \mathbb{R}^{2 \times 2}, x \in \mathbb{R}^2$ , hence

$$\begin{aligned}
f(Y + P_i v^\top) & = - \sum_{p=1}^L \operatorname{Tr} [Y_{pp} \operatorname{diag}(\mu_p^{(1)})] - v_i \cdot \mu_i^{(1)} + v_{i+1} \cdot \mu_{i+1}^{(1)} \\
& + 2 \sum_{p < q} \mathbf{OT}_{pq}(C_{pq} - Y_{pq}) - 2 \sum_{q=i+1}^L v_q \cdot \mu_q^{(1)} + 2 \sum_{q=i+2}^L v_q \cdot \mu_q^{(1)} \\
& = f(Y) - v_i \cdot \mu_i^{(1)} - v_{i+1} \cdot \mu_{i+1}^{(1)}.
\end{aligned}$$

Similarly

$$\begin{aligned}
f(Y + v P_i^\top) & = - \sum_{p=1}^L \operatorname{Tr} [Y_{pp} \operatorname{diag}(\mu_p^{(1)})] - \operatorname{Tr} [v_i \mathbf{1}_2^\top \operatorname{diag}(\mu_i^{(1)})] \\
& + \operatorname{Tr} [v_{i+1} \mathbf{1}_2^\top \operatorname{diag}(\mu_{i+1}^{(1)})] + 2 \sum_{p < q, q \notin \{i, i+1\}} \mathbf{OT}_{pq}(C_{pq} - Y_{pq}) \\
& + 2 \sum_{p=1}^{i-1} \mathbf{OT}_{pi}(C_{pi} - Y_{pi} - v_p \mathbf{1}_2^\top) \\
& + 2 \sum_{p=1}^i \mathbf{OT}_{p,i+1}(C_{p,i+1} - Y_{p,i+1} + v_p \mathbf{1}_2^\top) \\
& = f(Y) + v_i \cdot \mu_i^{(1)} + v_{i+1} \cdot \mu_{i+1}^{(1)}.
\end{aligned}$$

Since the identities

$$f(Y + P_i v^\top) = f(Y) - v_i \cdot \mu_i^{(1)} - v_{i+1} \cdot \mu_{i+1}^{(1)}, \quad f(Y + v P_i^\top) = f(Y) + v_i \cdot \mu_i^{(1)} + v_{i+1} \cdot \mu_{i+1}^{(1)}$$

hold for arbitrary  $Y$ , the claim Eq. (3.25) is proven.

Then from Eq. (3.25) it follows that

$$f(Y) = f(Y + PB + B^\top P^\top) \quad (3.26)$$

for arbitrary  $B \in \mathbb{R}^{(L-1) \times (2L)}$ .

Now let  $Q \in \mathbb{R}^{(2L) \times (L+1)}$  be defined by

$$Q = \begin{bmatrix} w_1 & 0 & \cdots & 0 & w_2 \\ 0 & w_1 & & \vdots & \vdots \\ \vdots & & \ddots & & \\ 0 & \cdots & & w_1 & w_2 \end{bmatrix}, \quad w_1 = \frac{1}{2} \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \quad w_2 = \frac{1}{2} \mathbf{1}_2,$$

and observe that  $Q$  is chosen so that each column of  $Q$  is orthogonal to each column of  $P$ . Moreover  $P$  and  $Q$  both have full rank, so it follows that  $R := [Q, P]$  is invertible.

Then for fixed  $Y$ , consider

$$\hat{Y} = R^\top Y R = \begin{pmatrix} Q^\top Y Q & Q^\top Y P \\ P^\top Y Q & P^\top Y P \end{pmatrix}.$$

We aim to choose  $B$  such that

$$R^\top (PB + B^\top P^\top) R = \begin{pmatrix} 0 & 0 \\ P^\top P B Q & P^\top P B P \end{pmatrix} + \begin{pmatrix} 0 & Q^\top B^\top P^\top P \\ 0 & P^\top B^\top P^\top P \end{pmatrix}$$

cancels  $\hat{Y}$  on all but the top-left block. Using  $Q^\top P = 0$  (and  $P^\top Q = 0$ ), one can readily check that such a choice is given by

$$-B = (P^\top P)^{-1} \hat{Y}_{21} (Q^\top Q)^{-1} Q^\top + \frac{1}{2} (P^\top P)^{-1} \hat{Y}_{22} (P^\top P)^{-1} P^\top.$$

By the identity (3.26), it follows that we can further restrict the feasible set by intersecting with

$$S' = \left\{ Y : R^\top Y R = \begin{pmatrix} * & 0 \\ 0 & 0 \end{pmatrix} \succeq 0, f(Y) \geq d_0 \right\}. \quad (3.27)$$

In fact  $S'$  is compact, and the proof is complete pending the proof of this claim, to which we now turn.

Observe that for  $(Y, \{\phi_{pq}, \psi_{pq}\}_{p < q})$  feasible, we may multiply Eq. (3.21) from the left by  $(\mu_p^{(1)})^\top$  and from the right by  $\mu_q^{(1)}$  to obtain

$$\begin{aligned} \phi_{pq} \cdot \mu_p^{(1)} + \psi_{pq} \cdot \mu_q^{(1)} &\leq (\mu_p^{(1)})^\top [C_{pq} - Y_{pq}] (\mu_q^{(1)}) \\ &= (\mu_q^{(1)})^\top [C_{pq} - Y_{pq}]^\top (\mu_p^{(1)}) \end{aligned}$$

$$= \operatorname{Tr} \left( [C_{pq} - Y_{pq}]^\top (\mu_p^{(1)}) (\mu_q^{(1)})^\top \right).$$

By substituting this inequality into the objective function  $F(Y, \{\phi_{pq}, \psi_{pq}\}_{p < q})$  as defined in (3.19), we see that

$$F(Y, \{\phi_{pq}, \psi_{pq}\}_{p < q}) \leq \operatorname{Tr}(CM) - \operatorname{Tr}(YM).$$

for  $(Y, \{\phi_{pq}, \psi_{pq}\}_{p < q})$  feasible, where

$$M_{pq} := \begin{cases} \operatorname{diag}(\mu_p^{(1)}), & p = q \\ (\mu_p^{(1)}) (\mu_q^{(1)})^\top, & p \neq q. \end{cases}$$

It follows then that

$$f(Y) \leq \operatorname{Tr}(CM) - \operatorname{Tr}(YM).$$

In fact  $M$  can be written  $M = Q\widetilde{M}Q^\top$ , where  $\widetilde{M} \succ 0$ . This can be verified directly by taking

$$\widetilde{M} = \begin{bmatrix} \widetilde{\rho}_1 \\ \vdots \\ \widetilde{\rho}_L \\ 1 \end{bmatrix} [\widetilde{\rho}_1 \quad \cdots \quad \widetilde{\rho}_L \quad 1] + \operatorname{diag}([1 - \widetilde{\rho}_1^2 \quad \cdots \quad 1 - \widetilde{\rho}_L^2 \quad 0]),$$

with

$$\widetilde{\rho}_p = 1 - 2\rho_p, \quad p = 1, \dots, L.$$

Note that  $\widetilde{M} \succ 0$  by the assumption (3.23). Hence

$$f(Y) \leq \operatorname{Tr}(CM) - \operatorname{Tr}(Q^\top Y Q \widetilde{M}).$$

Now since  $\widetilde{M} \succ 0$ , there exists a scalar  $K > 0$  such that if  $Y \succeq 0$  and  $Q^\top Y Q \not\leq K$ , then  $f(Y) < d_0$ . But  $Q^\top Y Q$  is the upper-left block of  $R^\top Y R$ , so it follows from the definition (3.27) of  $S'$  that that

$$S' \subset \left\{ Y : R^\top Y R = \begin{pmatrix} A & 0 \\ 0 & 0 \end{pmatrix}, 0 \preceq A \preceq K \right\},$$

from which it follows that  $S'$  is compact, and the proof is complete.  $\square$

*Remark 3.* Note that the proof of Theorem 2 guarantees that the domain of the dual problem (3.19) can be restricted to  $Y$  of the form  $Y = Q\widetilde{Y}Q^\top$ , yielding a ‘reduced’ dual problem in which  $\widetilde{Y}$  replaces  $Y$  as an optimization variable. In fact, one can also verify directly that any  $M$  feasible for the primal problem (3.9) satisfies  $MP = 0$ , hence the domain of the primal problem can be restricted to  $M$  of the form  $M = Q\widetilde{M}Q^\top$ , likewise yielding a reduced primal problem.

But despite this apparent symmetry, the latter observation need not imply the former in a more general SDP setting, and the arguments given in the proof of Theorem 2, which use more of the specific structure of our problem, do appear to be necessary to the proof of dual attainment for this problem.

Moreover, observe with caution that the dual of such a reduced primal problem is *not* the reduced dual problem!

## 4 Tighter lower bound via 3-marginals

In this section, we further tighten the convex relaxation proposed in Section 3 with a formulation that additionally involves the 3-marginals.

One defines the 3-marginals  $\mu_{pqr}^{(3)}$  (for  $p, q, r$ , distinct) induced by a probability measure  $\mu$  on  $\{0, 1\}^L$  via

$$\mu_{pqr}^{(3)}(s_p, s_q, s_r) := \sum_{s_1, \dots, s_L \setminus \{s_p, s_q, s_r\}} \mu(s_1, \dots, s_L). \quad (4.1)$$

There is no 3-marginal analog known to us of the semidefinite constraint that can be enforced using the 2-marginals. However, we can nonetheless use the 3-marginals to enforce additional necessary *local consistency* constraints. Indeed, the 2-marginals can themselves be written in terms of the 3-marginals via

$$\mu_{pq}^{(2)}(s_p, s_q) = \sum_{s_r} \mu_{pqr}^{(3)}(s_p, s_q, s_r). \quad (4.2)$$

Accordingly, we will include  $K = \{K_{pqr}\}$  for distinct  $p, q, r$  as optimization variables for the 3-marginals. Note that based on Eq. (4.1) we can enforce that  $K$  is *symmetric*, by which we mean that

$$K_{pqr}(s_p, s_q, s_r) = K_{\sigma(p)\sigma(q)\sigma(r)}(s_{\sigma(p)}, s_{\sigma(q)}, s_{\sigma(r)})$$

for any permutation  $\sigma$  on the letters  $\{p, q, r\}$ . If we were to extend  $K_{pqr}$  by zeros to  $p, q, r$  not distinct, then we could think of  $K \in \mathbb{R}^{(2L) \times (2L) \times (2L)}$  as a symmetric 3-tensor, with  $(p, q, r)$ -th  $2 \times 2 \times 2$  block given by  $K_{pqr}$ . In principle the imposition of symmetry removes some redundancy in the specification of  $K$ .

Then we arrive at the following *3-marginal SDP*:

$$\begin{aligned} & \underset{M \in \mathbb{R}^{(2L) \times (2L)}, K \in \mathbb{R}^{(2L) \times (2L) \times (2L)}}{\text{minimize}} && \text{Tr}(CM) && (4.3) \\ & \text{subject to} && M \succeq 0, \\ & && M_{pq} \geq 0 \text{ for } p \neq q, \\ & && M_{pq} \mathbf{1}_2 = \mu_p^{(1)} \text{ for } p \neq q, \\ & && M_{pp} = \text{diag}(\mu_p^{(1)}) \text{ for all } p, \\ & && K \geq 0, K \text{ symmetric,} \end{aligned}$$

$$M_{pq}(s_p, s_q) = \sum_{s_r} K_{pqr}(s_p, s_q, s_r) \text{ for } p, q, r \text{ distinct.}$$

Note that the blocks  $K_{pqr}$  for  $p, q, r$  not distinct are superfluous and can be discarded in an efficient optimization.

For simplicity, we omit discussion of the duality of (4.3). Since only linear constraints have been added, most of the interesting features from the mathematical viewpoint have already been discussed above. Indeed, as in Section 3.2, we may derive the dual of the 3-marginal problem (4.3), and we may certify as in Section 3.3 that the 3-marginal problem satisfies strong duality and dual attainment.

## 5 General MMOT with pairwise cost

As has been suggested both explicitly and via the notation, almost all of our discussion of relaxation methods for MMOT can be applied to general MMOT problems with pairwise cost functions. The main caveat is that specific references to the fact that the 1-marginal state space has two elements should be suitably generalized. For clarity, we now recapitulate our methods for the general MMOT problem with pairwise cost. The reader interested in general MMOT should still see the earlier sections for derivations, discussions, and proofs. Here we only summarize the methods.

We will consider a problem with  $L$  marginals, written  $\mu_p^{(1)}$  for  $p = 1, \dots, L$ . These quantities are fixed in advance and never varied in the following discussion. We let  $N_p$  be the size of the state space of the  $p$ -th marginal, so  $\mu_p^{(1)}$  is a probability vector of length  $N_p$ . Note that the marginals need not all have the same state space, i.e.,  $N_p$  can depend on  $p$ . We write the  $p$ -th state space as  $\mathcal{X}_p := \{1, \dots, N_p\}$ . Then the joint state space is given by  $\mathcal{X} := \prod_{p=1}^L \mathcal{X}_p$ , and we write  $\text{Pr}_p$  for the  $p$ -th projection  $\mathcal{X} \rightarrow \mathcal{X}_p$ . Suppose that we are given a pairwise cost function  $C_{pq} \in \mathbb{R}^{N_p \times N_q}$  for each pair  $p \neq q$  of marginals. (Without loss of generality we assume  $C_{pp} = 0$ .) Then we consider the problem

$$\min_{\mu \in \mathcal{P}(\mathcal{X})} \sum_{(s_1, \dots, s_L) \in \mathcal{X}} \sum_{p, q=1}^L C_{pq}(s_p, s_q) \mu(s_1, \dots, s_L), \quad \text{s.t. } (\text{Pr}_p) \# \mu = \mu_p^{(1)}, \quad p = 1, \dots, L. \quad (5.1)$$

Here  $\mu : \mathcal{X} \rightarrow \mathbb{R}$  can be thought of as an  $L$ -tensor whose  $p$ -th index ranges from  $1, \dots, N_p$ . Again, the objective function of such a MMOT problem can be rephrased in terms of the 2-marginals:

$$\min_{\mu \in \mathcal{P}(\mathcal{X})} \sum_{p \neq q}^L \text{Tr}(C_{pq} \mu_{pq}^{(2)}), \quad \text{s.t. } (\text{Pr}_p) \# \mu = \mu_p^{(1)}, \quad p = 1, \dots, L, \quad (5.2)$$

where the 2-marginals  $\mu_{pq}^{(2)}$  are here implicitly defined in terms of the optimization variable  $\mu$ .



Then we introduce the *2-marginal primal SDP*

$$\begin{aligned}
& \underset{M \in \mathbb{R}^{N_{\text{tot}} \times N_{\text{tot}}}}{\text{minimize}} && \text{Tr}(CM) && (5.3) \\
& \text{subject to} && M \succeq 0, \\
& && M_{pq} \geq 0 \text{ for all } p, q = 1, \dots, L \text{ (} p \neq q \text{)}, \\
& && M_{pq} \mathbf{1}_{N_q} = \mu_p^{(1)} \text{ for all } p, q = 1, \dots, L \text{ (} p \neq q \text{)}, \\
& && M_{pp} = \text{diag}(\mu_p^{(1)}) \text{ for all } p = 1, \dots, L.
\end{aligned}$$

Here  $N_{\text{tot}} := \sum_{p=1}^L N_p$  and  $\mathbf{1}_k$  denotes the vector of ones of length  $k$ . The dual of (5.3) is given by

$$\begin{aligned}
& \underset{Y, \{\phi_{pq}, \psi_{pq}\}_{p < q}}{\text{maximize}} && 2 \sum_{p < q} (\phi_{pq} \cdot \mu_p^{(1)} + \psi_{pq} \cdot \mu_q^{(1)}) - \sum_{p, s} Y_{pp}(s, s) \mu_p^{(1)}(s) && (5.4) \\
& \text{subject to} && Y \succeq 0, \\
& && \phi_{pq} \mathbf{1}_{N_q}^\top + \mathbf{1}_{N_p} \psi_{pq}^\top \leq C_{pq} - Y_{pq} \text{ for } p < q.
\end{aligned}$$

In (5.4) it is understood that  $Y \in \mathbb{R}^{N_{\text{tot}} \times N_{\text{tot}}}$  and moreover  $\phi_{pq} \in \mathbb{R}^{N_p}$ ,  $\psi_{pq} \in \mathbb{R}^{N_q}$ .

By generalizing the discussion of Theorem 2, we have strong duality for the 2-marginal SDP, hence the optimal values of (5.3) and (5.4) are equal, and moreover the dual problem admits a maximizer. (The primal problem admits a maximizer trivially because the feasible set is compact.)

Finally, we turn to the *3-marginal primal SDP*

$$\begin{aligned}
& \underset{M \in \mathbb{R}^{N_{\text{tot}} \times N_{\text{tot}}}, K \in \mathbb{R}^{N_{\text{tot}} \times N_{\text{tot}} \times N_{\text{tot}}}}{\text{minimize}} && \text{Tr}(CM) && (5.5) \\
& \text{subject to} && M \succeq 0, \\
& && M_{pq} \geq 0 \text{ for } p \neq q, \\
& && M_{pq} \mathbf{1}_{N_q} = \mu_p^{(1)} \text{ for } p \neq q, \\
& && M_{pp} = \text{diag}(\mu_p^{(1)}) \text{ for all } p, \\
& && K \geq 0, K \text{ symmetric,} \\
& && M_{pq}(s_p, s_q) = \sum_{s_r} K_{pqr}(s_p, s_q, s_r) \text{ for } p, q, r \text{ distinct.}
\end{aligned}$$

For simplicity we omit the concrete formulation of the corresponding dual problem, but we note that strong duality and dual attainment can be proved by methods similar to those applied in the 2-marginal case.

## 6 Numerical results

In this section, we numerically demonstrate the effectiveness of the proposed methods on model problems of strongly correlated fermionic systems.

## 6.1 One-dimensional spinless model

Here we consider a 1D spinless Hubbard-like model defined by the Hamiltonian of Eq. (2.6), in which we take

$$t_{pq} = \begin{cases} 1 & \text{if } |q - p| = 1, \\ 0 & \text{otherwise} \end{cases} \quad (6.1)$$

and consider two different cases of  $v$ , with next-nearest neighbor (NN) interaction,

$$v_{pq} = \begin{cases} U/2 & \text{if } |q - p| = 1, \\ U/40 & \text{if } |q - p| = 2, \\ 0 & \text{otherwise} \end{cases} \quad (6.2)$$

and next-next-nearest neighbor interaction (NNNN)

$$v_{pq} = \begin{cases} U/2 & \text{if } |q - p| = 1, \\ U/20 & \text{if } |q - p| = 2, \\ U/200 & \text{if } |q - p| = 3, \\ 0 & \text{otherwise.} \end{cases} \quad (6.3)$$

The reason why we omit the obvious scenario of the nearest neighbor (NN) interaction is that in such a case, we find that our convex relaxation becomes *numerically exact* and hence we consider the case to be not representative. We do not have a proof yet to explain why our convex relaxation scheme can be numerically exact.

We will compare the Kohn-Sham SCE energies yielded by our methods with one another, as well as with the exact ground-state energy (2.7), which is computed via exact diagonalization (ED) in the `OpenFermion` [72] software package. The MMOT problems arising in Kohn-Sham SCE and their SDP relaxations are solved in `MATLAB` with the `CVX` software package [39].

We refer to the exact self-consistent Kohn-Sham SCE solution obtained by solving the original linear programming (LP) problem for MMOT as the ‘LP’ solution. Hence the tightness of the Kohn-Sham SCE lower bound (2.15) *itself* can be evaluated by comparing the exact energy with the LP energy, while the tightness of our SDP *relaxations* of the relevant MMOT problems (which, in turn, yield lower bounds for the Kohn-Sham SCE energy) can be evaluated by comparing the LP energy with the 2- and 3-marginal SDP energies. We refer to these two sources of error, respectively, as the ‘Kohn-Sham SCE model error’ and the ‘error due to relaxation.’

In Figs. 33(a) and 34(a), we plot  $E/U$  with respect to  $U$  for  $v$  as in Eqs. (6.2) and (6.3), respectively. In these experiments,  $L = 14$  and  $N = 9$ . The energy differences of the Kohn-Sham SCE solutions from the exact energy are plotted in Figs. 33(b) and 34(b). It is confirmed numerically that the LP energy lower-bounds the exact energy,

and in turn the SDP energies lower-bound the LP energy. While the 3-marginal SDP lower bound is noticeably tighter than the 2-marginal SDP lower bound, the error due to relaxation is dominated by the Kohn-Sham SCE model error in both cases.

Since the effective potential is of interest in Kohn-Sham DFT, in Fig. 35 we plot the SCE potential (2.16) at self-consistency in the case of  $v$  as in Eq. (6.3). It can be seen that the 3-marginal SDP performs better than the 2-marginal SDP in this regard, as one might expect. (However, note carefully that although it is guaranteed *a priori* that the 3-marginal SDP provides a lower bound on the *energy* that is at least as tight as that of the 2-marginal SDP, no such comparison is theoretically guaranteed in advance for the effective potential.)

To study the scaling of energy in the thermodynamic limit  $L \rightarrow \infty$ , in Fig 36(a), we plot  $E/U$  as a function of  $L$  by fixing  $U = 5$  and a filling factor of  $N/L = 2/3$ . In Fig 36(b), we plot the total runtime of our methods on a MacBook Pro with a 2.3GHz Core I5 CPU and 16GB of memory.

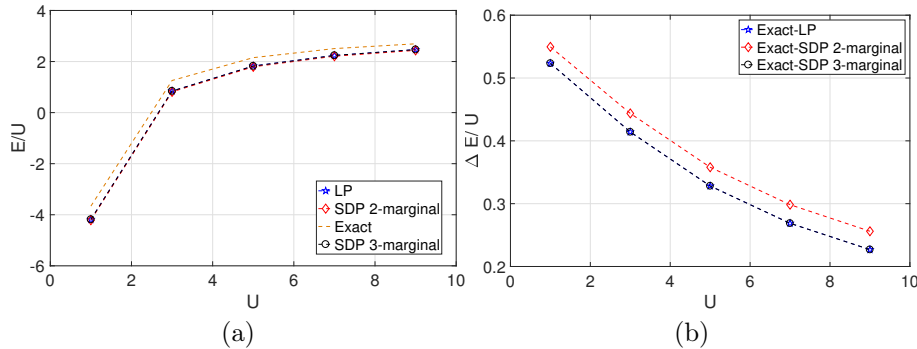


Figure 33: Spinless 1D fermionic lattice model with  $v$  as in Eq. (6.2),  $L = 14$ ,  $N = 9$ . (a)  $E/U$  as a function of  $U$ . (b) Difference between the exact energy and the Kohn-Sham SCE energies obtained from the unrelaxed LP and the SDP relaxations.

## 6.2 Two-dimensional spinful model

We consider a 2D generalized Hubbard type model defined by the Hamiltonian

$$\begin{aligned} \hat{H} = & - \sum_{i,j=1}^{L-1} \sum_{\sigma \in \{\uparrow, \downarrow\}} \left( \hat{a}_{i+1,j;\sigma}^\dagger \hat{a}_{i,j;\sigma} + \hat{a}_{i,j+1;\sigma}^\dagger \hat{a}_{i,j;\sigma} + \text{h.c.} \right) \\ & + U \sum_{i,j=1}^L \hat{n}_{i,j;\uparrow} \hat{n}_{i,j;\downarrow} + V \sum_{i,j=1}^{L-1} (\hat{n}_{i+1,j} \hat{n}_{i,j} + \hat{n}_{i,j+1} \hat{n}_{i,j}). \end{aligned} \quad (6.4)$$

Here  $\hat{n}_{i,j} := \hat{n}_{i,j;\uparrow} + \hat{n}_{i,j;\downarrow}$ . As discussed in section 2, although the creation and annihilation operators in Eq. (6.4) involve two spatial indices and one spin index, one

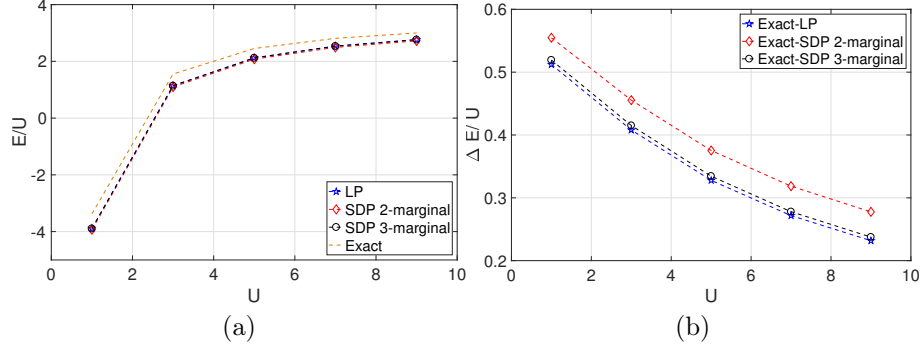


Figure 34: Spinless 1D fermionic lattice model with  $v$  as in Eq. (6.3),  $L = 14$ ,  $N = 9$ . (a)  $E/U$  as a function of  $U$ . (b) Difference between the exact energy and the Kohn-Sham SCE energies obtained from the unrelaxed LP and the SDP relaxations.

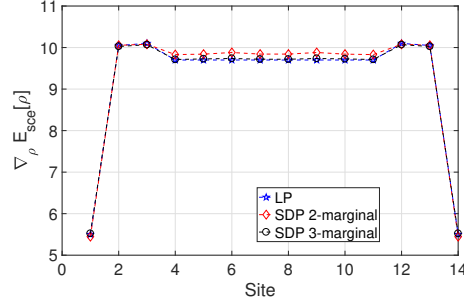


Figure 35: The effective potential for the spinless 1D fermionic lattice model with  $v$  as in Eq. (6.3),  $U = 5$ ,  $L = 14$ ,  $N = 9$ . The relative  $\ell^2$  errors for the 2- and 3-marginal formulations (compared to the unrelaxed LP formulation) are  $1.2 \times 10^{-2}$  and  $2.7 \times 10^{-3}$ , respectively.

may of course order the operators with a single index by defining

$$b_{(j-1)L+i} = a_{i,j;\uparrow}, \quad b_{(j-1)L+i+L^2} = a_{i,j;\downarrow}.$$

The new creation operators are fixed as the Hermitian adjoints of these new annihilation operators. The term associated with  $U$  is the on-site electron-electron interaction, while  $V$  specifies the nearest-neighbor electron-electron interaction. In the standard Hubbard model, we have  $V = 0$ . (However, in the case  $V = 0$ , the MMOT problem arising in the SCE framework becomes a trivial problem, since the interaction terms associated with different sites are decoupled.) Fig. 37 shows the energies for the generalized Hubbard model on a  $3 \times 3$  lattice, with  $V = 0.05 U$  and  $U$  ranging from 1.0 to 19.0. The number  $N$  of electrons is set to be 12. Here energies are obtained from the exact solution, the exact Kohn-Sham SCE solution obtained by linear programming

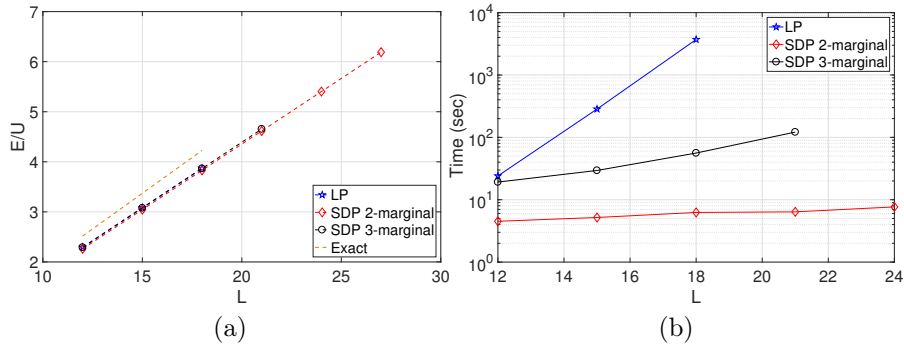


Figure 36: Spinless 1D fermionic lattice model with  $v$  as in Eq. (6.3),  $U = 5$ ,  $N/L = 2/3$ . (a)  $E/U$  as a function of  $L$ . (b) Running time as a function of  $L$ .

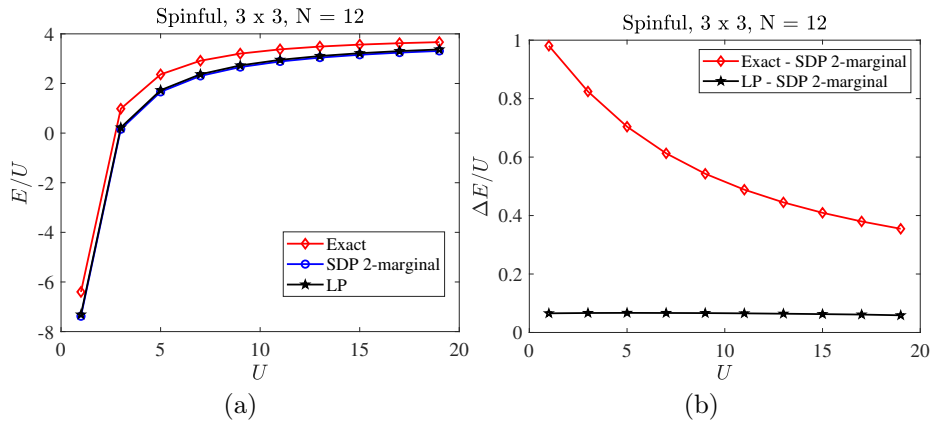


Figure 37: Spinful  $3 \times 3$  Hubbard model with  $N = 12$ .

(LP), and the approximate Kohn-Sham SCE solution obtained via the 2-marginal SDP relaxation. We find that the Kohn-Sham SCE formulation becomes asymptotically accurate when  $U$  becomes large. Furthermore, the error due to relaxation is much smaller than the Kohn-Sham SCE model error. Fig. 37(b) further shows that the energy difference between the LP and 2-marginal SDP solutions is approximately constant with respect to the on-site interaction strength  $U$ .

## 7 Conclusion

In this Part, we have considered the strictly correlated electron (SCE) limit of a fermionic quantum many-body system in the second-quantized formalism. To the extent of our knowledge, the setup of the SCE problem in this setting has not ap-

peared in the literature. Mathematically, the SCE limit requires the solution of a multi-marginal optimal transport problem over certain classical probability measures. We propose a relaxation that enforces constraints on the 2-marginals of these measures, and the relaxed problem can be solved efficiently via semi-definite programming (SDP). We prove that the SDP problem satisfies strong duality and moreover that the dual solution is attained, despite the fact that the primal problem does not possess a strictly feasible point. We consider a tighter relaxation involving the 3-marginals and discuss how our methods can be applied to completely general multi-marginal optimal transport problems with pairwise costs.

The relaxed formulation is not exact and provides only a lower bound to the SCE energy. Hence it is meaningful to compare the error due to relaxation with the Kohn-Sham SCE model error, i.e., the disparity between the Kohn-Sham SCE energy and the exact energy of the solution to the quantum many-body problem. Our numerical results for various fermionic lattice model problems indicate that the former can be much smaller than the latter, hence our convex relaxation scheme can be considered to be effective. On the other hand, as indicated in, e.g., [67], Kohn-Sham SCE is only the zero-th order approximation to the quantum many-body ground state energy in the limit of large interaction. Hence the SCE functional and SCE potential should be considered more properly as an “ingredient” for designing more accurate exchange-correlation functionals. From such a perspective, just as the exact formulation of SCE is only a model, it may even be appropriate to consider the relaxed SCE formulation as a model itself. It can capture certain strong correlation effects and can be solved efficiently.

One immediate extension of the current work is to include finite-temperature effects via entropic regularization. In fact, entropic regularization may be relevant for another reason as well. During our numerical studies, we observed that the self-consistent iteration for Kohn-Sham SCE (*not* the convex optimization problem solved within each iteration) can be difficult to converge. The convergence behavior may depend sensitively on the filling factor, the lattice size, and the form of the interaction. Such difficulty can arise for both the exact SCE formulation solved via linear programming and the relaxed formulations solved by SDP. Preliminary results show that entropic regularization can help make the loop easier to converge. We are not aware of any reports of such issues in the literature, and we plan to study such behavior more systematically in future work.

## Part IX

# Variational embedding for quantum spins and fermions

## 1 Introduction

In this section we strengthen and broaden the convex relaxation considered earlier in Part VIII. Recall that in Part VIII, we actually considered a sequence of two relaxations: the relaxation from the exact problem to exact SCE and the relaxation of exact SCE to approximate SCE (via relaxation of MMOT). We referred to the errors accrued in these two steps as the model error and the relaxation error, respectively.

In this Part we consider a relaxation of the quantum Gibbs variational principle (for quantum spin systems in addition to fermionic systems) which takes inspiration from the relaxation of MMOT introduced in Part VIII but acts directly on the exact variational principle, i.e., there is no analogous ‘model error.’ (Moreover, there is no analogous restriction to ‘generalized Coulomb interactions.’) As such these relaxations can be viewed as unrelated to the SCE formalism, though they may be motivated by our *method* for SCE.

In addition to providing variational lower bounds for the ground state energies of quantum ensembles, these relaxations can be viewed as *embedding methods* in the spirit of, e.g., the density matrix embedding theory [49] and the dynamical mean-field theory.<sup>19</sup> As far as we know, however, our relaxations are the only *variational* quantum embedding methods.

### 1.1 Outline

Please note that this Part is based on joint work in preparation with Lin Lin. In section 2, we introduce the variational embedding method for quantum spin systems. In this section 3, we introduce the analogous method for fermionic systems. In order to consider this analog, we must generalize our perspective from section 2 to the setting of star-algebras. In section 4 we show preliminary numerical results for these methods. We close in section 5 with a discussion of convex duality, which clarifies the embedding perspective on our methods.

---

<sup>19</sup>See Part VII for detailed discussion of DMFT.

## 2 Quantum spins

### 2.1 Preliminaries

Let  $i = 1, \dots, M$  index the sites of our model, and for each site  $i$  let  $X_i$  be the classical state space (discrete, for simplicity). For each site, the quantum state space is  $Q_i := \mathbb{C}^{X_i}$ , and the global quantum state space is

$$\mathcal{Q} := \bigotimes_{i=1}^M Q_i \simeq \mathbb{C}^{\mathcal{X}},$$

where  $\mathcal{X} := \prod_{i=1}^M X_i$ . Let  $H_i$  denote a Hermitian operator  $Q_i \rightarrow Q_i$ , and let  $H_{ij}$  denote a Hermitian operator  $Q_i \otimes Q_j \rightarrow Q_i \otimes Q_j$ . We will use the hatted notation  $\hat{H}_i$  to denote the operator  $\mathcal{Q} \rightarrow \mathcal{Q}$  obtained by tensoring  $H_i$  by the identity operator on all sites  $k \neq i$ , and likewise we identify  $\hat{H}_{ij}$  with the operator  $\mathcal{Q} \rightarrow \mathcal{Q}$  obtained by tensoring  $H_{ij}$  with the identity on all sites  $k \notin \{i, j\}$ . Then we consider a Hamiltonian  $\hat{H} : \mathcal{Q} \rightarrow \mathcal{Q}$  of the form

$$\hat{H} = \sum_i \hat{H}_i + \sum_{i < j} \hat{H}_{ij}.$$

*Remark 1.* We shall introduce several examples of interest in the case  $X_i = \{-1, 1\}$ , i.e., the case of quantum spin- $\frac{1}{2}$  systems. To this end, first recall the Pauli matrices:

$$\sigma^x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma^y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma^z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

which, together with the identity  $I_2$ , form a basis for Hermitian operators on  $\mathbb{C}^2$ . Now let  $\sigma_i^{x/y/z} \in \mathcal{H}(\bigotimes_i \mathbb{C}^2) \simeq \bigotimes_i \mathcal{H}(\mathbb{C}^2)$  be obtained by tensoring a copy of  $\sigma^{x/y/z}$  for the  $i$ -th site with the identity  $I_2$  on all the other sites. Given a graph structure on the site indices (with adjacency indicated by ‘ $\sim$ ’), we define two model Hamiltonians of interest—the transverse-field Ising (TFI) Hamiltonian and anti-ferromagnetic Heisenberg (AFH) Hamiltonian—as follows:

$$\hat{H}_{\text{TFI}} = -h \sum_i \sigma_i^x - \sum_{i \sim j} \sigma_i^z \sigma_j^z, \tag{2.1}$$

$$\hat{H}_{\text{AFH}} = \sum_{i \sim j} [\sigma_i^x \sigma_j^x + \sigma_i^y \sigma_j^y + \sigma_i^z \sigma_j^z]. \tag{2.2}$$

In the TFI Hamiltonian,  $h \in \mathbb{R}$  is a scalar parameter.

We are interested in the computing the ground-state energy

$$E_0 = \inf \left\{ \langle \Phi | \hat{H} | \Phi \rangle : |\Phi\rangle \in \mathcal{Q}, \langle \Phi | \Phi \rangle = 1 \right\},$$

which can be equivalently recast as



$$E_0 = \inf_{\rho \in \mathcal{D}(\mathcal{Q})} \text{Tr}[\hat{H}\rho],$$

where  $\mathcal{D}(\mathcal{Q})$  denotes the set of density operators on  $\mathcal{Q}$  (i.e., positive semidefinite linear operators  $\mathcal{Q} \rightarrow \mathcal{Q}$  of unit trace). Assuming that there exists a unique ground state  $|\Phi_0\rangle$ , the infimum is attained at  $\rho = |\Phi_0\rangle\langle\Phi_0|$ . Now we can write

$$E_0 = \inf_{\{\rho_{ij}\}_{i<j} \in \mathbf{QM}_2(\mathcal{Q})} \left( \sum_i \text{Tr}[H_i \rho_i] + \sum_{i<j} \text{Tr}[H_{ij} \rho_{ij}] \right), \quad (2.3)$$

where  $\mathbf{QM}_2(\mathcal{Q})$  denotes the set of collections  $\{\rho_{ij}\}_{i<j}$  of representable quantum two-marginals, i.e., those collections  $\{\rho_{ij}\}$  which can be obtained as reduced density operators of a single  $\rho \in \mathcal{D}(\mathcal{Q})$  via the partial trace, as in

$$\rho_{ij} = \text{Tr}_{\{1, \dots, M\} \setminus \{i, j\}}[\rho],$$

where  $i < j$ .

To clarify, here we view  $\rho$  as being equipped with labels  $1, \dots, M$  for its indices as  $\rho = \rho_{i_1 \dots i_M, j_1 \dots j_M}$ , and for any subset  $S \subset \{1, \dots, M\}$ ,  $\rho_S = \text{Tr}_{\{1, \dots, M\} \setminus S}[\rho]$  denotes the reduced density operator obtained by tracing out the indices contained in  $S$ , with the remaining labels maintained. We comment that the partial trace  $\rho_S$  may be equivalently defined as the unique operator on  $\bigotimes_{i \in S} Q_i$  such that  $\text{Tr}[\hat{A}\rho_S] = \text{Tr}[\hat{A}\rho]$  for all operators  $\hat{A}$  on  $\bigotimes_{i \in S} Q_i$  (alternatively viewed as operators on  $\mathcal{Q}$  by tensoring with the identity). This perspective illustrates the relationship between marginalization in the quantum spin setting (i.e., computing the partial trace) and the more abstract notion of marginalization that is necessary for the treatment of fermions in section 3 below.

For convenience, we denote  $\rho_{ij} = \rho_{\{i, j\}}$  for  $i < j$  as above. It is convenient to then define  $\rho_{ij}$  for  $i > j$  via the stipulation that  $\sigma_{ij}\rho_{ij}\sigma_{ji} = \rho_{ji}$ , where  $\sigma_{ij} : Q_i \otimes Q_j \rightarrow Q_j \otimes Q_i$  is the linear operator defined by  $\sigma_{ij}(\phi_i \otimes \phi_j) = \phi_j \otimes \phi_i$ . Finally, we remark that the one-marginals  $\rho_i = \text{Tr}_{\{1, \dots, M\} \setminus \{i\}}[\rho]$  are determined by the two-marginals via  $\rho_i = \text{Tr}_{\{j\}}[\rho_{ij}]$ , and this dependence is meant to be understood implicitly in (2.3). We will occasionally denote  $\rho_{ii} := \rho_i$ .

## 2.2 Local consistency constraints

Now it is of interest to determine necessary conditions satisfied by collections in  $\mathbf{QM}_2(\mathcal{Q})$ . By enforcing a set of necessary conditions as a proxy for membership in  $\mathbf{QM}_2(\mathcal{Q})$ , we can obtain a lower bound on the ground state energy.

To begin with, the  $\rho_{ij}$  are themselves density operators on  $Q_i \otimes Q_j$ , i.e.,  $\rho_{ij} \succeq 0$  with  $\text{Tr}[\rho_{ij}] = 1$ . Moreover, we must have  $\text{Tr}_j[\rho_{ij}] = \text{Tr}_{j'}[\rho_{ij'}]$  for all  $i$  and  $j, j' \neq i$ , and we must have  $\sigma_{ij}\rho_{ij}\sigma_{ji} = \rho_{ji}$ . These constraints define the set of *locally consistent quantum two-marginals*. Call this set  $\mathbf{LQM}_2(\mathcal{Q})$ . In practice we define auxiliary

variable  $\rho_i$  for the one-marginals, constrained to satisfy  $\rho_i = \text{Tr}_j[\rho_{ij}] = \text{Tr}_i[\rho_{ji}]$ . The constraints  $\text{Tr}[\rho_{ij}] = 1$  for all  $i, j$  can in fact be enforced by enforcing  $\text{Tr}[\rho_i] = 1$  for all  $i$ , since  $\text{Tr}[\rho_{ij}] = \text{Tr}[\text{Tr}_j[\rho_{ij}]]$ .

Note that the local consistency constraint  $\text{Tr}_j[\rho_{ij}] = \rho_i$  is equivalent to existing that  $\text{Tr}[\hat{A}\rho_{ij}] = \text{Tr}[\hat{A}\rho_i]$  for all operators  $\hat{A}$  on  $Q_i$  (considered also as operators on  $Q_i \otimes Q_j$  by tensoring with the identity). This perspective highlights the connection to the abstract local consistency constraints appearing in the discussion of fermionic systems in section 3 below.

### 2.3 Global semidefinite constraints and the two-marginal SDP

We can derive a further constraint, more global in nature, as follows. Consider operators  $\hat{O} : \mathcal{Q} \rightarrow \mathcal{Q}$  (not necessarily Hermitian) of the form  $\hat{O} = \sum_i \hat{O}_i$ , where each  $\hat{O}_i$  is a one-body operator on  $\mathcal{Q}$ , i.e., obtained by tensoring an operator  $O_i$  on  $Q_i$  with the identity. Now  $\hat{O}^\dagger \hat{O} \succeq 0$ , so

$$\text{Tr} \left[ \rho \hat{O}^\dagger \hat{O} \right] \geq 0 \quad (2.4)$$

for any  $\rho \in \mathcal{D}(\mathcal{Q})$ . We will expand the left-hand side to obtain a constraint on the quantum two-marginals, which can be phrased as a semidefinite matrix constraint. First compute

$$\begin{aligned} 0 &\leq \text{Tr} \left[ \rho \hat{O}^\dagger \hat{O} \right] \\ &= \text{Tr} \left[ \rho \sum_{ij} \hat{O}_i^\dagger \hat{O}_j \right] \\ &= \sum_i \text{Tr} \left[ \rho_i O_i^\dagger O_i \right] + \sum_{i \neq j} \text{Tr} \left[ \rho_{ij} O_i^\dagger \otimes O_j \right]. \end{aligned}$$

Now without loss of generality, we can identify  $X_i$  with  $\{1, \dots, m_i\}$  where  $m_i = |X_i|$ , hence we can think of  $O_i$  as an arbitrary complex matrix  $O_i = (O_{i,kl})_{k,l=1,\dots,m_i}$ . We will use square brackets to indicate entries of an operator as in  $[O_i]_{kl} = O_{i,kl}$ . Note that the two-marginal  $\rho_{ij}$  is an operator  $Q_i \otimes Q_j \rightarrow Q_i \otimes Q_j$ , so we denote its  $((k, p), (l, q))$  entry by  $[\rho_{ij}]_{kp,lq}$  for  $k, l = 1, \dots, m_i$  and  $p, q = 1, \dots, m_j$ . Finally, for  $i \neq j$ , observe that

$$\begin{aligned} \left[ O_i^\dagger \otimes O_j \right]_{kp,lq} &= [O_i^\dagger]_{kl} [O_j]_{pq} \\ &= O_{i,lk}^* O_{j,pq} \end{aligned}$$

Then we expand the  $i \neq j$  sum to obtain

$$\sum_{i \neq j} \text{Tr} \left[ \rho_{ij} O_i^\dagger \otimes O_j \right] = \sum_{i \neq j} \sum_{k,l=1}^{m_i} \sum_{p,q=1}^{m_j} [\rho_{ij}]_{lq,kp} \left[ O_i^\dagger \otimes O_j \right]_{kp,lq}$$

$$\begin{aligned}
&= \sum_{i \neq j} \sum_{k,l=1}^{m_i} \sum_{p,q=1}^{m_j} [\rho_{ij}]_{lq, kp} \overline{O_{i, lk}} O_{j, pq} \\
&= \sum_{i,j=1}^M \sum_{k,l=1}^{m_i} \sum_{p,q=1}^{m_j} (1 - \delta_{ij}) [\rho_{ij}]_{lq, kp} \overline{O_{i, lk}} O_{j, pq}
\end{aligned}$$

Next expand the  $i$  sum:

$$\begin{aligned}
\sum_i \text{Tr} \left[ \rho_i O_i^\dagger O_i \right] &= \sum_i \sum_{k=1}^{m_i} \sum_{q=1}^{m_i} [\rho_i]_{qk} \left[ O_i^\dagger O_i \right]_{kq} \\
&= \sum_i \sum_{k,l=1}^{m_i} \sum_{q=1}^{m_i} [\rho_i]_{qk} [\hat{O}_i^\dagger]_{kl} [\hat{O}_i]_{lq} \\
&= \sum_i \sum_{k,l=1}^{m_i} \sum_{q=1}^{m_i} [\rho_i]_{qk} \overline{O_{i, lk}} O_{i, lq} \\
&= \sum_i \sum_{k,l=1}^{m_i} \sum_{p,q=1}^{m_i} \delta_{lp} [\rho_i]_{qk} \overline{O_{i, lk}} O_{i, pq} \\
&= \sum_{i,j=1}^M \sum_{k,l=1}^{m_i} \sum_{p,q=1}^{m_j} \delta_{ij} \delta_{lp} [\rho_i]_{qk} \overline{O_{i, lk}} O_{i, pq}.
\end{aligned}$$

Therefore we have derived

$$\sum_{i,j=1}^M \sum_{k,l=1}^{m_i} \sum_{p,q=1}^{m_j} [\delta_{ij} \delta_{lp} [\rho_i]_{qk} + (1 - \delta_{ij}) [\rho_{ij}]_{lq, kp}] \overline{O_{i, lk}} O_{j, pq} \geq 0.$$

We can think of  $O_{j, pq}$  as a vector  $O \in \prod_{i=1}^M \mathbb{C}^{m_i \times m_i} \simeq \mathbb{C}^{\sum_{i=1}^M m_i^2}$ . The choice of such  $O$  was completely arbitrary. Therefore we have proved that the  $\left( \sum_{i=1}^M m_i^2 \right) \times \left( \sum_{i=1}^M m_i^2 \right)$  matrix  $G^{(2)} = G^{(2)}[\{\rho_{ij}\}_{i \leq j}]$  defined by

$$G_{ilk, jpq}^{(2)} := \delta_{ij} \delta_{lp} [\rho_i]_{qk} + (1 - \delta_{ij}) [\rho_{ij}]_{lq, kp}$$

is positive definite. This matrix can be thought of as a linear operator  $G^{(2)} : \prod_{i=1}^M \mathbb{C}^{m_i \times m_i} \rightarrow \prod_{i=1}^M \mathbb{C}^{m_i \times m_i}$ . (One can readily check that  $G^{(2)}$  is Hermitian.) For a quantum spin system, we have  $m_i = 2$  for all  $i$ , so this is a semidefinite constraint on a  $(4M) \times (4M)$  matrix, which is (relatively) efficient to enforce.

At last we have derived a semidefinite relaxation, which we shall call the *two-marginal SDP*:

$$E_0^{(2)} = \inf_{\{\rho_{ij}\}_{i < j} \in \text{LQM}_2(\mathcal{Q}) : G^{(2)}[\{\rho_{ij}\}_{i \leq j}] \geq 0} \left( \sum_i \text{Tr} [H_i \rho_i] + \sum_{i < j} \text{Tr} [H_{ij} \rho_{ij}] \right).$$

The relaxation yields the energy lower bound  $E_0 \geq E_0^{(2)}$ , as well as a minimizer  $\rho^{(2)}$  that is expected to approximate the exact two-marginals.

The two-marginal SDP can be written, in expanded form, as

$$\underset{\{\rho_i\}, \{\rho_{ij}\}_{i < j}}{\text{minimize}} \quad \sum_i \text{Tr} [H_i \rho_i] + \sum_{i < j} \text{Tr} [H_{ij} \rho_{ij}] \quad (2.5)$$

$$\text{subject to} \quad \rho_{ij} \succeq 0, \quad i, j = 1, \dots, M, \quad (2.6)$$

$$\rho_i = \text{Tr}_{\{j\}}[\rho_{ij}], \quad \rho_j = \text{Tr}_{\{i\}}[\rho_{ij}], \quad i, j = 1, \dots, M, \quad (2.7)$$

$$\text{Tr}[\rho_i] = 1, \quad i = 1, \dots, M, \quad (2.8)$$

$$G[\{\rho_{ij}\}_{i \leq j}] \succeq 0. \quad (2.9)$$

Although there are several ways to write constraints yielding the same feasible set, the dual SDP is actually influenced by the choice of constraints used to define this set. The choices made here will yield interesting dual structure, to be explored below.

## 2.4 Abstract perspective on the global semidefinite constraints

More abstractly, it is useful to think of  $G = G[\{\rho_{ij}\}]$  as being composed of blocks  $G_{ij}[\rho_{ij}]$  (indexed by marginal pairs  $i, j$ ), defined by

$$(G_{ij}[\rho_{ij}])_{\alpha\beta} = \begin{cases} \text{Tr} \left[ \rho_i O_{i,\alpha}^\dagger O_{i,\beta} \right] & i = j \\ \text{Tr} \left[ \rho_{ij} O_{i,\alpha}^\dagger \otimes O_{j,\beta} \right] & i \neq j, \end{cases}$$

where  $\{O_{i,\alpha}\}_\alpha$  is basis for the set of one-body operators on site  $i$ . (Note that this collection is of cardinality  $m_i^2$ .) By considering  $\alpha$  as a multi-index  $\alpha = (k, l)$  and choosing  $(O_{i,(k,l)})_{k',l'} = \delta_{kk'} \delta_{ll'}$  to be the ‘standard unit vectors’ in  $\mathbb{C}^{m_i \times m_i}$ , we exactly recover our former explicit representation of  $G[\{\rho_{ij}\}]$ .

*Remark 2.* (Restricted operator sets.) The more abstract perspective suggests a natural framework for further relaxation. Suppose that for each  $i = 1, \dots, M$ , we are given a linearly independent collection  $\{O_{i,\alpha}\}_{\alpha \in \mathcal{I}_i}$  of one-body operators for the  $i$ -th site, where  $\mathcal{I}_i$  is a given index set. Then we can define  $G = G[\rho^{(2)}]$  in terms of blocks as above, where the block  $G_{ij}[\rho_{ij}]$  is a  $|\mathcal{I}_i| \times |\mathcal{I}_j|$  matrix, defined once again by

$$(G_{ij}[\rho_{ij}])_{\alpha\beta} = \begin{cases} \text{Tr} \left[ \rho_i O_{i,\alpha}^\dagger O_{i,\beta} \right] & i = j \\ \text{Tr} \left[ \rho_{ij} \left( O_{i,\alpha}^\dagger \otimes O_{j,\beta} \right) \right] & i \neq j \end{cases}$$

for  $\alpha \in \mathcal{I}_i$ ,  $\beta \in \mathcal{I}_j$ . In principle one can consider restricted index sets with  $|\mathcal{I}_i| < m_i^2$  containing only the most physically important operators. Such restricted structure will correspond to interesting structure from the perspective of the dual problem to be considered below.

*Remark 3.* (Quasi-local constraints.) In order to improve the efficiency of the semidefinite introduced above, one could enforce the semidefiniteness of certain principal submatrices of  $G$ . E.g., for each  $k$ , one could define a submatrix  $G^{(k)}$  of  $G$  by restricting the block indices  $i, j$  to those satisfying  $d(i, k), d(j, k) \leq d_{\max}$ , where  $d$  is an appropriate notion of distance between indices (e.g., graph distance for a lattice model) and  $d_{\max}$  is a locality parameter. Then one enforces  $G^{(k)}[\{\rho_{ij}\}] \succeq 0$  for all  $k$ . For constant  $d_{\max}$  suitably large, in principle such constraints could achieve good performance while maintaining linear scaling in  $M$  of the SDP problem size for suitably local Hamiltonians, by omitting  $\rho_{ij}$  from the optimization variables for  $d(i, j) > d_{\max}$ .

## 2.5 Higher constraints

A tighter SDP relaxation can be derived by considering a set  $\{\rho_{ijk}\}_{i < j < k}$  of quantum three-marginals as the optimization variable. One may enforce the suitably defined local consistency constraints, denoted  $\{\rho_{ijk}\}_{i < j < k} \in \mathbf{LQM}_3(\mathcal{Q})$ , then defining variables  $\rho_{ij}$  in terms of the  $\rho_{ijk}$  via partial traces, additionally enforce  $G[\{\rho_{ij}\}_{i \leq j}] \succeq 0$ . We refer to the corresponding semidefinite relaxation as the *three-marginal SDP*.

To derive further semidefinite constraints, we have to keep track of the four-marginals. Suitable necessary conditions can be derived by enforcing  $\text{Tr} \left[ \rho \hat{O}^\dagger \hat{O} \right] \geq 0$  for all  $\hat{O}$  of the form  $\hat{O} = \sum_{i, i'} \hat{O}_{i, i'}$ , where the  $\hat{O}_{i, i'}$  are *two-body* operators. As such one may define the *four-marginal SDP*, and so on. Note that, e.g., the four-marginal SDP can in fact accommodate more general Hamiltonians, i.e., Hamiltonians including additional four-body terms.

## 2.6 Cluster perspective

In order to systematically improve the accuracy of the two-marginal SDP, instead of considering higher marginals we may alternately consider *increasing cluster size*. Formally, such considerations will yield problems that can still be accommodated as special cases of our previously introduced setting. However, the difference in perspective is noteworthy, and the generalization to the case of overlapping clusters (considered in the next section) is *not* accommodated as such a special case.

Suppose that our site index set is written as a union of cluster index sets  $C_\gamma$ , i.e.,

$$\{1, \dots, M\} = \bigcup_{\gamma=1}^{N_c} C_\gamma,$$

where the cluster index sets  $C_\gamma$  are *disjoint*. Then one can define

$$Y_\gamma := \prod_{i \in C_\gamma} X_i$$

to be the classical state space for the  $\gamma$ -th cluster. Then by considering the clusters now as *sites* with classical state spaces  $Y_\gamma$  and following the derivation of the two-marginal SDP, we may derive the *cluster two-marginal SDP*, relative to the cluster decomposition  $\{C_\gamma\}$ . Note that this problem may be viewed formally as a two-marginal SDP ; however, the distinction makes sense when we think of the limit of expanding clusters for a problem that is otherwise fixed. Higher-marginal cluster SDPs can be derived similarly.

## 2.7 Overlapping clusters

The treatment of overlapping clusters is more delicate. Suppose again that

$$\{1, \dots, M\} = \bigcup_{\gamma=1}^{N_c} C_\gamma,$$

but now relax the assumption that the  $C_\gamma$  are disjoint. Since the overlap of two clusters might even be a single site of the original model, we can no longer just ‘coarse-grain’ clusters and forget all of their intra-cluster structure. In particular, imposition of necessary local consistency constraints demands a bit more care.

Now the primary objects in our relaxation will be the two-*cluster* marginals, denoted  $\rho_{\gamma\delta} := \rho_{C_\gamma \cup C_\delta}$  for  $\gamma \leq \delta$ . Each  $\rho_{\gamma\delta}$  is an operator on the quantum state space specified by the *union* of sites  $C_\gamma \cup C_\delta$ , which may of course be smaller in size than  $|C_\gamma| + |C_\delta|$ . Then the one-cluster marginals  $\rho_\gamma := \rho_{C_\gamma}$  (which we sometimes also denote by  $\rho_{\gamma\gamma}$ ) are obtained in terms of the two-cluster marginals via

$$\rho_\gamma = \text{Tr}_{C_\delta \setminus C_\gamma} [\rho_{\gamma\delta}], \quad \rho_\delta = \text{Tr}_{C_\gamma \setminus C_\delta} [\rho_{\gamma\delta}].$$

These identities yield consistency constraints analogous to the local consistency constraints introduced earlier. However, we can also include the *overlap* constraints by introducing the variable  $\rho_{(\gamma\delta) \cap (\gamma'\delta')}$  representing the marginal corresponding to the set  $(C_\gamma \cup C_\delta) \cap (C_{\gamma'} \cup C_{\delta'})$ , for all  $\gamma < \delta$ ,  $\gamma' < \delta'$ , constrained by

$$\rho_{(\gamma\delta) \cap (\gamma'\delta')} = \text{Tr}_{(C_\gamma \cup C_\delta) \setminus (C_{\gamma'} \cup C_{\delta'})} [\rho_{\gamma\delta}] = \text{Tr}_{(C_{\gamma'} \cup C_{\delta'}) \setminus (C_\gamma \cap C_\delta)} [\rho_{\gamma'\delta'}].$$

Note that these constraints are nontrivial only if the intersection  $(C_\gamma \cup C_\delta) \cap (C_{\gamma'} \cup C_{\delta'})$  of cluster pairs is nonempty.

To complete the discussion of the overlapping cluster two-marginal SDP, we need to derive the semidefinite constraint. This is derived by observing the necessary condition  $\text{Tr} [\rho \hat{O}^\dagger \hat{O}] \geq 0$  for all  $\hat{O}$  of the form  $\hat{O} = \sum_\gamma \hat{O}_\gamma$ , where  $\hat{O}_\gamma$  is a one-cluster operator, i.e., an operator on  $\bigotimes_{i \in C_\gamma} Q_i$ , interpreted also (abusing notation slightly) as an operator on  $\mathcal{Q}$  by tensoring with the identity on all sites outside of  $C_\gamma$ .

In fact, given a collection of one-cluster operators  $\{O_{\gamma,\alpha}\}_{\alpha \in \mathcal{I}_\gamma}$  for the  $\gamma$ -th cluster (i.e., operators on  $\bigotimes_{i \in C_\gamma} Q_i$ ), we build  $G[\{\rho_{\gamma\delta}\}]$  blockwise by defining

$$(G_{\gamma\delta}[\rho_{\gamma\delta}])_{\alpha\beta} = \text{Tr} \left[ \rho_{\gamma\delta} \tilde{O}_{\gamma,\alpha}^\dagger \tilde{O}_{\delta,\beta} \right]$$

for  $\alpha \in \mathcal{I}_\gamma$ ,  $\beta \in \mathcal{I}_\delta$ , and  $\gamma < \delta$  (extending to  $\gamma > \delta$  by hermiticity), where  $\tilde{O}_{\gamma,\alpha}$  is an operator on  $\bigotimes_{i \in C_\gamma \cup C_\delta} Q_i$  obtained from  $O_{\gamma,\alpha}$  by tensoring with the identity operator over all sites in  $C_\delta \setminus C_\gamma$ . For example, if  $C_\gamma = \{1, 2\}$  and  $C_\delta = \{2, 3\}$ , then we can represent  $\tilde{O}_{\gamma,\alpha} = O_{\gamma,\alpha} \otimes I_{m_3}$  and  $\tilde{O}_{\delta,\beta} = I_{m_1} \otimes O_{\delta,\beta}$  (recall that here  $O_{\gamma,\alpha}$  is an operator on  $Q_1 \otimes Q_2$  and  $O_{\delta,\beta}$  is an operator on  $Q_2 \otimes Q_3$ ).

The semidefinite constraint is, as before,  $G[\{\rho_{\gamma\delta}\}] \succeq 0$ . The resulting SDP can accommodate Hamiltonians of the form

$$\hat{H} = \sum_{\gamma} \hat{H}_{\gamma} + \sum_{\gamma < \delta} \hat{H}_{\gamma\delta},$$

where  $\hat{H}_{\gamma}$  and  $\hat{H}_{\gamma\delta}$  are one-cluster and two-cluster operators, respectively.

Suitable analogous relaxations with higher overlapping cluster marginal constraints may also be derived.

## 2.8 Translation-invariant setting

In this section we describe how translation-invariant structure can be exploited in a natural way in our semidefinite relaxation framework. For simplicity we focus only on the case of the two-marginal SDP for a translation-invariant Hamiltonian in one dimension. Extension to higher dimensions is straightforward.

For the purposes of this section it is convenient to adopt a zero-indexing convention for our site indices (usually denoted by  $i, j$ ), i.e., we index our sites as  $i = 0, \dots, M-1$ . We obtain a translation-invariant Hamiltonian by assuming that  $\hat{H}_i = \hat{H}_0$  for all  $i$  and  $\hat{H}_{ij} = \hat{H}_{0,j-i}$  for all  $i < j$ . In turn we are guaranteed translation-invariance of the ground-state density operator (note: symmetry-breaking cannot occur for systems of finite size). In particular, we have  $\rho_i = \rho_0$  for all  $i$  and  $\rho_{ij} = \rho_{0,j-i}$  for all  $i < j$ , and it follows that we can constrain the matrix  $G = G[\{\rho_{ij}\}]$  to be block-circulant, so that the block  $G_{ij}$  depends only on  $i - j \pmod{M}$ . Hence all of the information of  $G$  is contained in the first row of blocks, and moreover  $G$  can be block-diagonalized by taking the blockwise discrete Fourier transform of the first row of blocks. Indeed, these diagonal blocks are obtained as

$$\tilde{G}_k = \frac{1}{\sqrt{M}} \sum_{j=0}^{M-1} \exp\left(\iota \frac{2\pi jk}{M}\right) G_{0j},$$

$k = 0, \dots, M-1$ , where we use ‘ $\iota$ ’ to denote the imaginary unit to avoid confusion with our indexing notation. Now the constraint  $G \succeq 0$  is equivalent to the constraint that  $\tilde{G}_k \succeq 0$  for all  $k$ . Hence we arrive at the periodic two-marginal SDP:

$$\underset{\rho_0, \{\rho_{0j}\}_{j=0, \dots, M-1}}{\text{minimize}} \quad \text{Tr} [H_0 \rho_0] + \sum_{j=1}^{M-1} \text{Tr} [H_{0j} \rho_{0j}]$$

$$\begin{aligned}
\text{subject to } \quad & \rho_{0j} \succeq 0, \quad j = 0, \dots, M-1, \\
& \rho_0 = \text{Tr}_{\{j\}}[\rho_{0j}], \quad \rho_0 = \text{Tr}_{\{0\}}[\rho_{0j}], \quad j = 0, \dots, M-1, \\
& \text{Tr}[\rho_0] = 1, \\
& \sum_{j=0}^{M-1} \exp\left(i\frac{2\pi jk}{M}\right) G_{0j}[\rho_{0j}] \succeq 0, \quad k = 0, \dots, M-1.
\end{aligned}$$

Notice that we have economized significantly on optimization variables, and, moreover, we have exchanged a semidefinite constraint of size  $\sim M$  for  $M$  semidefinite constraints of size constant in  $M$ . Moreover, a careful implementation of a solver for this SDP should be able to exploit the FFT in the implementation of the semidefinite constraints.

### 2.8.1 Periodicity constraints

If our sites are obtained as composite sites representing non-overlapping clusters (as discussed in section 2.6) and if, moreover, our Hamiltonian is translation-invariant with respect to these *underlying* sites, then we can impose further constraints to enforce the *internal* translation-invariance of our cluster marginals. To wit, in addition to our optimization variables  $\{\rho_{0\delta}^C\}$  for the two-*cluster* marginals, we can define additional optimization variables  $\{\rho_{0j}\}$  for the two-*site* marginals and then enforce, for all  $i \in C_0, j \in \{1, \dots, M\}$ , that  $\rho_{0,j-i} = \text{Tr}_{C_0 \cup C_{\delta(j)} \setminus \{i,j\}} \left[ \rho_{0,\delta(j)}^C \right]$ , where  $\delta(j)$  is the index of the cluster containing site  $j$ . We refer to these additional constraints as *periodicity constraints*.

## 3 Fermions

### 3.1 Preliminaries

The fundamental objects in fermionic second quantization are the creation operators  $a_1^\dagger, \dots, a_M^\dagger$  and their Hermitian adjoints, the annihilation operators  $a_i$ , which act on the Fock space  $\mathcal{F}$  and satisfy the canonical anticommutation relations

$$\{a_i, a_j^\dagger\} = \delta_{ij}, \quad \{a_i, a_j\} = \{a_i^\dagger, a_j^\dagger\} = 0,$$

where  $\{\cdot, \cdot\}$  denotes the anticommutator. One defines the number operators by  $\hat{n}_i := a_i^\dagger a_i$  and the total number operator by  $\hat{N} := \sum_{i=1}^M \hat{n}_i$ .

These objects can be concretely realized via the identification of Hilbert spaces  $\mathcal{F} \simeq \bigotimes^M \mathbb{C}^2 \simeq \mathbb{C}^{2^M}$ , under which the annihilation operators correspond to quantum spin- $\frac{1}{2}$  operators as

$$a_i^\dagger \rightsquigarrow \underbrace{\sigma^z \otimes \dots \otimes \sigma^z}_{i-1 \text{ factors}} \otimes \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \otimes I_2 \otimes \dots \otimes I_2.$$



This identification of operators defines the Jordan-Wigner transformation (JWT). Note that the JWT depends on the ordering of the states in the sense that permuting the states before the JWT is not equivalent to permuting the tensor factors after the JWT.

After specifying a particle-number-conserving Hamiltonian  $\hat{H}$ , i.e., a Hermitian operator on the Fock space which commutes with  $\hat{N}$ , and a fixed particle number  $N$ , we are interested in computing the  $N$ -particle ground state energy

$$E_0(N) = \inf \left\{ \langle \psi | \hat{H} | \psi \rangle : |\psi\rangle \in \mathcal{F}, \langle \psi | \psi \rangle = 1, \langle \psi | \hat{N} | \psi \rangle = N \right\}.$$

It is equivalent to solve

$$E_0(N) = \inf_{\rho \in \mathcal{D}(\mathcal{F}) : \text{Tr}[\hat{N}\rho] = N} \text{Tr} \left[ \hat{H} \rho \right],$$

where  $\mathcal{D}(\mathcal{F})$  indicates the set of density operators on the Fock space (i.e., positive semidefinite Hermitian operators  $\mathcal{F} \rightarrow \mathcal{F}$  of unit trace).

Observe that although  $\mathcal{F}$  can be identified with a quantum-spin state space, the creation operators are *not* one-qubit operators in the sense of quantum spin systems, nor are hopping operators  $a_i^\dagger a_j + a_j^\dagger a_i$  generically two-qubit operators. Moreover, the complexity of such operators after the JWT can depend unphysically on the ordering of the sites. Hence most second-quantized problems of interest (with the exception of local one-dimensional models) *simply do not* fit into the framework of variational embedding introduced above for quantum spin systems.

To illustrate this point and provide some concrete examples, we now describe several Hamiltonians of interest in this setting. Of particular note is the Hubbard model, whose states we enumerate via the orbital-spin index  $(i, \sigma)$ , where  $i = 1, \dots, M$ ,  $\sigma = \uparrow, \downarrow$ .

$$\hat{H} = -t \sum_{ij\sigma} A_{ij} a_{i\sigma}^\dagger a_{j\sigma} + U \sum_i \hat{n}_{i\uparrow} \hat{n}_{i\downarrow}, \quad (3.1)$$

where  $A_{ij}$  is the adjacency matrix of a graph with vertex set  $\{1, \dots, M\}$ , e.g., a square lattice.

More generally, one can consider a ‘generalized Coulomb model’ of the form

$$\hat{H} = \sum_{ij\sigma} h_{ij} a_{i\sigma}^\dagger a_{j\sigma} + \sum_{ij\sigma\tau} U_{ij} \hat{n}_{i\sigma} \hat{n}_{j\tau},$$

which includes in particular the Hubbard model and variants with longer-range interactions. In fact, via certain choices of orbital bases such as the recently introduced Gausslets [105], electronic structure problems in the continuum can be mapped to second-quantized Hamiltonians of this form. As we shall see, the generalized Coulomb model is accommodated naturally within the framework of fermionic variational embedding.

Broadening our view further still, consider a general two-body Hamiltonian  $\hat{H}$ , written as

$$\hat{H} = \sum_{ij} h_{ij} a_i^\dagger a_j + \frac{1}{2} \sum_{ijkl} v_{ijkl} a_i^\dagger a_j^\dagger a_l a_k.$$

Electronic structure problems in first quantization can be mapped to such Hamiltonians via an arbitrary choice of orbital basis  $\{\phi_i\}$  for (a subspace of)  $L^2(\mathbb{R}^d)$ , where  $d$  is the physical dimension. If the basis functions have compact support, then  $v_{ijkl}$  can be nonzero only if both  $\text{supp}(\phi_i) \cap \text{supp}(\phi_k) \neq \emptyset$  and  $\text{supp}(\phi_j) \cap \text{supp}(\phi_l) \neq \emptyset$ . It will follow that after a suitable choice of overlapping clusters (chosen so that each pair of intersecting basis functions), such Hamiltonians can also be accommodated within fermionic variational embedding. We leave investigation of *ab initio* quantum chemistry problems by these means to future work.

In order to define a convex relaxation of the fermionic Gibbs variational principle that is analogous to our relaxation for quantum spin systems, we adopt a more abstract (and indeed general) perspective in section 3.2, allowing for the derivation of a suitable two-cluster-marginal SDP in 3.3. We will in fact see in section 3.4 that our relaxation is tight for noninteracting Hamiltonians, i.e., Hamiltonians that are quadratic in the creation and annihilation operators. This feature has no analog in the quantum spin setting because there is no related notion of noninteracting systems. Then in section 3.5, we will describe how one can translate our abstract convex optimization problem into an explicit SDP that can be implemented on a computer.

## 3.2 Abstract perspective

The fundamental objects of interest in the abstract perspective is the *algebra of operators* on the Fock space. In fact, the Fock space itself plays no direct role in the following developments, nor does any global JWT. Marginalization will make use of the notion of a *subalgebra* subordinate to each cluster. It is in the details of how these subalgebras lie within the global algebra that the quantum-spin and fermionic cases differ.

Now let

$$\mathcal{A} := \langle 1, a_1, \dots, a_M, a_1^\dagger, \dots, a_M^\dagger \rangle$$

denote the unital star-algebra over the complex numbers<sup>20</sup> generated by the creation and annihilation operators subject to the canonical anticommutation relations  $\{a_i, a_j^\dagger\} = \delta_{ij}$ ,  $\{a_i, a_j\} = \{a_i^\dagger, a_j^\dagger\} = 0$ . (Throughout we will use angle brackets to denote such generated algebras.) We let  $\hat{n}_i = a_i^\dagger a_i$  denote the corresponding number operators and let  $\hat{N} = \sum_i \hat{n}_i$  denote the total number operator.

---

<sup>20</sup>Recall that a star-algebra over  $\mathbb{C}$  is essentially an associative algebra over  $\mathbb{C}$  in which one can take adjoints, where the adjoints satisfy their usual algebraic properties. We will use no deep results from the theory of star-algebras but nonetheless find the perspective to be clarifying. In specific, it is useful to view our algebra of fermionic operators independently from any Fock space on which it acts, and in fact the notion of the Fock space does not play any explicit role in our developments.

In fact the algebra  $\mathcal{A}$  comes equipped with a  $\mathbb{Z}_2$ -grading, i.e., we can write  $\mathcal{A}$  as a direct sum of vector spaces  $\mathcal{A} = \mathcal{A}^e \oplus \mathcal{A}^o$ , where  $\mathcal{A}^e$  and  $\mathcal{A}^o$  denote the sets of even and odd operators, respectively. An operator is even (resp., odd) if it can be written as a sum of even (resp., odd) monomials in  $a_1, \dots, a_M, a_1^\dagger, \dots, a_M^\dagger$ . (The reader can check that this notion is well-defined.) The  $\mathbb{Z}_2$ -grading refers to the fact that  $\mathcal{A}^e \mathcal{A}^e \subset \mathcal{A}^e$ ,  $\mathcal{A}^o \mathcal{A}^o \subset \mathcal{A}^e$ ,  $\mathcal{A}^e \mathcal{A}^o \subset \mathcal{A}^o$ , and  $\mathcal{A}^o \mathcal{A}^e \subset \mathcal{A}^o$ .

For any subset  $C \subset \{1, \dots, M\}$ . Let  $\mathcal{A}_C$  denote the subalgebra

$$\mathcal{A}_C := \left\langle \{1\} \cup \{a_i, a_i^\dagger : i \in C\} \right\rangle,$$

and let the even and odd components  $\mathcal{A}_C^e$  and  $\mathcal{A}_C^o$  be defined accordingly. Suppose that our site index set is written as a union of cluster index sets  $C_\gamma$ , i.e.,

$$\{1, \dots, M\} = \bigcup_{\gamma=1}^{N_c} C_\gamma,$$

where the cluster index sets  $C_\gamma$  are disjoint, for simplicity.

We comment that, in contrast to our exposition for the case of quantum spins, we shall directly work with general clusters (as opposed to clusters consisting of a single site). The reason is that in the quantum spin setting, it was possible to view non-overlapping clusters as single sites (with enlarged local state spaces). Such a reduction is not natural in the fermionic setting. Hence we retain the index notation  $\gamma, \delta$  for clusters and  $i, j$  for individual sites (of which the clusters are comprised).

Now our structural assumption on our Hamiltonian  $\hat{H} \in \mathcal{A}$  is that it can be written as a sum of one-cluster and two-cluster operators, i.e., as

$$\hat{H} = \sum_{\gamma} \hat{H}_\gamma + \sum_{\gamma < \delta} \hat{H}_{\gamma\delta},$$

where  $\hat{H}_\gamma \in \mathcal{A}_{C_\gamma}$  and  $\hat{H}_{\gamma\delta} \in \mathcal{A}_{C_\gamma \cup C_\delta}$ .

Note carefully for context that the subalgebra  $\mathcal{A}_{C_\gamma}$  corresponds in our earlier setting of quantum spin systems to the subalgebra of operators on  $\bigotimes_{i \in C_\gamma} Q_i$ , viewed as operators on  $\mathcal{Q}$  by tensoring with the identity. Clearly, even by viewing the fermionic system as a spin system via JWT, this subalgebra is inequivalent to the fermionic subalgebra above defined. The reader should keep this perspective on the developments of section 2 in mind as we transpose them to the fermionic setting.

Next we turn to defining our notion of a statistical ensemble and its marginals. For this task we turn to the language of star-algebras. The role of our full ensemble is played by the *state*, a linear functional  $\omega : \mathcal{A} \rightarrow \mathbb{C}$  such that  $\omega(1) = 1$  and  $\omega(A^\dagger A) \geq 0$  for any  $A \in \mathcal{A}$ . In our setting (which is finite-dimensional), the action of a state can be viewed as nothing more than tracing against a density operator on the Fock space, as can be verified readily via the Riesz representation theorem. However, the abstract perspective will be useful in defining the notion of a marginal because if we try to

directly borrow the corresponding notion from the setting of quantum spins, i.e., the partial trace, then we find ourselves in need of a global JWT to proceed.

We let  $\Omega$  denote the set of states on  $\mathcal{A}$ . Then in star-algebraic language, the  $N$ -particle ground-state energy  $E_0(N)$  minimization problem is naturally recast as

$$E_0(N) = \inf_{\omega \in \Omega : \omega(\hat{N})=N} \omega(\hat{H}). \quad (3.2)$$

Next, our notion of a marginal in this setting is simply the restriction of a state to a subalgebra. That is, for a subset  $C \subset \{1, \dots, M\}$ , we define the marginal  $\omega_C$  via

$$\omega_C := \omega|_{\mathcal{A}_C}.$$

Of course,  $\omega_C$  is itself a state on  $\mathcal{A}_C$ . We let  $\Omega_C$  denote the set of states on  $\mathcal{A}_C$ . Notice that, as follows immediately from the definition, the sets  $\Omega, \Omega_C$  are *convex*.

Note that in the quantum spin setting of section 2, the action  $\omega(\hat{A})$  of the state corresponds to the trace  $\text{Tr}[\hat{A}\rho]$  against the density operator  $\rho$ . For  $\hat{A}$  an operator on  $\bigotimes_{i \in C} Q_i$ , we have  $\omega_C(\hat{A}) = \omega(\hat{A}) = \text{Tr}[\hat{A}\rho] = \text{Tr}[\hat{A}\rho_C]$ , i.e., our notion of marginalization—applied to a cluster subalgebra in the quantum spin setting—precisely recovers the partial trace operation.

### 3.3 The two-cluster-marginal SDP

In this section we shall derive an ‘abstract SDP’ without describing how it can be realized on a computer. Later, in section 3.5, we will describe how to achieve such realization (which makes use of JWTs only for each pair of clusters). For simplicity, we will only derive a relaxation that analogizes the (*nonoverlapping*) *two-cluster-marginal SDP*. Further analogs can be derived by straightforward (though perhaps tedious) modifications of the arguments presented below.

For simplicity we denote the one-cluster marginals by  $\omega_\gamma := \omega_{C_\gamma}$  and the two-cluster marginals by  $\omega_{\gamma\delta} := \omega_{C_\gamma \cup C_\delta}$ . Note carefully from the definitions here that  $\omega_{\gamma\delta} = \omega_{\delta\gamma} : \mathcal{A}_{C_\gamma \cup C_\delta} \rightarrow \mathbb{C}$  and that  $\omega_{\gamma\gamma} = \omega_\gamma : \mathcal{A}_{C_\gamma} \rightarrow \mathbb{C}$ . Our one- and two-cluster marginals evidently satisfy the local consistency constraints

$$\omega_\gamma = \omega_{\gamma\delta}|_{\mathcal{A}_{C_\gamma}}, \quad \omega_\delta = \omega_{\gamma\delta}|_{\mathcal{A}_{C_\delta}}$$

via nested restriction.

Now, by analogy to (2.4), our semidefinite constraint will be derived from the observation that for any  $\hat{A} \in \mathcal{A}$  of the form  $\hat{A} = \sum_\gamma \hat{A}_\gamma$ , where  $\hat{A}_\gamma \in \mathcal{A}_{C_\gamma}$  for all  $\gamma$ ,

$$0 \leq \omega(A^\dagger A) = \omega \left( \left[ \sum_\gamma \hat{A}_\gamma \right]^\dagger \left[ \sum_\delta \hat{A}_\delta \right] \right) = \sum_{\gamma\delta} \omega \left( \hat{A}_\gamma^\dagger \hat{A}_\delta \right).$$

Therefore the two-cluster marginals satisfy

$$\sum_{\gamma\delta} \omega_{\gamma\delta} \left( \hat{A}_\gamma^\dagger \hat{A}_\delta \right) \geq 0$$

for all choices of  $\{A_\gamma\}_{\gamma=1}^{N_c}$  for which  $\hat{A}_\gamma \in \mathcal{A}_{C_\gamma}$  for all  $\gamma$ .

More specifically, for each cluster  $\gamma$  consider a *list*  $\left\{ \hat{A}_{\gamma,\alpha} \right\}_{\alpha \in \mathcal{I}_\gamma}$  of operators in  $\mathcal{A}_{C_\gamma}$ , possibly (but not necessarily) spanning the space of all operators in  $\mathcal{A}_{C_\gamma}$ . (Compare to the perspective of section 2.4 on the global semidefinite constraints in the quantum spin setting.) Then one obtains  $G[\{\omega_{\gamma\delta}\}] \succeq 0$ , where  $G = (G_{\gamma\delta})$  is specified blockwise by

$$(G_{\gamma\delta}[\omega_{\gamma\delta}])_{\alpha\beta} = \omega_{\gamma\delta} \left( \hat{A}_{\gamma,\alpha}^\dagger \hat{A}_{\delta,\beta} \right).$$

In fact,  $G = G[\{\omega_{\gamma\delta}\}_{\gamma \leq \delta}]$  depends only on  $\omega_{\gamma\delta}$  for  $\gamma \leq \delta$  because the lower triangular part can be obtained from the upper triangular part via hermiticity.

Then we have derived the following relaxation of the variational principle (3.2), in which the  $\omega_\gamma$  and  $\omega_{\gamma\delta}$  are considered as optimization variables:

$$\begin{aligned} E_0^{(2)}(N) &:= \underset{\{\omega_\gamma\}, \{\omega_{\gamma\delta}\}_{\gamma < \delta}}{\text{minimize}} && \sum_{\gamma} \omega_\gamma \left( \hat{H}_\gamma \right) + \sum_{\gamma < \delta} \omega_{\gamma\delta} \left( \hat{H}_{\gamma\delta} \right), && (3.3) \\ &\text{subject to} && \omega_{\gamma\delta} \in \Omega_{C_\gamma \cup C_\delta}, \quad 1 \leq \gamma < \delta \leq N_c, \\ &&& \omega_\gamma = \omega_{\gamma\delta}|_{\mathcal{A}_{C_\gamma}}, \quad \omega_\delta = \omega_{\gamma\delta}|_{\mathcal{A}_{C_\delta}}, \quad 1 \leq \gamma < \delta \leq N_c, \\ &&& N = \sum_{\gamma} \omega_\gamma (\hat{N}_\gamma), \\ &&& G[\{\omega_{\gamma\delta}\}_{\gamma \leq \delta}] \succeq 0, \end{aligned}$$

where  $\hat{N}_\gamma := \sum_{i \in C_\gamma} \hat{n}_i$  denotes the  $\gamma$ -th cluster number operator. Since the constraints are convex, we have specified an abstract convex optimization problem. Now that we know that this relaxation makes sense in principle, our hope is to express it later as a concrete semidefinite program.

It is computationally useful to realize a simplification. Physical fermionic Hamiltonians are always even (including the anomalous, or particle-number-*non*conserving, Hamiltonians that arise in effective descriptions of superconductivity), and hence one expects the action of a physical state on an *odd* operator in fact always yields zero. Hence

$$(G_{\gamma\delta}[\omega_{\gamma\delta}])_{\alpha\beta} = \omega_{\gamma\delta} \left( \hat{A}_{\gamma,\alpha}^\dagger \hat{A}_{\delta,\beta} \right)$$

is zero unless  $\hat{A}_{\gamma,\alpha}$  and  $\hat{A}_{\delta,\beta}$  are either both even or both odd. It follows that we can reduce the size of the semidefinite constraint by splitting our operator lists into even and odd subsets which we denote  $\left\{ \hat{A}_{\gamma,\alpha}^e \right\}_{\alpha \in \mathcal{I}_\gamma^e}$  and  $\left\{ \hat{A}_{\gamma,\alpha}^o \right\}_{\alpha \in \mathcal{I}_\gamma^o}$ , respectively. Then we

define separate matrices  $G^e$  and  $G^o$  blockwise by

$$\left(G_{\gamma\delta}^{e/o}[\omega_{\gamma\delta}]\right)_{\alpha\beta} = \omega_{\gamma\delta} \left( \left[ \hat{A}_{\gamma,\alpha}^{e/o} \right]^\dagger \left[ \hat{A}_{\delta,\beta}^{e/o} \right] \right). \quad (3.4)$$

Then we may equivalently substitute our semidefinite constraint  $G \succeq 0$  with two semidefinite constraints  $G^{e/o} \succeq 0$ , each of half (assuming that complete operator lists are chosen) the original size.

### 3.4 Exactness for noninteracting problems

In this section we assume that  $\hat{H}$  is noninteracting, i.e., of the form  $\hat{H} = \sum_{ij} \overline{h_{ij}} a_i^\dagger a_j$ , where  $h = (h_{ij})$  is Hermitian. We want to show that in this setting  $E_0^{(2)}(N) = E_0(N)$ , i.e., the relaxation just introduced is tight, under the meager further assumption that for each  $i \in \{1, \dots, M\}$ , the operators  $a_i, a_i^\dagger$  are contained in some cluster's operator list.

Indeed, under this latter assumption it is not hard to see that the matrices  $D(\omega_{\{i,j\}}) := \left( \omega_{\{i,j\}}(a_i^\dagger a_j) \right)_{i,j=1}^M$  and  $D'(\omega_{\{i,j\}}) := \left( \omega_{\{i,j\}}(a_i a_j^\dagger) \right)_{i,j=1}^M$  appear as principal submatrices of  $G^o[\{\omega_{\gamma\delta}\}]$ , where the two-site marginals  $\omega_{\{i,j\}}$  are suitably obtained in terms of the two-cluster marginals  $\omega_{\gamma\delta}$  by appropriate restriction. Note that by the fermionic anticommutation relations, in fact  $D'(\omega_{\{i,j\}}) = I_M - D(\omega_{\{i,j\}})$ . Hence for any feasible solution to our SDP, we have  $0 \preceq D(\omega_{\{i,j\}}) \preceq I_M$ . Then it follows that  $E_0^{(2)}(N)$  is itself an upper bound for the optimal value  $E_0'(N)$  of the following (further relaxed) SDP:

$$\begin{aligned} E_0'(N) &:= \underset{D \in \mathbb{C}^{M \times M}}{\text{minimize}} && \text{Tr}[hD] \\ &\text{subject to} && 0 \preceq D \preceq I_M, \\ &&& \text{Tr}[D] = N. \end{aligned}$$

But it is not hard to show that  $E_0'(N) = \sum_{i=1}^N \lambda_i(h)$ , where  $\lambda_i(h)$  indicates the  $i$ -th lowest eigenvalue of  $h$ . But as is well known for noninteracting problems, this is precisely the value of  $E_0(N)$ . Hence we have shown  $E_0(N) \geq E_0^{(2)}(N) \geq E_0'(N) = E_0(N)$ , from which it follows that  $E_0^{(2)}(N) = E_0(N)$ .

For many problems one also expects asymptotic tightness in the limit of strong interaction. For example, in the  $t \rightarrow 0$  limit of the Hubbard model, the sites completely decouple, and we conjecture that our SDP is tight in this scenario.

### 3.5 Concrete perspective

We want to figure out how to represent  $\omega_{\gamma\delta}$  in concrete terms. Note that  $\omega_{\gamma\delta}$  is defined by its action on  $\mathcal{A}_{C_\gamma \cup C_\delta}$ . It is at this point that we introduce for computational purposes the JWT, though only for restricted fermionic algebras. After specifying

ordering the sites of  $C_\gamma \cup C_\delta$ , i.e., a labeling map  $\kappa_{\gamma\delta} : C_\gamma \cup C_\delta \rightarrow \{1, \dots, L_{\gamma\delta}\}$  where  $L_{\gamma\delta} := |C_\gamma \cup C_\delta|$ , the corresponding JWT fixes an algebra isomorphism  $\mathcal{J}_{\gamma\delta} : \mathcal{A}_{C_\gamma \cup C_\delta} \rightarrow \text{End} \left( \bigotimes_{i=1}^{L_{\gamma\delta}} \mathbb{C}^2 \right) \simeq \mathbb{C}^{2^{L_{\gamma\delta}}}$ , and we define  $c_i^{\gamma\delta} \in \text{End} \left( \bigotimes_{i=1}^{L_{\gamma\delta}} \mathbb{C}^2 \right)$  to be the image of  $a_i$  under this isomorphism for  $i \in C_\gamma \cup C_\delta$ . More specifically, the transformation  $\mathcal{J}_{\gamma\delta}$  is specified by setting  $\mathcal{J}_{\gamma\delta}(a_{\kappa_{\gamma\delta}^{-1}(i)}) = c_i^{\gamma\delta}$ , where

$$c_i^{\gamma\delta} := \underbrace{\sigma_z \otimes \cdots \otimes \sigma_z}_{(i-1) \text{ factors}} \otimes \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \otimes \underbrace{I_2 \otimes \cdots \otimes I_2}_{(L_{\gamma\delta}-i) \text{ factors}}.$$

Here  $\sigma_z$  is the usual Pauli matrix  $\sigma_z := \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ . Notice that the case  $\gamma = \delta$  makes perfect sense according to the above definitions, though we will also introduce the alternative notation  $\mathcal{J}_\gamma := \mathcal{J}_{\gamma\gamma}$ .

Moreover,  $F_{\gamma\delta} := \mathcal{J}_{\gamma\delta} \circ \omega_{\gamma\delta} \circ \mathcal{J}_{\gamma\delta}^{-1}$  is a linear functional on  $\text{End} \left( \bigotimes_{i=1}^{L_{\gamma\delta}} \mathbb{C}^2 \right)$  satisfying  $F_{\gamma\delta}(\text{Id}) = 1$  and  $F_{\gamma\delta}(A^\dagger A) \geq 0$  for any  $A \in \text{End} \left( \bigotimes_{i=1}^{L_{\gamma\delta}} \mathbb{C}^2 \right)$ . It follows (via Riesz representation) that there exists a unique  $\rho_{\gamma\delta} \succeq 0$  with  $\text{Tr}[\rho_{\gamma\delta}] = 1$  such that  $F_{\gamma\delta}(A) = \text{Tr}[A\rho_{\gamma\delta}]$  for all  $A \in \text{End} \left( \bigotimes_{i=1}^{L_{\gamma\delta}} \mathbb{C}^2 \right)$ . That is to say,  $\omega_{\gamma\delta}(\hat{A}) = \text{Tr}[A\rho_{\gamma\delta}]$  whenever  $A = \mathcal{J}_{\gamma\delta}(\hat{A})$ . Again, we introduce the alternative notation  $\rho_\gamma = \rho_{\gamma\gamma}$  for conceptual clarity.

Motivated by the preceding, we shall replace our optimization over *states*  $\omega_{\gamma\delta} : \mathcal{A}_{C_\gamma \cup C_\delta} \rightarrow \mathbb{C}$  with optimization over *density operators*  $\rho_{\gamma\delta} \in \text{End} \left( \bigotimes_{i=1}^{L_{\gamma\delta}} \mathbb{C}^2 \right)$ . Crucially, the correspondence between states and density operators has relied on a separate JWT for *each* pair  $(\gamma, \delta)$ , not a single global JWT that maps the global fermionic state to a global density operator.

Under this correspondence  $G_{\gamma\delta}^{e/o}[\omega_{\gamma\delta}]$  as defined (3.4) can be obtained as

$$\left( G_{\gamma\delta}^{e/o}[\rho_{\gamma\delta}] \right)_{\alpha\beta} = \text{Tr} \left( \left[ \mathcal{J}_{\gamma\delta} \left( \hat{A}_{\gamma,\alpha}^{e/o} \right) \right]^\dagger \left[ \mathcal{J}_{\gamma\delta} \left( \hat{A}_{\delta,\beta}^{e/o} \right) \right] \rho_{\gamma\delta} \right),$$

where we abuse notation slightly by identifying  $G_{\gamma\delta}^{e/o}[\rho_{\gamma\delta}]$  with  $G_{\gamma\delta}^{e/o}[\omega_{\gamma\delta}]$ .

In order to write down a concrete realization of the optimization problem (3.3), the hurdle that remains is to encode the local consistency constraints  $\omega_\gamma = \omega_{\gamma\delta}|_{\mathcal{A}_{C_\gamma}}$  and  $\omega_\delta = \omega_{\gamma\delta}|_{\mathcal{A}_{C_\delta}}$  for  $\gamma < \delta$ , which require us to ‘marginalize’ our fermionic states.

To see how to do this, we first assume that the labeling map  $\kappa_{\gamma\delta}$  satisfies  $\kappa_{\gamma\delta}(C_\gamma) < \kappa_{\gamma\delta}(C_\delta)$  in the sense that every element of the left-hand side is less than every element of the right-hand side. (In the case of overlapping clusters, the relevant generalization ensures that  $\kappa_{\gamma\delta}(C_\gamma) < \kappa_{\gamma\delta}([C_\gamma \cup C_\delta] \setminus C_\gamma)$ .) For simplicity we also assume that  $\kappa_{\gamma\delta}|_{C_\gamma} = \kappa_{\gamma\gamma}$  for all  $\gamma < \delta$ , and from now on we think of the labeling maps  $\kappa_{\gamma\delta}$  as fixed. It is always possible to choose a labeling that satisfies these assumptions.

Then it follows from the definition of the JWT that any element  $A = \mathcal{J}_{\gamma\delta}(\hat{A})$  of  $\mathcal{J}_{\gamma\delta}(\mathcal{A}_{C_\delta})$  is of the form

$$A = B \otimes \text{Id}_{\bigotimes_{i=1}^{|C_\delta|} \mathbb{C}^2} = B \otimes \underbrace{I_2 \otimes \cdots \otimes I_2}_{|C_\delta| \text{ factors}},$$

where  $B = \mathcal{J}_\gamma(\hat{A}) \in \text{End}\left(\bigotimes_{i=1}^{L_\gamma} \mathbb{C}^2\right)$ . Then

$$\omega_{\gamma\delta}(\hat{A}) = \text{Tr}[A\rho_{\gamma\delta}] = \text{Tr}[B\tilde{\rho}_\gamma],$$

where  $\tilde{\rho}_\gamma := \text{Tr}_{\kappa_{\gamma\delta}(C_\gamma)}[\rho_{\gamma\delta}]$ . Meanwhile, we have  $\mathcal{J}_\gamma(\hat{A}) = B$ , and  $\omega_\gamma(\hat{A}) = \text{Tr}[B\rho_\gamma]$ . Hence the constraint  $\omega_\gamma = \omega_{\gamma\delta}|_{\mathcal{A}_{C_\gamma}}$  for  $\gamma < \delta$  is equivalent to the stipulation that  $\text{Tr}[B\tilde{\rho}_\gamma] = \text{Tr}[B\rho_\gamma]$  for all  $B$ , i.e., that

$$\rho_\gamma = \text{Tr}_{\kappa_{\gamma\delta}(C_\delta)}[\rho_{\gamma\delta}].$$

Meanwhile, for any  $A = \mathcal{J}_{\gamma\delta}(\hat{A})$  where  $\hat{A} \in \mathcal{A}_{C_\delta}^e$  is *even*, we can write

$$A = \underbrace{I_2 \otimes \cdots \otimes I_2}_{|C_\gamma| \text{ factors}} \otimes B,$$

where  $B = \mathcal{J}_\delta(\hat{A}) \in \text{End}\left(\bigotimes_{i=1}^{L_\delta} \mathbb{C}^2\right)$ . Hence for all  $\hat{A} \in \mathcal{A}_{C_\delta}^e$ , we derive as above that  $\omega_\delta(\hat{A}) = \text{Tr}[B\tilde{\rho}_\delta]$ , where  $\tilde{\rho}_\delta := \text{Tr}_{\kappa_{\gamma\delta}(C_\delta)}[\rho_{\gamma\delta}]$ . But for  $\hat{A} \in \mathcal{A}_{C_\delta}^o$ , as mentioned above we can assume  $\omega_{\gamma\delta}(\hat{A}) = \omega_\delta(\hat{A}) = 0$  (because this identity is a necessary condition satisfied by the exact marginals) and hence also that  $\text{Tr}[B\rho_\delta] = 0 = \text{Tr}[B\tilde{\rho}_\delta]$  for all  $B \in \mathcal{J}_\delta(\mathcal{A}_{C_\delta}^o)$ . Thus the constraint  $\omega_\delta = \omega_{\gamma\delta}|_{\mathcal{A}_{C_\delta}}$  for  $\gamma < \delta$  is equivalent to the stipulation that  $\text{Tr}[B\tilde{\rho}_\delta] = \text{Tr}[B\rho_\delta]$  for all  $B$ , i.e., that

$$\rho_\delta = \text{Tr}_{\kappa_{\gamma\delta}(C_\delta)}[\rho_{\gamma\delta}].$$

Finally, note that the constraint  $\text{Tr}[\rho_{\gamma\delta}] = 1$  can simply be encoded, given our first local consistency constraint, by  $\text{Tr}[\rho_\gamma] = 1$ . Then we obtain the following concrete realization of (3.3):

$$\begin{aligned} E_0^{(2)}(N) &:= \underset{\{\rho_\gamma\}, \{\rho_{\gamma\delta}\}_{\gamma < \delta}}{\text{minimize}} && \sum_{\gamma} \text{Tr} \left[ \mathcal{J}_\gamma \left( \hat{H}_\gamma \right) \rho_\gamma \right] + \sum_{\gamma < \delta} \text{Tr} \left[ \mathcal{J}_{\gamma\delta} \left( \hat{H}_{\gamma\delta} \right) \rho_{\gamma\delta} \right], \\ &\text{subject to} && \rho_{\gamma\delta} \succeq 0, \quad 1 \leq \gamma < \delta \leq N_c, \\ &&& \rho_\gamma = \text{Tr}_{\kappa_{\gamma\delta}(C_\delta)}[\rho_{\gamma\delta}], \quad \rho_\delta = \text{Tr}_{\kappa_{\gamma\delta}(C_\gamma)}[\rho_{\gamma\delta}], \quad \gamma < \delta, \\ &&& \text{Tr}[\rho_\gamma] = 1, \quad \gamma = 1, \dots, N_c, \\ &&& N = \sum_{\gamma} \text{Tr} \left[ \mathcal{J}_\gamma \left( \hat{N}_\gamma \right) \rho_\gamma \right], \\ &&& G[\{\rho_{\gamma\delta}\}_{\gamma \leq \delta}] \succeq 0. \end{aligned}$$



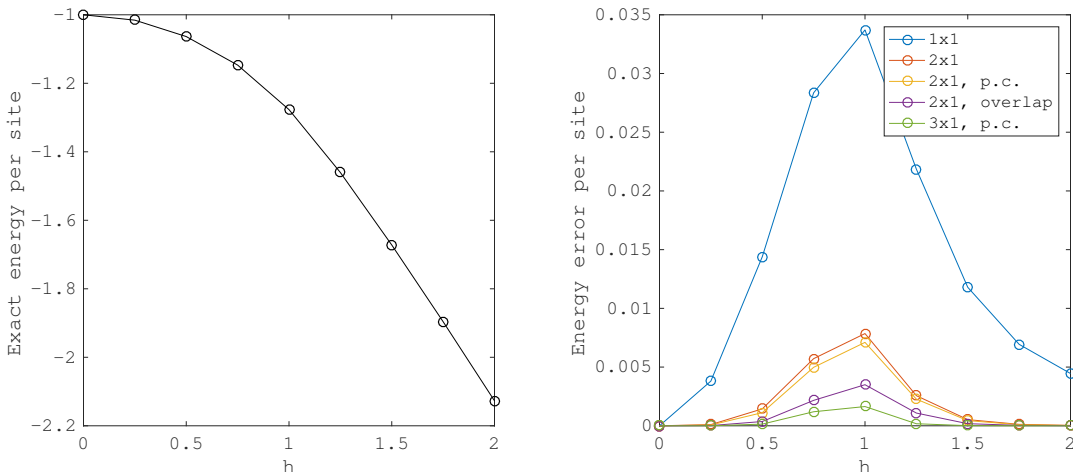


Figure 38: TFI model on periodic  $12 \times 1$  lattice. Approximate energies are computed via the two-cluster-marginal relaxation. Note that ‘p.c.’ indicates the inclusion of the periodicity constraints introduced in section 2.8.1, and ‘overlap’ indicates that the choice of overlapping  $2 \times 1$  clusters, i.e.,  $\{1, 2\}, \{2, 3\}, \{3, 4\}, \dots, \{11, 12\}, \{12, 1\}$ .

## 4 Numerical results

All numerical results were computed in MATLAB with CVX [39]. We limit our experiments to problems that are small enough to validate by exact diagonalization. In particular, we will illustrate numerically the fact that all of our relaxations must yield lower bounds for the exact energy. As discussed in section 5.3 below, a more scalable implementation should be possible, but such an implementation (as well as an accompanying numerical study of properties of larger systems, e.g., approaching a thermodynamic limit) will be left to future work.

### 4.1 Transverse-field Ising model

First we consider the transverse-field Ising (TFI) model (2.1) on a periodic  $12 \times 1$  square lattice, comparing results of the two-cluster-marginal SDP for various cluster sizes. We also test the periodicity constraints of 2.8.1 and the case of overlapping clusters. The results are shown in Figure 38.

Next we consider the TFI model on a periodic  $4 \times 3$  square lattice, comparing results of the two-cluster-marginal SDP for various cluster sizes. The results are shown in Figure 39.

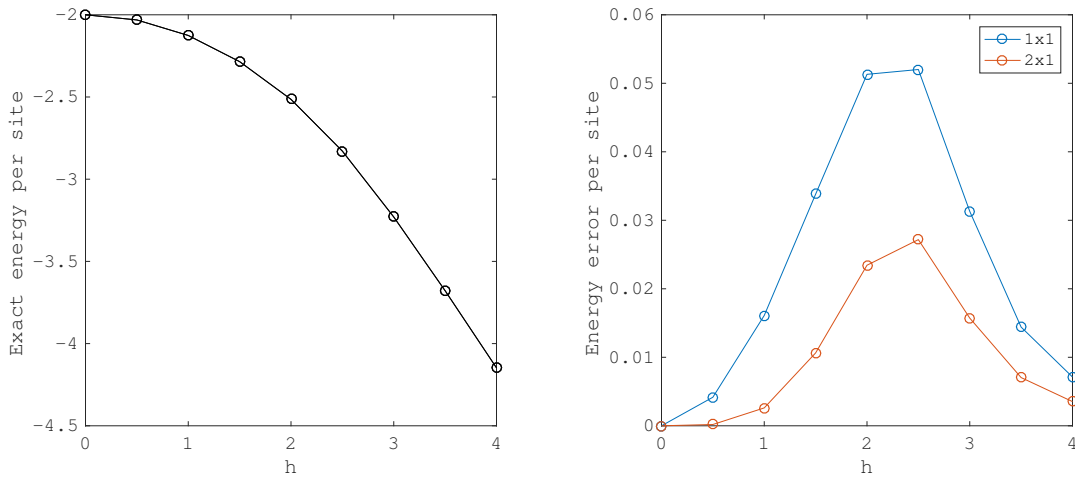


Figure 39: TFI model on periodic  $4 \times 3$  lattice. Approximate energies are computed via the two-cluster-marginal relaxation.

	Energy error per site				
Exact energy per site	$1 \times 1$	$2 \times 1$	$2 \times 1$ , p.c.	$2 \times 1$ , overlap	$3 \times 1$ , p.c.
-1.7958	0.6017	0.0634	0.0462	0.0159	0.0048

Table 1: AFH model on periodic  $12 \times 1$  lattice. Approximate energies are computed via the two-cluster-marginal relaxation. Note that ‘p.c.’ indicates the inclusion of the periodicity constraints introduced in section 2.8.1, and ‘overlap’ indicates that the choice of overlapping  $2 \times 1$  clusters, i.e.,  $\{1, 2\}, \{2, 3\}, \{3, 4\}, \dots, \{11, 12\}$ .

## 4.2 Anti-ferromagnetic Heisenberg model

Here we consider the anti-ferromagnetic Heisenberg model (2.2) on a periodic  $12 \times 1$  square lattice, comparing results of the two-cluster-marginal SDP for various cluster sizes. We also test the periodicity constraints of 2.8.1 and the case of overlapping clusters. The results are shown in Table 1.

In Table 2 we show results for the AFH model on a periodic  $4 \times 3$  lattice for various cluster sizes.

## 4.3 Hubbard model

Finally we consider the Hubbard model (3.1) on a non-periodic  $8 \times 1$  lattice with particle numbers  $N = 6, 7, 8, 9, 10$  and interaction strengths  $U \in [0, 12]$ . In Figure 40, we plot results for the two-cluster-marginal relaxation with  $1 \times 1$  clusters  $C_i := \{(i, \uparrow), (i, \downarrow)\}$ . Observe that the for large  $U$ , the energy error (even before normalizing by  $U$ ) is decreasing in  $U$ .

Exact energy per site	Energy error per site		
	1 × 1 clusters	2 × 1 clusters	1 × 3 clusters
-2.4561	1.0439	0.3937	0.0410

Table 2: AFH model on periodic  $4 \times 3$  lattice. Approximate energies are computed via the two-cluster-marginal relaxation.

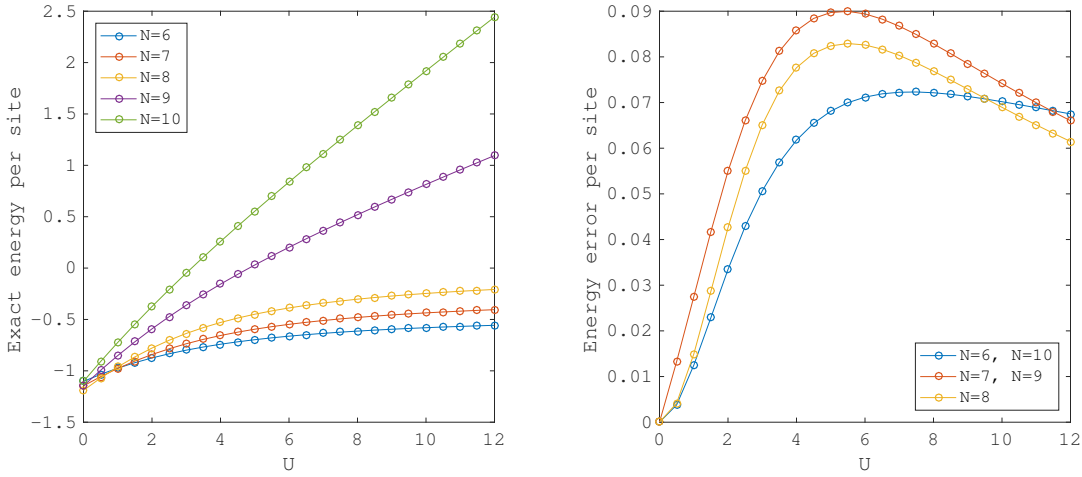


Figure 40: Hubbard model on non-periodic  $8 \times 1$  lattice. Approximate energies are computed via the two-cluster-marginal relaxation with  $1 \times 1$  clusters  $C_i := \{(i, \uparrow), (i, \downarrow)\}$ . Note that the energy errors in the cases  $N = 6$  and  $N = 7$  coincide with the errors in the cases  $N = 10$  and  $N = 9$ , respectively.

## 5 Duality and the effective Hamiltonian perspective

For simplicity, we consider duality only for the two-marginal SDP in the quantum spin setting, and it will be convenient to take the ‘abstract perspective’ of section 2.4, with possibly restricted operator sets as in Remark 2. Duality in other settings can be approached by similar means.

### 5.1 The quantum Kantorovich problem

In preparation for our discussion of the duality of the two-marginal SDP, we first introduce the notion of the quantum Kantorovich problem, which is a direct quantum analog (and in fact generalization) of the Kantorovich problem of optimal transport. The analogy is defined by replacing probability measures with density operators, a cost function with a cost operator  $\hat{C}$ , and classical marginalization with quantum marginalization (i.e., the partial trace). Given operators  $\mu_i \in \text{End}(Q_i)$  for  $i = 1, 2$  of unit trace, we may define the optimal quantum Kantorovich cost via the SDP

$$\begin{aligned} \mathbf{QK}[\hat{C}; \mu_1, \mu_2] &:= \underset{\pi \in \text{End}(Q_1 \otimes Q_2)}{\text{minimize}} && \text{Tr}[C\pi] \\ &\text{subject to} && \pi \succeq 0 \\ &&& \mu_1 = \text{Tr}_{\{2\}}[\pi], \mu_2 = \text{Tr}_{\{1\}}[\pi]. \end{aligned}$$

Note that if  $\mu_1 \not\geq 0$  or  $\mu_2 \not\geq 0$ , then since  $\pi \succeq 0$  implies that  $\text{Tr}_{\{i\}}[\pi] \succeq 0$ , the problem is infeasible, i.e.,  $\mathbf{QK}[C; \mu_1, \mu_2] = +\infty$ . Hence without loss of generality one may assume that  $\mu_i \succeq 0$ , i.e., that the  $\mu_i$  are indeed density operators on  $Q_i$ . Nonetheless, the slightly relaxed perspective will be of some use below. In fact, conversely, the program *is* feasible whenever  $\mu_1, \mu_2 \succeq 0$  because in this case  $\pi = \mu_1 \otimes \mu_2$  is a feasible point.

There is a notion of quantum Kantorovich duality that analogizes the usual notion, as follows. Let the *Hermitian* operators  $A \in \text{End}(Q_1)$  and  $B \in \text{End}(Q_2)$  be dual variables for the first and second marginal constraints, respectively. These will be the ‘quantum Kantorovich potentials.’ Dualizing these constraints yields the Lagrangian

$$\mathcal{L}_{\text{QK}}(\pi, A, B) = \text{Tr}[C\pi] + \text{Tr}[A(\mu_1 - \text{Tr}_{\{2\}}[\pi])] + \text{Tr}[B(\mu_2 - \text{Tr}_{\{1\}}[\pi])]$$

still constrained by  $\pi \succeq 0$ . Using the fact that  $\text{Tr}[A \text{Tr}_{\{2\}}[\pi]] = \text{Tr}[(A \otimes \text{Id})\pi]$  and  $\text{Tr}[B \text{Tr}_{\{1\}}[\pi]] = \text{Tr}[(\text{Id} \otimes B)\pi]$ , we obtain

$$\mathcal{L}_{\text{QK}}(\pi, A, B) = \text{Tr}[A\mu_1] + \text{Tr}[B\mu_2] + \text{Tr}[(C - A \otimes \text{Id} - \text{Id} \otimes B)\pi].$$

Now for fixed  $A, B$ , we have

$$\inf_{\pi \succeq 0} \text{Tr}[(C - A \otimes \text{Id} - \text{Id} \otimes B)\pi] = \begin{cases} 0, & C - A \otimes \text{Id} - \text{Id} \otimes B \succeq 0 \\ -\infty, & \text{otherwise.} \end{cases}$$

Hence we have derived the Kantorovich dual problem

$$\begin{aligned} & \underset{A, B \text{ Hermitian}}{\text{maximize}} && \text{Tr}[A\mu_1] + \text{Tr}[B\mu_2] && (5.1) \\ & \text{subject to} && A \otimes \text{Id} + \text{Id} \otimes B \preceq C. \end{aligned}$$

Strong duality holds by Sion's minimax theorem (together with the compactness of the feasible set of the primal problem).

Let  $\pi$  be the minimizer for the primal problem, and suppose that the dual problem admits a maximizer  $(A, B)$ . Then let  $M = C - A \otimes \text{Id} - \text{Id} \otimes B$ , so

$$\text{Tr}[M\pi] = \text{Tr}[C\pi] - \text{Tr}[(A \otimes \text{Id})\pi] - \text{Tr}[(\text{Id} \otimes B)\pi] = \text{Tr}[C\pi] - \text{Tr}[A\mu_1] - \text{Tr}[B\mu_2] = 0,$$

by primal and dual optimality. But  $\pi \succeq 0$ , so we can write  $\pi = \sum_{i=1}^m p_i \phi_i \phi_i^*$  where  $p_i > 0$ , and  $\text{Tr}[M\pi] = \sum_{i=1}^m p_i \phi_i^* M \phi_i$ . But also  $M \succeq 0$ , so  $p_i \phi_i^* M \phi_i \geq 0$  for all  $i = 1, \dots, m$ . Then since  $\text{Tr}[M\pi] = 0$  it follows that  $\phi_i^* M \phi_i = 0$  for all  $i = 1, \dots, m$ , and since  $M \succeq 0$  this means that  $M\phi_i = 0$  for all  $i = 1, \dots, m$ .

Therefore  $\pi$  is a convex combination of orthogonal projectors onto mutually-orthogonal, degenerate ground state eigenvectors of the Hamiltonian  $C - A \otimes \text{Id} - \text{Id} \otimes B$ . This observation lifts the corresponding observation in the classical setting on the support of the Kantorovich coupling, i.e., that  $\pi_{ij} \geq 0$  only if  $\phi_i + \psi_j = c_{ij}$ , where  $\pi = (\pi_{ij})$ ,  $\phi = (\phi_i)$  and  $\psi = (\psi_i)$  are the Kantorovich potentials, and  $c = (c_{ij})$  is the cost matrix.

In fact, one can consider a regularization of the primal problem by a von Neumann entropy penalty (scaled by  $\beta$ ), for which the solution can be shown to be of the form

$$\pi_\beta = \frac{1}{Z_\beta} \exp[-\beta(C - A_\beta \otimes \text{Id} - \text{Id} \otimes B_\beta)],$$

where  $A_\beta$  and  $B_\beta$  are the unique operators chosen to yield the desired marginals  $\mu_1, \mu_2$ . In the 'zero-temperature' limit  $\beta \rightarrow \infty$  one expects  $\pi_\beta \rightarrow \pi$ ,  $A_\beta \rightarrow A$ , and  $B_\beta \rightarrow B$ .

## 5.2 Partial duality

Now we turn to the discussion of duality. Referring to (2.5), we first consider a *partial* Lagrangian obtained by dualizing *only* the constraint (2.9):

$$\mathcal{L}_{\text{part}}(\{\rho_i\}, \{\rho_{ij}\}, X) = \sum_i \text{Tr}[H_i \rho_i] + \sum_{i < j} \text{Tr}[H_{ij} \rho_{ij}] - \text{Tr}(G[\{\rho_{ij}\}] X),$$

whose domain is defined by  $X \in \mathbb{C}^{(\sum_i m_i^2) \times (\sum_i m_i^2)}$  Hermitian positive semidefinite and  $\{\rho_i\}, \{\rho_{ij}\}$  satisfying constraints (2.6), (2.7), and (2.8).

Now

$$\begin{aligned}
\text{Tr}(G[\{\rho_{ij}\}]X) &= \sum_{ij} \text{Tr}(G_{ij}[\rho_{ij}]X_{ji}) \\
&= \sum_i \sum_{\alpha\beta} \text{Tr} \left[ \rho_i O_{i,\alpha}^\dagger O_{i,\beta} \right] (X_{ii})_{\beta\alpha} \\
&\quad + \sum_{i \neq j} \sum_{\alpha\beta} \text{Tr} \left[ \rho_{ij} \left( O_{i,\alpha}^\dagger \otimes O_{j,\beta} \right) \right] (X_{ji})_{\beta\alpha} \\
&= \sum_i \sum_{\alpha\beta} \text{Tr} \left[ \rho_i O_{i,\alpha}^\dagger O_{i,\beta} \right] (X_{ii})_{\beta\alpha} \\
&\quad + \sum_{i < j} \sum_{\alpha\beta} \left\{ \text{Tr} \left[ \rho_{ij} \left( O_{i,\alpha}^\dagger \otimes O_{j,\beta} \right) \right] (X_{ji})_{\beta\alpha} \right. \\
&\quad \quad \left. + \text{Tr} \left[ \rho_{ji} \left( O_{j,\beta}^\dagger \otimes O_{i,\alpha} \right) \right] (X_{ij})_{\alpha\beta} \right\}.
\end{aligned}$$

Now by the hermiticity of  $X$  we have  $(X_{ji})_{\beta\alpha} = \overline{(X_{ij})_{\alpha\beta}}$ , and we also have the identity

$$\text{Tr} \left[ \rho_{ji} \left( O_{j,\beta}^\dagger \otimes O_{i,\alpha} \right) \right] = \text{Tr} \left[ \rho_{ij} \left( O_{i,\alpha} \otimes O_{j,\beta}^\dagger \right) \right].$$

Therefore

$$\text{Tr}(G[\{\rho_{ij}\}]X) = \sum_i \text{Tr}[Y_i(X_{ii})\rho_i] + \sum_{i < j} \text{Tr}[Y_{ij}(X_{ij})\rho_{ij}],$$

where we have defined the functions  $Y_i : \mathbb{C}^{|\mathcal{I}_i| \times |\mathcal{I}_i|} \rightarrow \text{End}(Q_i)$  and  $Y_{ij} : \mathbb{C}^{|\mathcal{I}_j| \times |\mathcal{I}_i|} \rightarrow \text{End}(Q_i \otimes Q_j)$  by

$$Y_i(M) = \sum_{\alpha\beta} \overline{M}_{\alpha\beta} O_{i,\alpha}^\dagger O_{i,\beta}, \quad Y_{ij}(M) = \left[ \sum_{\alpha\beta} \overline{M}_{\alpha\beta} \left( O_{i,\alpha}^\dagger \otimes O_{j,\beta} \right) \right] + \text{h.c.},$$

where ‘h.c.’ denotes the Hermitian conjugate. Note that if  $M$  is Hermitian, then  $Y_i(M)$  is Hermitian as well, hence  $Y_i(X_{ii})$  and  $Y_{ij}(X_{ij})$  are Hermitian operators.

By applying Sion’s minimax theorem [51] and then separating the infimum over  $\{\rho_i\}, \{\rho_{ij}\}$  into an outer infimum over  $\{\rho_i\}$  (subject to constraint (2.8)) and an inner infimum over  $\{\rho_{ij}\}$  (subject to constraints (2.6) and (2.7)), we may rewrite the two-marginal SDP energy as

$$E_0^{(2)} = \sup_{X \geq 0} \inf_{\{\rho_i\} : \text{Tr}[\rho_i] = 1 \forall i} \mathcal{F}(X, \{\rho_i\}), \quad (5.2)$$

where

$$\mathcal{F}(X, \{\rho_i\}) := \sum_i \text{Tr}[(H_i - Y_i(X_{ii}))\rho_i] + \sum_{i < j} \mathbf{QK}[H_{ij} - Y_{ij}(X_{ij}); \rho_i, \rho_j]. \quad (5.3)$$

This is the form of a concave-convex maxmin problem. The effective domain of the minimization over  $\{\rho_i\}$  is in fact specified by the constraints  $\text{Tr}[\rho_i] = 1$ ,  $\rho_i \succeq 0$  for all  $i$ , because if  $\rho_i \not\succeq 0$  for some  $i$ , then at least one of the quantum Kantorovich problems in the expression for  $\mathcal{F}(X, \{\rho_i\})$  is infeasible, i.e., of infinite optimal cost. The significance of this form is that for fixed  $X, \{\rho_i\}$ , the two-marginals  $\rho_{ij}$  have been entirely decoupled from one another in the evaluation of  $\mathcal{F}(X, \{\rho_i\})$ . Moreover, for each pair  $i < j$ , we see the emergence of an effective Hamiltonians  $H_i^{\text{eff}}(X_{ii}) := H_i - Y_i(X_{ii})$  and  $H_{ij}^{\text{eff}}(X_{ij}) := H_{ij} - Y_{ij}(X_{ij})$  on  $Q_i$  and  $Q_i \otimes Q_j$ , respectively. Notice that the new contributions to these effective Hamiltonians are linear combinations of operators of the form  $O_{i,\alpha}^\dagger O_{i,\beta}$  and  $O_{i,\alpha}^\dagger \otimes O_{i,\beta}$ , respectively. Thus we see how our choice of effective operator lists is reflected in the richness of our class of possible effective Hamiltonians.

### 5.3 Computational significance of partial duality

From the computational point of view, the partial dual perspective is of no small importance. Although general results guarantee that the complexity of solving the two-marginal SDP (2.5) is only polynomial in  $M$ , direct solution (by, e.g., interior-point methods) may still scale quite poorly in practice. One might hope that the complexity should be limited only by  $O(M^3)$  per iteration, i.e., the cost of diagonalizing a matrix of size proportional to  $M$ , since the SDP constraint (2.9) concerns a matrix of size proportional to  $M$ . However, since the semidefinite matrix  $G$  is entangled with further equality constraints, the best guarantees for interior-point methods are far more pessimistic. One can interpret our discussion of duality thus far as revealing a special structure of these equality constraints that allows us in principle to design methods achieving a cost of  $O(M^3)$  per iteration. (We remark that similar considerations could be expected to achieve a cost of  $O(M)$  per iteration for the quasi-local two-marginal SDP with fixed  $d_{\max}$  as discussed in section 3, though we omit details for simplicity.)

Though we will leave implementation along these lines to future work, we will describe how one can compute gradients of  $\mathcal{F}(X, \{\rho_i\})$ , enabling the application of projected gradient ascent-descent methods. For fixed  $X, \{\rho_i\}$ , let  $(A_{ij}^*, B_{ij}^*)$  be the unique dual optimizer (assuming that it exists) for the Kantorovich dual formulation of  $\mathbf{QK}[H_{ij} - Y_{ij}(X_{ij}); \rho_i, \rho_j]$ . Then it follows that

$$\frac{\partial \mathcal{F}}{\partial \rho_k}(X, \{\rho_i\}) = H_k - Y_k(X_{kk}) + \sum_{j>k} A_{kj}^* + \sum_{i<k} B_{ik}^*$$

(Note that if the dual optimizer is not unique, one only gets a supergradient.) One may take a gradient descent step for  $\rho_k$  in the direction of the *traceless part* of  $\frac{\partial \mathcal{F}}{\partial \rho_k}$ , adjusting the step size if necessary to guarantee that  $\rho_k \succeq 0$ . Moreover, letting  $\rho_{ij}^*$  be the primal solution of the Kantorovich problem indicated by  $\mathbf{QK}[H_{ij} + Y_{ij}(X_{ij}); \rho_i, \rho_j]$ ,

we have

$$\frac{\partial \mathcal{F}}{\partial (\overline{X_{ii}})_{\alpha\beta}}(X, \{\rho_i\}) = -\text{Tr}[O_{i,\alpha}^\dagger O_{i,\beta} \rho_i], \quad \frac{\partial \mathcal{F}}{\partial (\overline{X_{ij}})_{\alpha\beta}}(X, \{\rho_i\}) = -\text{Tr} \left[ \left( O_{i,\alpha}^\dagger \otimes O_{j,\beta} \right) \rho_{ij}^* \right].$$

(If the primal optimizer is not unique, one only gets a subgradient.) After taking a gradient ascent step in  $X$ , one may project onto the feasible domain  $\{X \succeq 0\}$  by diagonalizing  $X$  and zeroing all negative eigenvalues.

Efficient methods for solving the primal and dual quantum Kantorovich problems (beyond black-box SDP solvers) will be explored in future work. In particular, preliminary results indicate promise for a quantum analog of the classical Sinkhorn scaling algorithm, for which the computational cost per iteration is roughly given by the cost of diagonalizing certain operators on  $Q_i \otimes Q_j$ .

## 5.4 Full duality

We now turn to deriving the full dual to the original two-marginal SDP. For now we proceed formally, postponing a discussion of strong duality until the end of the section. First introduce dual variables  $\lambda_i \in \mathbb{R}$  for the constraints  $\text{Tr}[\rho_i] = 1$  appearing in the minimization within (5.2), and then exchange supremum and infimum to obtain the problem:

$$\begin{aligned} & \sup_{X \succeq 0, \lambda} \inf_{\{\rho_i\}} \left\{ \sum_i \lambda_i (1 - \text{Tr}[\rho_i]) + \mathcal{F}(X, \{\rho_i\}) \right\} \\ = & \sup_{X \succeq 0, \lambda} \left\{ \sum_i \lambda_i + \inf_{\{\rho_i\}} \left\{ \sum_i \text{Tr}[(H_i - Y_i(X_{ii}) - \lambda_i)\rho_i] + \right. \right. \\ & \left. \left. \sum_{i < j} \mathbf{QK}[H_{ij} - Y_{ij}(X_{ij}); \rho_i, \rho_j] \right\} \right\}. \end{aligned}$$

Now by substituting the Kantorovich dual expression (5.1) for  $\mathbf{QK}$  and then exchanging maximization and minimization, we obtain the problem:

$$\begin{aligned} & \underset{X \succeq 0, \lambda \in \mathbb{R}^M, \{A_{ij}\}, \{B_{ij}\}}{\text{maximize}} \quad \sum_i \lambda_i + \inf_{\{\rho_i\}} \left\{ \sum_i \text{Tr}[(H_i - Y_i(X_{ii}) - \lambda_i)\rho_i] \right. \\ & \qquad \qquad \qquad \left. + \sum_{i < j} \text{Tr}[A_{ij}\rho_i] + \sum_{i < j} \text{Tr}[B_{ij}\rho_j] \right\} \\ & \text{subject to} \quad A_{ij} \otimes \text{Id} + \text{Id} \otimes B_{ij} \preceq H_{ij} - Y_{ij}(X_{ij}), \quad i < j, \\ & \qquad \qquad \qquad X \succeq 0. \end{aligned}$$

Now the expression within the infimum in the objective function can be rewritten

$$\sum_i \text{Tr} \left[ \left( H_i - Y_i(X_{ii}) - \lambda_i + \sum_{j > i} A_{ij} + \sum_{j < i} B_{ji} \right) \rho_i \right],$$



so carrying out the infimum within the objective function, we arrive at the full dual:

$$\begin{aligned}
& \underset{X \succeq 0, \lambda \in \mathbb{R}^M, \{A_{ij}\}, \{B_{ij}\}}{\text{maximize}} && \mathbf{1}^\top \lambda \\
& \text{subject to} && H_i - Y_i(X_{ii}) - \lambda_i + \sum_{j>i} A_{ij} + \sum_{j<i} B_{ji} = 0, \quad i = 1, \dots, M, \\
& && A_{ij} \otimes \text{Id} + \text{Id} \otimes B_{ij} \preceq H_{ij} - Y_{ij}(X_{ij}), \quad 1 \leq i < j \leq M, \\
& && X \succeq 0,
\end{aligned}$$

where the optimization variables  $A_{ij} \in \text{End}(Q_i)$  and  $B_{ij} \in \text{End}(Q_j)$  are understood to be Hermitian.

Strong duality can be understood as follows. Though we have taken an indirect path, the same dual problem could have been derived by directly dualizing the original primal problem (2.5) in the usual fashion. Since the feasible domain for  $\{\rho_i\}, \{\rho_{ij}\}_{i<j}$  in this problem is compact, Sion's minimax theorem applies, and strong duality holds. The question of whether the dual optimizer is attained is more subtle and will be deferred to future work, though see Part VIII for the discussion of strong duality in a similar setting.

## Bibliography

- [1] A. ALTLAND AND B. D. SIMONS, *Condensed matter field theory*, Cambridge Univ. Pr., 2010.
- [2] D. J. AMIT AND V. MARTIN-MAYOR, *Field theory, the renormalization group, and critical phenomena: graphs to computers*, World Scientific Publishing Co Inc, 2005.
- [3] D. G. ANDERSON, *Iterative procedures for nonlinear integral equations*, J. Assoc. Comput. Mach., 12 (1965), pp. 547–560.
- [4] H. ARAKI AND E. H. LIEB, *Entropy inequalities*, Comm. Math. Phys., 18 (1970), pp. 160–170.
- [5] F. ARYASETIAWAN AND O. GUNNARSSON, *The GW method*, Rep. Prog. Phys., 61 (1998), p. 237.
- [6] E. J. BAERENDS, *Exact exchange-correlation treatment of dissociated  $H_2$  in density functional theory*, Phys. Rev. Lett., 87 (2001), p. 133004.
- [7] G. BAYM AND L. P. KADANOFF, *Conservation laws and correlation functions*, Phys. Rev., 124 (1961), p. 287.
- [8] A. D. BECKE, *Density-functional exchange-energy approximation with correct asymptotic behavior*, Phys. Rev. A, 38 (1988), pp. 3098–3100.
- [9] J.-D. BENAMOU, G. CARLIER, AND L. NENNA, *A numerical method to solve multi-marginal optimal transport problems with Coulomb cost*, in Splitting Methods in Communication, Imaging, Science, and Engineering, Springer, 2016, pp. 577–601.
- [10] A. BENLAGRA, K.-S. KIM, AND C. PÉPIN, *The Luttinger-Ward functional approach in the Eliashberg framework: a systematic derivation of scaling for thermodynamics near the quantum critical point*, J. Phys. Condens. Matter, 23 (2011), p. 145601.
- [11] M. BERNARDI, M. PALUMMO, AND J. C. GROSSMAN, *Extraordinary sunlight absorption and one nanometer thick photovoltaics using two-dimensional monolayer materials*, Nano Lett., 13 (2013), pp. 3664–3670.
- [12] S. BIERMANN, F. ARYASETIAWAN, AND A. GEORGES, *First-principles approach to the electronic structure of strongly correlated systems: Combining the GW approximation and dynamical mean-field theory*, Phys. Rev. Lett., 90 (2003), p. 086402.

- [13] P. BILLINGSLEY, *Probability and measure*, Wiley, 2012.
- [14] J.-P. BLAIZOT AND G. RIPKA, *Quantum theory of finite systems*, The MIT Press, 1985.
- [15] P. E. BLÖCHL, T. PRUSCHKE, AND M. POTTHOFF, *Density-matrix functionals from Green's functions*, Phys. Rev. B, 88 (2013), p. 205139.
- [16] M. BORN AND R. OPPENHEIMER, *Zur quantentheorie der molekeln*, Ann. Phys. (Leipzig), 389 (1927), pp. 457–484.
- [17] STEPHEN BOYD AND LIEVEN VANDENBERGHE, *Convex optimization*, Cambridge Univ. Pr., 2004.
- [18] O. BRATTELI AND D. W. ROBINSON, *Operator algebras and quantum statistical mechanics 1*, Springer, 1987.
- [19] R. BULLA, T. A. COSTI, AND T. PRUSCHKE, *Numerical renormalization group method for quantum impurity systems*, Rev. Mod. Phys., 80 (2008), pp. 395–450.
- [20] GIUSEPPE BUTTAZZO, LUIGI DE PASCALE, AND PAOLA GORI-GIORGI, *Optimal-transport formulation of electronic density-functional theory*, Phys. Rev. A, 85 (2012), p. 062502.
- [21] KLAUS CAPELLE AND VIVALDO L. CAMPO, *Density functionals and model Hamiltonians: Pillars of many-particle physics*, Phys. Rep., 528 (2013), pp. 91–159.
- [22] HUAJIE CHEN, GERO FRIESECKE, AND CHRISTIAN B MENDEL, *Numerical methods for a kohn-sham density functional model based on optimal transport*, J. Chem. Theory Comput., 10 (2014), pp. 4360–4368.
- [23] J. P. COE, *Lattice density-functional theory for quantum chemistry*, Phys. Rev. B, 99 (2019), p. 165118.
- [24] A JOHN COLEMAN, *Structure of fermion density matrices*, Rev. Mod. Phys., 35 (1963), p. 668.
- [25] H.D. CORNEAN, V. MOLDOVEANU, AND C.-A. PILLET, *A mathematical account of the negf formalism*, Ann. Henri Poincaré, 19 (2018), p. 411.
- [26] CODINA COTAR, GERO FRIESECKE, AND CLAUDIA KLÜPPELBERG, *Density functional theory and optimal transportation with coulomb cost*, Commun. Pure Appl. Math., 66 (2013), pp. 548–599.
- [27] RAFAEL D. L. M., *The role of the rigged Hilbert space in quantum mechanics*, Eur. J. Phys., 26 (2005), pp. 287–312.

- [28] N. E. DAHLEN, R. VAN LEEUWEN, AND U. VON BARTH, *Variational energy functionals of the green function tested on molecules*, in Int. J. Quantum Chem., vol. 101, 2005, pp. 512–519.
- [29] P. DELIGNE, P. ETINGOF, D. S. FREED, L. C. JEFFREY, D. KAZHDAN, J. W. MORGAN, D. R. MORRISON, AND E. WITTEN, eds., *Quantum fields and strings: A course for mathematicians. Vol. 1, 2*, Amer. Math. Soc., 1999.
- [30] B. DEWITT, *Supermanifolds*, Cambridge University Press, 1984.
- [31] W. F. DONOGHUE, *Monotone Matrix Functions and Analytic Continuation*, Springer, 1974.
- [32] R. ELDER, *Comment on “Non-existence of the Luttinger-Ward functional and misleading convergence of skeleton diagrammatic series for Hubbard-like models”*, arXiv:1407.6599, (2014).
- [33] A. L. FETTER AND J. D. WALECKA, *Quantum theory of many-particle systems*, Courier Corp., 2003.
- [34] R.P. FEYNMAN AND A.R. HIBBS, *Quantum Mechanics and Path Integrals*, McGraw-Hill New York, 1965.
- [35] R.P FEYNMAN AND F.L. VERNON, *The theory of a general quantum system interacting with a linear dissipative system*, Ann. Phys., 24 (1963), p. 118.
- [36] G.B. FOLLAND, *Quantum field theory: a tourist guide for mathematicians*, American Mathematical Society, Providence, 2008.
- [37] A. GEORGES, G. KOTLIAR, W. KRAUTH, AND M. J. ROZENBERG, *Dynamical mean-field theory of strongly correlated fermion systems and the limit of infinite dimensions*, Rev. Mod. Phys., 68 (1996), p. 13.
- [38] K. J. H. GIESBERTZ AND M. RUGGENTHALER, *One-body reduced density-matrix functional theory in finite basis sets at elevated temperatures*, Physics Reports, (2019), pp. 1–47.
- [39] M. GRANT AND S. BOYD, *CVX: Matlab software for disciplined convex programming*, 2013.
- [40] JURI GROSSI, DERK P KOOI, KLAAS JH GIESBERTZ, MICHAEL SEIDL, ARON J COHEN, PAULA MORI-SÁNCHEZ, AND PAOLA GORI-GIORGI, *Fermionic statistics in the strongly correlated limit of density functional theory*, J. Chem. Theory Comput., 13 (2017), pp. 6089–6100.

- [41] E. GULL, A.J. MILLIS, A.I. LICHTENSTEIN, A.N. RUBTSOV, M. TROYER, AND P. WERNER, *Continuous-time monte carlo methods for quantum impurity models*, Rev. Mod. Phys., 83 (2011), p. 349.
- [42] O. GUNNARSSON, G. ROHRINGER, T. SCHÄFER, G. SANGIOVANNI, AND A. TOSCHI, *Breakdown of Traditional Many-Body Theories for Correlated Electrons*, Phys. Rev. Lett., 119 (2017), p. 056402.
- [43] L. HEDIN, *New method for calculating the one-particle Green's function with application to the electron-gas problem*, Phys. Rev., 139 (1965), p. A796.
- [44] P. HOHENBERG AND W. KOHN, *Inhomogeneous electron gas*, Phys. Rev., 136 (1964), pp. B864–B871.
- [45] S. ISMAIL-BEIGI, *Correlation energy functional within the GW -RPA: Exact forms, approximate forms, and challenges*, Phys. Rev. B, 81 (2010), pp. 1–21.
- [46] L. ISSERLIS, *On a formula for the product-moment coefficient of any order of a normal frequency distribution in any number of variables*, Biometrika, 12 (1918), pp. 134–139.
- [47] Y. KHOO, L. LIN, M. LINDSEY, AND L. YING, *Semidefinite relaxation of multi-marginal optimal transport for strictly correlated electrons in second quantization*, arXiv:1905.08322, (2019).
- [48] Y. KHOO AND L. YING, *Convex relaxation approaches for strictly correlated density functional theory*, arXiv:1808.04496, (2018).
- [49] G. KNIZIA AND G. CHAN, *Density matrix embedding: A simple alternative to dynamical mean-field theory*, Phys. Rev. Lett., 109 (2012), p. 186404.
- [50] W. KOHN AND L. SHAM, *Self-consistent equations including exchange and correlation effects*, Phys. Rev., 140 (1965), pp. A1133–A1138.
- [51] H. KOMIYA, *Elementary proof for Sion's minimax theorem*, Kodai Math. J., 11 (1988), pp. 5–7.
- [52] G. KOTLIAR, S. SAVRASOV, G. PÁLSSON, AND G. BIROLI, *Cellular dynamical mean field approach to strongly correlated systems*, Phys. Rev. Lett., 87 (2001).
- [53] G. KOTLIAR, S. Y. SAVRASOV, K. HAULE, V. S. OUDOVENKO, O. PARCOLLET, AND C. A. MARIANETTI, *Electronic structure calculations with dynamical mean-field theory*, Rev. Mod. Phys., 78 (2006), p. 865.
- [54] E. KOZIK, M. FERRERO, AND A. GEORGES, *Nonexistence of the Luttinger-Ward Functional and Misleading Convergence of Skeleton Diagrammatic Series for Hubbard-Like Models*, Phys. Rev. Lett., 114 (2015), p. 156402.

- [55] C. LEE, W. YANG, AND R. G. PARR, *Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density*, Phys. Rev. B, 37 (1988), pp. 785–789.
- [56] M. LEVY, *Universal variational functionals of electron densities, first-order density matrices, and natural spin-orbitals and solution of the  $v$ -representability problem*, Proc. Natl. Acad. Sci., 76 (1979), pp. 6062–6065.
- [57] Y. LI AND J. LU, *Bold diagrammatic Monte Carlo in the lens of stochastic iterative methods*, arXiv:1710.00966, (2017).
- [58] A.I. LICHTENSTEIN AND M.I. KATSNELSON, *Antiferromagnetism and  $d$ -wave superconductivity in cuprates: A cluster dynamical mean-field theory*, Phys. Rev. B, 62 (2000), p. R9283.
- [59] E. H. LIEB, *Density functional for Coulomb systems*, Int J. Quantum Chem., 24 (1983), p. 243.
- [60] N. A. LIMA, L. N. OLIVEIRA, AND K. CAPELLE, *Density-functional study of the Mott gap in the Hubbard model*, Europhys. Lett., 60 (2002), pp. 601–607.
- [61] L. LIN AND M. LINDSEY, *Bold Feynman diagrams and the Luttinger-Ward formalism via Gibbs measures. Part I: Perturbative approach*, arXiv:arXiv:1809.02900, (2018).
- [62] ———, *Bold Feynman diagrams and the Luttinger-Ward formalism via Gibbs measures. Part II: Non-perturbative analysis*, arXiv:arXiv:1809.02901, (2018).
- [63] ———, *Sparsity pattern of the self-energy for classical and quantum impurity problems*, arXiv:1902.04796, (2019).
- [64] L. LIN AND C. YANG, *Elliptic preconditioner for accelerating self consistent field iteration in Kohn-Sham density functional theory*, SIAM J. Sci. Comp., 35 (2013), pp. S277–S298.
- [65] J. M. LUTTINGER AND J. C. WARD, *Ground-state energy of a many-fermion system. II*, Phys. Rev., 118 (1960), p. 1417.
- [66] F. MALET AND P. GORI-GIORGI, *Strong Correlation in Kohn-Sham Density Functional Theory*, Phys. Rev. Lett., 109 (2012), p. 246402.
- [67] FRANCESC MALET, ANDRÉ MIRTSCHINK, KLAAS JH GIESBERTZ, LUCAS O WAGNER, AND PAOLA GORI-GIORGI, *Exchange–correlation functionals from the strong interaction limit of DFT: applications to model chemical systems*, Phys. Chem. Chem. Phys., 16 (2014), pp. 14551–14558.

- [68] R. M. MARTIN, L. REINING, AND D. M. CEPERLEY, *Interacting Electrons*, Cambridge Univ. Pr., 2016.
- [69] DAVID MAZZIOTTI, *Realization of quantum chemistry without wave functions through first-order semidefinite programming*, Phys. Rev. Lett., 93 (2004), p. 213001.
- [70] —, *Structure of fermionic density matrices: Complete  $N$ -representability conditions*, Phys. Rev. Lett., 108 (2012), p. 263002.
- [71] DAVID A MAZZIOTTI, *Contracted Schrödinger equation: Determining quantum energies and two-particle density matrices without wave functions*, Phys. Rev. A, 57 (1998), p. 4219.
- [72] J. R. MCCLEAN ET AL., *OpenFermion: the electronic structure package for quantum computers*, arXiv:1710.07629, (2017).
- [73] C. MEJUTO-ZAERA, L. ZEPEDA-NÚÑEZ, M. LINDSEY, N. TUBMAN, B. WHALEY, AND L. LIN, *Efficient hybridization fitting for dynamical mean-field theory via semi-definite relaxation*, arXiv:1907.07191.
- [74] C. MENDEL AND L. LIN, *Kantorovich dual solution for strictly correlated electrons in atoms and molecules*, Phys. Rev. B, 87 (2013), p. 125106.
- [75] CHRISTIAN B MENDEL, FRANCESC MALET, AND PAOLA GORI-GIORGI, *Wigner localization in quantum dots from kohn-sham density functional theory without symmetry breaking*, Phys. Rev. B, 89 (2014), p. 125106.
- [76] N.D. MERMIN, *Thermal properties of the inhomogeneous electron gas*, Phys. Rev., 137 (1965), p. A1441.
- [77] J. W. NEGELE AND H. ORLAND, *Quantum many-particle systems*, Westview, 1988.
- [78] M. NIELSEN AND I. CHUANG, *Quantum computation and quantum information*, Cambridge University Press, 2001.
- [79] G. ONIDA, L. REINING, AND A. RUBIO, *Electronic excitations: density-functional versus many-body Green's-function approaches*, Rev. Mod. Phys., 74 (2002), p. 601.
- [80] B. PASS, *Multi-marginal optimal transport: theory and applications*, ESAIM: Mathematical Modelling and Numerical Analysis, 49 (2015), pp. 1771–1790.
- [81] J. P. PERDEW, K. BURKE, AND M. ERNZERHOF, *Generalized gradient approximation made simple*, Phys. Rev. Lett., 77 (1996), pp. 3865–3868.

- [82] J. P. PERDEW AND A. ZUNGER, *Self-interaction correction to density-functional approximations for many-electron systems*, Phys. Rev. B, 23 (1981), pp. 5048–5079.
- [83] M. E. PESKIN AND D. V. SCHROEDER, *An introduction to quantum field theory*, (1995).
- [84] M. POLYAK, *Feynman diagrams for pedestrians and mathematicians*, in Graphs and Patterns in Mathematics and Theoretical Physics, 2004, pp. 1–27.
- [85] M. POTTHOFF, *Non-perturbative construction of the Luttinger-Ward functional*, Condens. Matter Phys., 9 (2006), pp. 557–567.
- [86] NV PROKOF'EV AND BV SVISTUNOV, *Bold diagrammatic Monte Carlo: A generic sign-problem tolerant technique for polaron models and possibly interacting many-body problems*, Physical Rev. B, 77 (2008), p. 125101.
- [87] P. PULAY, *Convergence acceleration of iterative sequences: The case of SCF iteration*, Chem. Phys. Lett., 73 (1980), pp. 393–398.
- [88] F. RASSOUL-AGHA AND T. SEPPÄLÄINEN, *A course on large deviations with an introduction to Gibbs measures*, American Mathematical Society, 2015.
- [89] M. REED AND B. SIMON, *Functional analysis*, Academic Press, 1980.
- [90] J. F. RENTROP, V. MEDEN, AND S. G. JAKOBS, *Renormalization group flow of the Luttinger-Ward functional: Conserving approximations and application to the Anderson impurity model*, Phys. Rev. B, 93 (2016), p. 195160.
- [91] R. T. ROCKAFELLAR, *Convex analysis*, Princeton University Press, 1970.
- [92] R. T. ROCKAFELLAR AND R. J.-B. WETS, *Variational analysis*, Springer, 2009.
- [93] K. SCHÖNHAMMER, O. GUNNARSSON, AND R. M. NOACK, *Density-functional theory on a lattice: Comparison with exact numerical results for a model with strongly correlated electrons*, Phys. Rev. B, 52 (1995), p. 2504.
- [94] M. SEIDL, P. GORI-GIORGI, AND A. SAVIN, *Strictly correlated electrons in density-functional theory: A general formulation with applications to spherical densities*, Phys. Rev. A, 75 (2007), p. 042511.
- [95] MICHAEL SEIDL, JOHN P PERDEW, AND MEL LEVY, *Strictly correlated electrons in density-functional theory*, Phys. Rev. A, 59 (1999), p. 51.
- [96] D. SÉNÉCHAL, *An introduction to quantum cluster methods*, arXiv:0806.2690, (2008).



- [97] BRUNO SENJEAN, NAOKI NAKATANI, MASAHISA TSUCHIIZU, AND EMANUEL FROMAGER, *Multiple impurities and combined local density approximations in site-occupation embedding theory*, Theor. Chem. Acc., 137 (2018), pp. 1–21.
- [98] S. SHARMA, J. K. DEWHURST, N. N. LATHIOTAKIS, AND E. K. U. GROSS, *Reduced density matrix functional for many-electron systems*, Phys. Rev. B, 78 (2008), p. 201103.
- [99] P. STAAR, T. MAIER, AND T. C. SCHULTHESS, *Dynamical cluster approximation with continuous lattice self-energy*, Phys. Rev. B, 88 (2013), p. 115101.
- [100] G. STEFANUCCI AND R. VAN LEEUWEN, *Nonequilibrium many-body theory of quantum systems: a modern introduction*, Cambridge Univ. Pr., 2013.
- [101] Q. SUN AND G. K.-L. CHAN, *Quantum embedding theories*, Acc. Chem. Res., 49 (2016), pp. 2705–2712.
- [102] A. SZABO AND N.S. OSTLUND, *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*, McGraw-Hill, New York, 1989.
- [103] W. TARANTINO, P. ROMANIELLO, J. A. BERGER, AND L. REINING, *Self-consistent Dyson equation and self-energy functionals: An analysis and illustration on the example of the Hubbard atom*, Phys. Rev. B, 96 (2017), p. 045124.
- [104] C. VILLANI, *Optimal transport: old and new*, Springer, 2009.
- [105] S. R. WHITE, *Hybrid grid/basis set discretizations of the schrödinger equation*, J. Chem. Phys., 147 (2017), p. 244102.
- [106] F.A. WOLF, A. GO, I.P. MCCULLOCH, A.J. MILLIS, AND U. SCHOLLWÖCK, *Imaginary-time matrix product state impurity solver for dynamical mean-field theory*, Phys. Rev. X, 5 (2015), p. 041032.
- [107] R. WONG, *Asymptotic Approximations of Integrals*, SIAM, 2001.
- [108] J. ZINN-JUSTIN, *Quantum Field Theory and Critical Phenomena*, Clarendon Pr., 2002.

# Appendices

## A Proof of Isserlis-Wick theorem

*Proof.* For any  $b \in \mathbb{R}^N$ , define

$$Z_0(b) = \int_{\mathbb{R}^N} e^{-\frac{1}{2}x^T A x + b^T x} dx. \quad (\text{A.1})$$

Clearly  $Z_0 = Z_0(0)$ . Performing the change of variable  $\tilde{x} = x + A^{-1}b$ , we find

$$Z_0(b) = e^{\frac{1}{2}b^T A^{-1}b} Z_0. \quad (\text{A.2})$$

Observe that, for integers  $1 \leq \alpha_1, \dots, \alpha_m \leq N$ ,

$$\left. \frac{\partial^m Z_0(b)}{\partial b_{\alpha_1} \cdots \partial b_{\alpha_m}} \right|_{b=0} = \int_{\mathbb{R}^N} x_{\alpha_1} \cdots x_{\alpha_m} e^{-\frac{1}{2}x^T A x} dx. \quad (\text{A.3})$$

Combining with Eq. (A.2), we have that

$$\langle x_{\alpha_1} \cdots x_{\alpha_m} \rangle_0 = \left. \frac{\partial^m}{\partial b_{\alpha_1} \cdots \partial b_{\alpha_m}} e^{\frac{1}{2}b^T A^{-1}b} \right|_{b=0}. \quad (\text{A.4})$$

Now write the Taylor expansion

$$e^{\frac{1}{2}b^T A^{-1}b} = \sum_{n=0}^{\infty} \frac{1}{n! 2^n} (b^T A^{-1}b)^n. \quad (\text{A.5})$$

Since the expansion contains no odd powers of the components of  $b$ , the right-hand side of Eq. (A.4) vanishes whenever  $m$  is odd. If  $m$  is an even number, only the term  $n = m/2$  in the Taylor expansion will contribute to the right-hand side of Eq. (A.4). This gives

$$\left. \frac{\partial^m}{\partial b_{\alpha_1} \cdots \partial b_{\alpha_m}} e^{\frac{1}{2}b^T A^{-1}b} \right|_{b=0} = \frac{1}{2^{m/2} (m/2)!} \frac{\partial^m}{\partial b_{\alpha_1} \cdots \partial b_{\alpha_m}} (b^T A^{-1}b)^{m/2}. \quad (\text{A.6})$$

Now one can write

$$\frac{\partial^m}{\partial b_{\alpha_1} \cdots \partial b_{\alpha_m}} = \frac{\partial^m}{\partial^{m_1} b_{\beta_1} \cdots \partial^{m_p} b_{\beta_p}}$$

where the indices  $\beta_1, \dots, \beta_p$  are distinct and  $\sum_{j=1}^p m_j = m$ . Then to further simplify the right-hand side of Eq. (A.6), we are interested in computing the coefficient of the  $b_{\beta_1}^{m_1} \cdots b_{\beta_p}^{m_p}$  term of the polynomial  $(b^T A^{-1}b)^{m/2}$ . Expand  $(b^T A^{-1}b)^{m/2}$  into a sum of monomials as

$$(b^T A^{-1} b)^{m/2} = \sum_{j_1, k_1, \dots, j_{m/2}, k_{m/2}} \prod_{i=1}^{m/2} (A^{-1})_{j_i, k_i} b_{j_i} b_{k_i}.$$

Thus each distinct permutation  $(j_1, k_1, \dots, j_{m/2}, k_{m/2})$  of the multiset  $\{\alpha_1, \dots, \alpha_m\}$ <sup>21</sup> contributes  $\prod_{i=1}^{m/2} (A^{-1})_{j_i, k_i}$  to our desired coefficient.

Since

$$\frac{\partial^m}{\partial^{m_1} b_{\beta_1} \dots \partial^{m_p} b_{\beta_p}} b_{\beta_1}^{m_1} \dots b_{\beta_p}^{m_p} = m_1! \dots m_p!,$$

we obtain from Eq. (A.4) and Eq. (A.6)

$$\langle x_{\alpha_1} \dots x_{\alpha_m} \rangle_0 = \frac{m_1! \dots m_p!}{2^{m/2} (m/2)!} \sum_{(j_1, k_1, \dots, j_{m/2}, k_{m/2})} \prod_{i=1}^{m/2} (A^{-1})_{j_i, k_i}, \quad (\text{A.7})$$

where the sum is understood to be taken over multiset permutations of  $\{\alpha_1, \dots, \alpha_m\}$ .

Now every permutation of  $\mathcal{I}_m$  can be thought of as inducing a permutation of the multiset  $\{\alpha_1, \dots, \alpha_m\}$  via the map  $(l_1, \dots, l_m) \mapsto (\alpha_{l_1}, \dots, \alpha_{l_m})$ . By this map each multiset permutation of  $\{\alpha_1, \dots, \alpha_m\}$  is associated with  $m_1! \dots m_p!$  different permutations of  $\mathcal{I}_m$ .

Moreover, every permutation of  $\mathcal{I}_m$  can be thought of as inducing a pairing on  $\mathcal{I}_m$  by pairing the first two indices in the permutation, then the next two, etc. For example, if  $m = 4$ , then the permutation  $(1, 4, 3, 2)$  induces the pairing  $(1, 4)(3, 2)$ . By this map, each pairing on  $\mathcal{I}_m$  is associated with  $2^{m/2} (m/2)!$  permutations of  $\mathcal{I}_m$  since a pairing does not distinguish the ordering of the pairs, nor the order of the indices within each pair.

Finally, observe that if we take any two permutations of  $\mathcal{I}_m$  associated with a pairing on  $\mathcal{I}_m$  and then consider the (possibly distinct) multiset permutations  $(j_1, k_1, \dots, j_{m/2}, k_{m/2})$  and  $(j'_1, k'_1, \dots, j'_{m/2}, k'_{m/2})$  of  $\{\alpha_1, \dots, \alpha_m\}$  associated to these permutations, then in fact  $\prod_{i=1}^{m/2} (A^{-1})_{j_i, k_i} = \prod_{i=1}^{m/2} (A^{-1})_{j'_i, k'_i}$ .

Thus we can replace the sum over multisets in Eq. (A.7) with a sum over pairings on  $\mathcal{I}_m$ . Each term must be counted  $\frac{2^{m/2} (m/2)!}{m_1! \dots m_p!}$  times because each pairing on  $\mathcal{I}_m$  induces  $2^{m/2} (m/2)!$  permutations of  $\mathcal{I}_m$ , each of which is redundant by a factor of  $m_1! \dots m_p!$ . This factor cancels with the factor in Eq. (A.7) to yield

$$\langle x_{\alpha_1} \dots x_{\alpha_m} \rangle_0 = \sum_{\sigma \in \Pi(\mathcal{I}_m)} \prod_{i \in \mathcal{I}_m / \sigma} (A^{-1})_{\alpha_i, \alpha_{\sigma(i)}}.$$

Recalling that  $G^0 = A^{-1}$ , this completes the proof of Theorem 2.  $\square$

*Remark 4.* In field theories, the auxiliary variable  $b$  is often interpreted as a coupling external field.

<sup>21</sup>For instance, if  $\alpha_1 = 1, \alpha_2 = 1, \alpha_3 = 2$ , there are only three distinct multiset permutations:  $(1, 1, 2)$ ,  $(1, 2, 1)$ , and  $(2, 1, 1)$ .

## B Proof of skeleton decomposition

*Proof.* As suggested in the statement of the proposition, let  $\Gamma^{(k)}$  be the maximal insertions admitted by  $\Gamma$ , where we do not count separately any insertions that differ only by their external labeling (see Remark 25). Let  $h_1^{(k)}$  and  $h_2^{(k)}$  be half edges such that  $\Gamma$  admits the insertion  $\Gamma^{(k)}$  at  $(h_1^{(k)}, h_2^{(k)})$ . Let  $e_1^{(k)}, e_2^{(k)}$  be the external half-edges of the truncated Green's function diagram  $\Gamma^{(k)}$  paired with  $h_1^{(k)}, h_2^{(k)}$ , respectively, in the overall diagram  $\Gamma$ .

First we aim to show that the  $\Gamma^{(k)}$  are disjoint, i.e., share no half-edges. In this case we say that the  $\Gamma^{(k)}$  do not *overlap*. In fact we will show additionally that the external half-edges of the  $\Gamma^{(k)}$  do not touch one another (i.e., are not paired in  $\Gamma$ ), and accordingly we say that the  $\Gamma^{(k)}$  do not *touch*.

To this end, let  $j \neq k$ . The idea is to consider the diagrammatic structure formed by collecting all of the half-edges appearing in  $\Gamma^{(j)}$  and  $\Gamma^{(k)}$  and then arguing by maximality that  $\Gamma^{(j)}$  and  $\Gamma^{(k)}$  cannot overlap or touch, let this structure admit  $\Gamma^{(j)}$  and  $\Gamma^{(k)}$  as insertions.

We now formalize this notion. Let  $H^{(j)}$  and  $H^{(k)}$  be the half-edge sets of  $\Gamma^{(j)}$  and  $\Gamma^{(k)}$ , respectively, and consider the union  $H^{(j,k)} := H^{(j)} \cup H^{(k)}$ , together with a *partial* pairing  $\Pi^{(j,k)}$  on  $H^{(j,k)}$ , i.e., a collection of disjoint pairs in  $H^{(j,k)}$ , defined by the rule that  $\{h_1, h_2\} \in \Pi^{(j,k)}$  if and only if  $\{h_1, h_2\} \in \Pi_\Gamma$  and  $h_1, h_2 \in H^{(j,k)}$ . In other words, the structure  $(H^{(j,k)}, \Pi^{(j,k)})$  is formed by taking all of the half-edges appearing in  $\Gamma^{(j)}$  and  $\Gamma^{(k)}$  and then pairing the ones that are paired in the overall diagram  $\Gamma$ .

Let  $E^{(j,k)}$  be the subset of unpaired half-edges in  $H^{(j,k)}$ , i.e., those half-edges that *do not* appear in any pair in  $\Pi^{(j,k)}$ . Since all half-edges in  $\Gamma^{(k)}$  besides  $e_1^{(k)}, e_2^{(k)}$  are paired in the diagram  $\Gamma^{(k)}$ , we must have  $E^{(j,k)} \subset \{e_1^{(j)}, e_2^{(j)}, e_1^{(k)}, e_2^{(k)}\}$ .

**Lemma 5.**  $|E^{(j,k)}| = 4$ .

*Proof.* We claim that  $|E^{(j,k)}|$  is even. To see this, we first establish that  $|H^{(j,k)}|$  is even. Notice that a truncated Green's function diagram (in particular,  $\Gamma^{(j)}$  and  $\Gamma^{(k)}$ ) contains an even number of half-edges (more specifically  $4n$ , where  $n$  is the order of the diagram). Thus  $|H^{(j)}|, |H^{(k)}|$  are even. Moreover,  $\Gamma^{(j)}$  and  $\Gamma^{(k)}$  share a half-edge if and only if they share the vertex (or interaction line) associated with this half-edge, in which case  $\Gamma^{(j)}$  and  $\Gamma^{(k)}$  share all four half-edges emanating from this vertex. Thus the number  $|H^{(j)} \cap H^{(k)}|$  of half-edges common to  $\Gamma^{(j)}$  and  $\Gamma^{(k)}$  is four times the number of common interaction lines, in particular an even number. So  $|H^{(j,k)}| = |H^{(j)}| + |H^{(k)}| - |H^{(j)} \cap H^{(k)}|$  is even, as desired. Now any partial pairing on  $H^{(j,k)}$  includes an even number of distinct elements, so the number of leftover elements, i.e.,  $|E^{(j,k)}|$ , must be even as well, as claimed.

Next we rule out the cases  $|E^{(j,k)}| \in \{0, 2\}$ .

Suppose that  $|E^{(j,k)}| = 0$ . Then the structure  $(H^{(j,k)}, \Pi^{(j,k)})$  defines a *closed* sub-diagram of  $\Gamma$ , disconnected from the rest of  $\Gamma$ . Since  $\Gamma$  is not closed, the sub-diagram

cannot be all of  $\Gamma$ . This conclusion contradicts the connectedness of  $\Gamma$ .<sup>22</sup>

Next suppose that  $|E^{(j,k)}| = 2$ . Then the structure  $(H^{(j,k)}, \Pi^{(j,k)})$  has two unpaired half-edges, hence defines a truncated Green's function diagram  $\Gamma^{(j,k)}$  that contains both  $\Gamma^{(j)}$  and  $\Gamma^{(k)}$  and admits both as insertions. But since  $\Gamma^{(j)} \neq \Gamma^{(k)}$ ,  $\Gamma^{(j,k)}$  is neither  $\Gamma^{(j)}$  nor  $\Gamma^{(k)}$  (i.e., the containment is proper). This conclusion contradicts the maximality of  $\Gamma^{(j)}$  and  $\Gamma^{(k)}$ .  $\square$

From Lemma 5 it follows that  $E^{(j,k)} = \{e_1^{(j)}, e_2^{(j)}, e_1^{(k)}, e_2^{(k)}\}$ . (And moreover we have  $e_1^{(j)}, e_2^{(j)}, e_1^{(k)}, e_2^{(k)}$  are distinct, which was not assumed!) This establishes one of our original claims, i.e., that the external half-edges of  $\Gamma^{(j)}$  and  $\Gamma^{(k)}$  do not touch (i.e., are not paired.)

We need to establish the other part of our original claim, i.e., that the half-edge sets  $H^{(j)}$  and  $H^{(k)}$  are disjoint. To see this, suppose otherwise, so  $\Gamma^{(j)}$  and  $\Gamma^{(k)}$  share some half-edge  $h$ . Since  $\Gamma^{(j)} \neq \Gamma^{(k)}$ , one of  $H^{(j)}$  and  $H^{(k)}$  does not contain the other, so assume without loss of generality that  $H^{(j)}$  does not contain  $H^{(k)}$ , so there is some half-edge  $h' \in H^{(k)} \setminus H^{(j)}$ . Since  $\Gamma^{(k)}$  is connected there must be some path in  $\Gamma^{(k)}$  connecting  $h$  with  $h'$ .<sup>23</sup> Now  $\Gamma^{(j)}$  can be disconnected from the rest of  $\Gamma$  by deleting the links  $\{e_1^{(j)}, h_1^{(j)}\}$  and  $\{e_2^{(j)}, h_2^{(j)}\}$  from the pairing  $\Pi_\Gamma$ , so evidently our path in  $\Gamma^{(k)}$  connecting  $h$  and  $h'$  must contain either  $(e_1^{(j)}, h_1^{(j)})$  or  $(e_2^{(j)}, h_2^{(j)})$  as a 'subpath.' But then either  $e_1^{(j)}$  or  $e_2^{(j)}$  is paired in  $\Gamma^{(k)}$ , hence also by  $\Pi^{(j,k)}$ , contradicting its inclusion in  $E^{(j,k)}$ .

In summary we have shown that the  $\Gamma^{(k)}$  are disjoint, i.e., share no half-edges, and that the external half-edges of the  $\Gamma^{(k)}$  do not touch one another, i.e., are not paired in  $\Gamma$ . We can define a truncated Green's function diagram  $\Gamma_s$  by considering  $\Gamma$ , then omitting all half-edges and vertices appearing in the  $\Gamma^{(k)}$ . Since the  $\Gamma^{(k)}$  are disjoint and do not touch, this leaves behind the half-edges  $h_1^{(k)}, h_2^{(k)}$  for all  $k$ , which are now left unpaired. Then we complete the construction of  $\Gamma_s$  by adding the pairings  $\{h_1^{(k)}, h_2^{(k)}\}$ . In short, we form  $\Gamma_s$  from  $\Gamma$  by replacing each insertion  $\Gamma^{(k)}$  with the edge  $\{h_1^{(k)}, h_2^{(k)}\}$ . Eq. (4.2) then holds by construction, for a suitable external labeling of the  $\Gamma^{(k)}$ .

Moreover, we find that  $\Gamma_s$  is 2PI. It is not hard to check first that  $\Gamma_s$  is 1PI. Indeed, the unlinking of any half-edge pair in  $\Gamma_s$  that is *not* one of the  $\{h_1^{(k)}, h_2^{(k)}\}$  can be lifted to the unlinking of the same half-edge pair in the original diagram  $\Gamma$ . Since  $\Gamma$  is 1PI, this unlinking does not disconnect  $\Gamma$ . Re-collapsing the maximal insertions once again does not affect the connectedness of the result, so  $\Gamma_s$  does not become disconnected by the unlinking. On the other hand, the unlinking of a half-edge pair  $\{h_1^{(k)}, h_2^{(k)}\}$  were to disconnect  $\Gamma_s$ , then necessarily the unlinking of either  $\{e_1^{(k)}, h_1^{(k)}\}$  or  $\{e_2^{(k)}, h_2^{(k)}\}$  would disconnect  $\Gamma$ , which contradicts the premise that  $\Gamma$  is 1PI.

<sup>22</sup>Interestingly, if one were to try to construct a unique skeleton decomposition for *closed* connected diagrams, i.e., free energy diagrams, this is the place where the argument would fail; see section 4.9.

<sup>23</sup>By such a 'path' we mean a sequence of half-edges  $(h_1, h_2, \dots, h_{2m-1}, h_{2m})$  in  $\Gamma^{(k)}$  such that  $h_1 = h; h_{2m} = h'; h_l, h_{l+1}$  share an interaction line for  $l$  odd; and  $h_l, h_{l+1}$  are paired by  $\Gamma^{(k)}$ .

Thus  $\Gamma_s$  is 1PI, and two-particle irreducibility is equivalent to the absence of any insertions. But if  $\Gamma_s$  admits an insertion containing either  $h_1^{(k)}$  or  $h_2^{(k)}$ , then this contradicts the maximality of the insertion  $\Gamma^{(k)}$  in  $\Gamma$ . Moreover, if  $\Gamma_s$  admits an insertion containing *none* of the  $h_1^{(k)}, h_2^{(k)}$ , then this insertion lifts to an insertion in the original diagram  $\Gamma$ , hence this insertion (i.e., all of its interaction lines and half-edges) must have been omitted from  $\Gamma_s$  (contradiction). We conclude that  $\Gamma_s$  admits no insertions, hence is 2PI.

Finally it remains to prove the uniqueness of the decomposition of Eq. (4.2). To this end, let  $\Gamma_s \in \mathfrak{F}_2^{2\text{PI}}$  and  $\Gamma^{(k)} \in \mathfrak{F}_2^{c,t}$  for  $k = 1, \dots, K$ , and let  $\{h_1^{(k)}, h_2^{(k)}\}$  be distinct half-edge pairs in  $\Gamma_s$  for  $k = 1, \dots, K$ . Then *define*  $\Gamma$  via Eq. (4.2). The uniqueness claim then follows if we can show that the  $\Gamma^{(k)}$  are the maximal insertions in  $\Gamma$ .

Suppose that  $\Gamma^{(k)}$  is not maximal for some  $k$ . Then by definition the diagram  $\Gamma'$  formed from  $\Gamma$  by collapsing the insertion  $\Gamma^{(k)}$  admits an insertion containing  $h_1^{(k)}$  or  $h_2^{(k)}$  (assume  $h_1^{(k)}$  without loss of generality). In fact let  $\Gamma_m^{(k)}$  be a *maximal* insertion containing  $h_1^{(k)}$ . Note also that  $\Gamma'$  still admits the  $\Gamma^{(j)}$  for  $j \neq k$  as insertions.

Then for  $j \neq k$  form the structure  $(H^{(j,k)}, \Pi^{(j,k)})$  by ‘merging’  $\Gamma^{(j)}$  and  $\Gamma_m^{(k)}$  via the same construction as above (now within the overall diagram  $\Gamma'$ ). By the same reasoning as in Lemma 5, the set of unpaired half-edges  $E^{(j,k)}$  must be of even cardinality, and we can rule out the case  $|E^{(j,k)}| = 0$ .

In the case  $|E^{(j,k)}| = 2$ , the structure  $(H^{(j,k)}, \Pi^{(j,k)})$  once again defines a truncated Green’s function diagram  $\Gamma^{(j,k)}$  admitting both  $\Gamma^{(j)}$  and  $\Gamma_m^{(k)}$  as insertions. But since  $\Gamma_m^{(k)}$  is maximal,  $\Gamma^{(j,k)} = \Gamma_m^{(k)}$ , and  $\Gamma^{(j)}$  is contained in  $\Gamma_m^{(k)}$ .

In the case  $|E^{(j,k)}| = 4$ , since  $\Gamma^{(j)}$  does not contain  $\Gamma_m^{(k)}$  (i.e., the half-edge set of the former does not contain that of the latter), the same reasoning as above guarantees that  $\Gamma^{(j)}$  and  $\Gamma_m^{(k)}$  do not overlap or touch.

Then consider the diagram  $\Gamma''$  formed from  $\Gamma'$  by collapsing the insertion  $\Gamma^{(j)}$ . In both of our cases (namely, that  $\Gamma^{(j)}$  is contained in  $\Gamma_m^{(k)}$  and that  $\Gamma^{(j)}$  and  $\Gamma_m^{(k)}$  do not overlap or touch), the insertion  $\Gamma_m^{(k)}$  *descends* to an insertion in  $\Gamma''$  containing  $h_1^{(k)}$ .

Iteratively repeating this procedure for all  $j \neq k$  (i.e., collapsing all of the insertions  $\Gamma^{(j)}$  to obtain the original skeleton diagram  $\Gamma_s$ ), we find that  $\Gamma_m^{(k)}$  descends to an insertion in  $\Gamma_s$  containing  $h_1^{(k)}$ , contradicting the fact that  $\Gamma_s$  is 2PI.  $\square$

## C Definitions and results from convex analysis

In this section we review some definitions and results from convex analysis. In this work many results are stated for concave functions, i.e., functions  $f$  such that  $-f$  are convex. The standard results of convex analysis can always be applied by considering negations. We state results below for convex functions to maintain consistency with the literature. Many results are stated in somewhat more generality than is needed for the purposes of this work (e.g., we do not simply conflate proper and non-proper convex functions). This is done to make sure that the reader can refer to the cited references. The discussion follows developments from Rockafellar [91] and Rockafellar and Wets [92].

### C.1 Convex sets and functions

We begin with the definition of convex sets and functions.

**Definition 6.** A set  $C \subset \mathbb{R}^n$  is *convex* if  $(1-t)x + ty \in C$  for every  $x, y \in C$  and all  $t \in [0, 1]$ .

**Definition 7.** An extended real-valued function  $f$  on a convex set  $C$ , i.e., a function  $f : C \rightarrow [-\infty, \infty] = \mathbb{R} \cup \{-\infty, +\infty\}$ , is *convex* if

$$f((1-t)x + ty) \leq (1-t)f(x) + tf(y)$$

for all  $x, y \in C$  and all  $t \in (0, 1)$ , where we interpret  $\infty - \infty = +\infty$  if necessary. We say that  $f$  is *strictly convex* on the convex set  $C$  if this inequality holds strictly whenever  $x \neq y$ .

**Definition 8.** The (*effective*) *domain* of a convex function  $f$  on  $S$ , denoted  $\text{dom } f$ , is the set  $\text{dom } f = \{x \in S : f(x) < +\infty\}$ .

The following is an immediate consequence of the preceding definitions:

**Lemma 9.** *Let  $f$  be convex on  $S \subset \mathbb{R}^n$ . Then  $\text{dom } f$  is convex.*

We note that when  $f \in C^2(C)$ , our definition of convexity coincides with the definition from multivariate calculus:

**Theorem 10.** *Let  $f \in C^2(C)$ , where  $C \subset \mathbb{R}^n$  is open and convex. Then  $f$  is convex on  $C$  if and only if the Hessian matrix  $\nabla^2 f(x)$  is positive semi-definite for all  $x \in C$ .*

*Proof.* See Theorem 4.5 of Rockafellar [91]. □

Notice that for  $f$  convex on a convex set  $C \subset \mathbb{R}^n$ , we can extend to  $\tilde{f}$  defined on  $\mathbb{R}^n$  by taking  $\tilde{f}|_{\mathbb{R}^n \setminus C} \equiv +\infty$ . It is immediate that  $\tilde{f}$  is convex on  $\mathbb{R}^n$ . Thus one loses no generality by considering only functions that are convex on  $\mathbb{R}^n$ .

The following definitions are helpful for ruling out pathologies:

**Definition 11.** A convex function  $f$  is called *proper* if  $\text{dom } f \neq \emptyset$  and  $f(x) > -\infty$  for all  $x$ .

We will only ever need to consider proper convex functions.

**Definition 12.** If  $f$  is a proper convex function, then  $f$  is called *closed* if it is also lower semi-continuous. (If  $f$  is a non-proper convex function, then  $f$  is called *closed* if it is either  $f \equiv +\infty$  or  $f \equiv -\infty$ .)

*Remark 13.* For the fact that this can be taken as the definition, see Theorem 7.1 of [91].

The convexity of a function guarantees its continuity in a certain sense:

**Theorem 14.** A convex function  $f$  on  $\mathbb{R}^n$  is continuous relative to any relatively open convex set in  $\text{dom } f$ . In particular,  $f$  is continuous on  $\text{int dom } f$ . In fact, it holds that a proper convex function  $f$  is locally Lipschitz on  $\text{int dom } f$ .

*Proof.* See Theorems 10.1 and 10.4 of Rockafellar [91]. □

## C.2 First-order properties of convex functions

There is an extension of the notion of differentiability that is fundamental to the analysis of convex functions.

**Definition 15.** Let  $f$  be a convex function on  $\mathbb{R}^n$ .  $y \in \mathbb{R}^n$  is called a *subgradient* of  $f$  at  $x \in \text{dom } f$  if  $f(z) \geq f(x) + \langle y, z - x \rangle$  for all  $z \in \mathbb{R}^n$ . The *subdifferential* of  $f$  at  $x \in \text{dom } f$ , denoted  $\partial f(x)$ , is the set of all subgradients of  $f$  at  $x$ . By convention  $\partial f(x) = \emptyset$  for  $x \notin \text{dom } f$ .

**Theorem 16.** Let  $f$  be a proper convex function.  $\partial f(x)$  is a non-empty bounded set if and only if  $x \in \text{int dom } f$ .

*Proof.* See Theorem 23.4 of Rockafellar [91]. □

It is perhaps no surprise that the derivative and the subdifferential of a convex function coincide wherever it is differentiable.

**Theorem 17.** Let  $f$  be a convex function, and let  $x \in \mathbb{R}^n$  such that  $f(x)$  is finite. If  $f$  is differentiable at  $x$ , then  $\nabla f(x)$  is the unique subgradient of  $f$  at  $x$ , where  $\nabla$  is the gradient defined with respect to the inner product used to define the subgradient. Conversely, if  $f$  has a unique subgradient at  $x$ , then  $f$  is differentiable at  $x$ .

*Proof.* See Theorem 25.1 of Rockafellar [91]. □



### C.3 The convex conjugate

A fundamental notion of convex analysis is convex conjugation, which extends the older notion of Legendre transformation.

**Definition 18.** Let  $f$  be a function  $\mathbb{R}^n \rightarrow [-\infty, +\infty]$ . Then the *convex conjugate* (or, *Legendre-Fenchel transform*)  $f^* : \mathbb{R}^n \rightarrow [-\infty, +\infty]$  with respect to an inner product  $\langle \cdot, \cdot \rangle$  on  $\mathbb{R}^n$  is defined by

$$f^*(y) = \sup_x \{\langle x, y \rangle - f(x)\} = -\inf_x \{f(x) - \langle x, y \rangle\}.$$

**Theorem 19.** *Let  $f$  be a convex function. Then  $f^*$  is a closed convex function, proper if and only if  $f$  is proper. Furthermore, if  $f$  is closed, then  $f^{**} = f$ .*

*Proof.* See Theorem 12.2 of Rockafellar [91]. □

It is an important fact that the subgradients of  $f$  and  $f^*$  are, in a sense, inverse mappings.

**Theorem 20.** *If  $f$  is a closed proper convex function, then  $x \in \partial f^*(y)$  if and only if  $y \in \partial f(x)$ .*

*Proof.* See Corollary 23.5.1 of Rockafellar [91]. □

Roughly speaking, differentiability of a convex function corresponds to the strict convexity of its conjugate. Indeed:

**Theorem 21.** *If  $f$  is a closed proper convex function, then the following are equivalent:*

1. *int dom  $f$  is nonempty,  $f$  is differentiable on int dom  $f$ , and  $\partial f(x) = \emptyset$  for all  $x \in \text{dom } f \setminus \text{int dom } f$ .*
2.  *$f^*$  is strictly convex on all convex subsets of  $\text{dom } \partial f^* := \{y : \partial f^*(y) \neq \emptyset\}$ .*

*Proof.* See Theorem 11.13 of [92]. □

Note that for proper convex  $f$ , if  $\text{dom } f^*$  is open, then  $\text{dom } \partial f^* = \text{dom } f^*$  by Theorem 16, and under the additional assumption that  $\text{dom } f$  is open, Theorem 21 simplifies to the following:

**Theorem 22.** *Let  $f$  is a lower semi-continuous, proper convex function, and suppose that  $\text{dom } f$  and  $\text{dom } f^*$  are open. Then the following are equivalent:*

1.  *$f$  is differentiable on  $\text{dom } f$ .*
2.  *$f^*$  is strictly convex on  $\text{dom } f^*$ .*

## C.4 Sequences of convex functions

Pointwise convergence of convex functions entails a kind of convergence of their subgradients.

**Theorem 23.** *Let  $f$  be a convex function on  $\mathbb{R}^n$ , and let  $C$  be an open convex set on which  $f$  is finite. Let  $f_1, f_2, \dots$  be a sequence of convex functions finite on  $C$  and converging pointwise to  $f$  on  $C$ . Let  $x \in C$ , and let  $x_1, x_2, \dots$  be a sequence of points in  $C$  converging to  $x$ . Then for any  $\varepsilon > 0$ , there exists  $N$  such that*

$$\partial f_i(x_i) \subset \partial f(x) + B_\varepsilon(0)$$

for all  $i \geq N$ .

*Proof.* See Theorem 24.5 of Rockafellar [91]. □

Besides pointwise convergence, there is in fact another nature of convergence for convex functions. This is the notion of *epi-convergence*, which is defined (even for non-convex functions) as follows:

**Definition 24.** Let  $f_i, f$  be extended-real-valued functions on  $\mathbb{R}^n$ . Then we say that the sequence  $\{f_i\}$  *epi-converges* to  $f$ , written as  $f = e \lim_{i \rightarrow \infty} f_i$  or  $f_i \xrightarrow{e} f$  as  $i \rightarrow \infty$ , if for all  $x \in \mathbb{R}^n$ , the following two conditions are satisfied:

$$\begin{aligned} \liminf_i f_i(x_i) &\geq f(x) && \text{for every sequence } x_i \rightarrow x \\ \limsup_i f_i(x_i) &\leq f(x) && \text{for some sequence } x_i \rightarrow x. \end{aligned}$$

We say that the sequence  $\{f_i\}$  *hypo-converges* to  $f$ , written as  $f = h \lim_{i \rightarrow \infty} f_i$  or  $f_i \xrightarrow{h} f$  as  $i \rightarrow \infty$ , if  $\{-f_i\}$  epi-converges to  $-f$ .

The notion of epi-convergence is particularly natural in the theory of convex functions; accordingly hypo-convergence is more relevant to concave functions. Note also that epi-convergence is neither stronger nor weaker than pointwise convergence. However, there is a useful theorem that relates the pointwise convergence and epi-convergence of convex functions.

**Theorem 25.** *Let  $f_i$  be a sequence of convex functions on  $\mathbb{R}^n$ , and let  $f$  be a lower semi-continuous convex function on  $\mathbb{R}^n$  such that  $\text{dom } f$  has non-empty interior. Then  $f = e \lim_{i \rightarrow \infty} f_i$  if and only if the  $f_i$  converge uniformly to  $f$  on every compact set  $C$  that does not contain a boundary point of  $\text{dom } f$ .*

*Proof.* See Theorem 7.17 of Rockafellar and Wets [92]. □

Under certain mild conditions, the epi-convergence of a sequence of convex functions is equivalent to the epi-convergence of the corresponding sequence of conjugate functions. Indeed, the following theorem is a natural motivation for considering epi-convergence as opposed to pointwise convergence.

**Theorem 26.** *Let  $f_i$  and  $f$  be lower semi-continuous, proper convex functions on  $\mathbb{R}^n$ . Then the  $f_i$  epi-converge to  $f$  if and only if the  $f_i^*$  epi-converge to  $f^*$ .*

*Proof.* See Theorem 11.34 of Rockafellar and Wets [92]. □

Finally, under certain circumstances one can upgrade mere pointwise convergence of convex functions to uniform convergence on compact subsets:

**Theorem 27.** *Let  $f_i$  and  $f$  be finite convex functions on an open convex set  $O \subset \mathbb{R}^n$ , and suppose that  $f_i \rightarrow f$  pointwise on  $O$ . Then  $f_i$  converges uniformly to  $f$  on every compact subset of  $O$ .*

*Proof.* See Corollary 7.18 of Rockafellar and Wets [92]. □

## D Classical results on weak convergence of probability measures

For completeness we recall here several classical results on the weak convergence of measures. For reference, see, e.g., Billingsley [13].

Let  $S$  be a metric space, and let  $\mathcal{P}(S)$  denote the set of probability measures on  $S$  (equipped with the Borel  $\sigma$ -algebra). We say that a sequence  $\mu_k \in \mathcal{P}(S)$  converges weakly to  $\mu \in \mathcal{P}(S)$ , denoted  $\mu_k \Rightarrow \mu$ , if  $\int f \, d\mu_k \rightarrow \int f \, d\mu$  as  $k \rightarrow \infty$  for all bounded, continuous functions  $f : S \rightarrow \mathbb{R}$ . A number of equivalent characterizations of weak convergence are given in the following result, often known as the Portmanteau theorem:

**Theorem 28** (Portmanteau). *Let  $S$  be a metric space, and let  $\mu_k, \mu \in \mathcal{P}(S)$ . The following are all equivalent conditions for the weak convergence  $\mu_k \Rightarrow \mu$ :*

1.  $\lim_{k \rightarrow \infty} \int f \, d\mu_k = \int f \, d\mu$  for all bounded, continuous functions  $f : S \rightarrow \mathbb{R}$ .
2.  $\liminf_{k \rightarrow \infty} \int f \, d\mu_k \geq \int f \, d\mu$  for all lower semi-continuous functions  $f : S \rightarrow \mathbb{R}$  bounded from below.
3.  $\liminf_{k \rightarrow \infty} \mu_k(U) \geq \mu(U)$  for all open sets  $U \subset S$ .

*Remark 29.* There are several other equivalent conditions often included in the statement of this result.

A condition for extracting a weakly convergent subsequence, as guaranteed by Prokhorov's theorem below, is given by the following notion of tightness:

**Definition 30.** Let  $S$  be a metric space equipped with the Borel  $\sigma$ -algebra. A set  $\mathcal{C}$  of measures on  $S$  is called *tight* if for any  $\varepsilon > 0$ , there exists a compact subset  $K \subset S$  such that  $\mu(K) > 1 - \varepsilon$  for all  $\mu \in \mathcal{C}$ . A sequence of measures is called *tight* if the set of terms in the sequence is tight.

**Theorem 31** (Prokhorov). *Let  $S$  be a metric space equipped with the Borel  $\sigma$ -algebra. Then any tight sequence in  $\mathcal{P}(S)$  admits a weakly convergent subsequence.*

## E Proofs of lemmas from Part III

### E.1 Lemma 8

*Proof.* Suppose  $\mu \ll \lambda$  is in  $\mathcal{M}_2$  and write  $d\mu = \rho dx$  where  $\rho$  is the probability density. Since  $\mu \ll \lambda$ ,  $\text{Cov}(\mu)$  must be positive definite. Let  $\mu_G$  be the Gaussian measure with the same mean and covariance as  $\mu$ , and let  $\rho_G$  be the corresponding probability density. Then one can compute that

$$\int \rho \log \rho_G dx = -\frac{1}{2} \log ((2\pi e)^N \det \text{Cov}(\mu))$$

(and in particular this integral is absolutely convergent). Now

$$\rho \log \rho = \rho \log \rho_G + \rho \log \frac{\rho}{\rho_G}.$$

The first term on the right-hand side of this equation is absolutely integrable, and the integral of the second term exists (in particular, the integral of the negative part of the second term is finite, and the value of the full integral is in fact  $-H_{\mu_G}(\mu)$ ). Therefore the integral  $\int \rho \log \rho dx \in (-\infty, \infty]$  exists. Moreover

$$\begin{aligned} H(\mu) &= -\int \rho \log \rho dx \\ &= \frac{1}{2} \log ((2\pi e)^N \det \text{Cov}(\mu)) + H_{\mu_G}(\mu) \\ &\leq \frac{1}{2} \log ((2\pi e)^N \det \text{Cov}(\mu)), \end{aligned}$$

with equality if and only if  $\mu_G = \mu$ .

To prove the second inequality in the statement of the lemma, define  $\bar{\mu} := \int x d\mu$  to be the mean of  $\mu$ . Then  $\text{Cov}(\mu) = \mathcal{G}(\mu) - \bar{\mu}\bar{\mu}^T$ , so in particular  $\det \text{Cov}(\mu) \leq \det \mathcal{G}(\mu)$ , with equality if and only if  $\bar{\mu} = 0$ .  $\square$

### E.2 Lemma 9

*Proof.* Without loss of generality we can assume that  $\mu_j = \rho_j dx$  for all  $j$ .

First, by the Portmanteau theorem for weak convergence of measures (Theorem 28) we have, for any  $z \in \mathbb{R}^n$ , that

$$\begin{aligned} z^T \mathcal{G}(\mu) z &= \int (z^T x)^2 d\mu \leq \liminf_{j \rightarrow \infty} \int (z^T x)^2 d\mu^{(j)} \\ &= \liminf_{j \rightarrow \infty} \int z^T x x^T z d\mu^{(j)} = \liminf_{j \rightarrow \infty} z^T \mathcal{G}(\mu_j) z \leq C \|z\|^2. \end{aligned}$$

It follows that  $\mu \in \mathcal{M}_2$  (and moreover  $\mathcal{G}(\mu) \preceq C \cdot I_n$ ).

Our goal is to put ourselves in a position to use the upper semi-continuity (note our sign convention) of the *relative* entropy with respect to the topology of weak convergence (see Fact 7). Let  $\beta > 0$ , and let  $Z_\beta = \int e^{-\beta\|x\|^2} dx$ . Let  $\gamma_\beta$  be the Gaussian measure with density proportional to  $e^{-\beta\|x\|^2}$ . Then

$$\begin{aligned} H(\mu_j) &= - \int \rho_j \log \rho_j dx \\ &= \log(Z_\beta) - \int \rho_j(x) \log \frac{\rho_j(x)}{\frac{1}{Z_\beta} e^{-\beta\|x\|^2}} dx + \beta \int \rho_j(x) \|x\|^2 dx \\ &= \log(Z_\beta) + H_{\gamma_\beta}(\mu_j) + \beta \text{Tr}[\mathcal{G}(\mu_j)]. \end{aligned}$$

Then by the upper semi-continuity of the *relative* entropy with respect to the topology of weak convergence, we have

$$\limsup_{j \rightarrow \infty} H(\mu_j) \leq \log(Z_\beta) + H_{\gamma_\beta}(\mu) + \beta CN = H(\mu) + \beta (CN - \text{Tr}[\mathcal{G}(\mu)]).$$

Since this inequality holds for any  $\beta > 0$ , the lemma follows.  $\square$

### E.3 Fact 11

*Proof.* We can assume that  $\mu$  is absolutely continuous with respect to the Lebesgue measure, i.e., has a density  $\rho$  (otherwise  $H(\mu) = -\infty$  and the inequality is trivial). It follows that  $\mu_i := \pi_i \# \mu$  are absolutely continuous with respect to the Lebesgue measure, i.e., have densities  $\rho_i$ , for  $i = 1, 2$ . Let  $x = (x_1, x_2)$  denote the splitting of  $x \in \mathbb{R}^N$  according to the product structure  $\mathbb{R}^n = \mathbb{R}^p \times \mathbb{R}^{N-p}$ . Then using the fact that  $\mu_1 \times \mu_2$  has density  $\rho_1(x_1)\rho_2(x_2)$ , one directly computes that

$$\begin{aligned} &H(\mu_1) + H(\mu_2) + H_{\mu_1 \times \mu_2}(\mu) \\ &= \int \rho_1(x_1) \log \rho_1(x_1) dx_1 + \int \rho_2(x_2) \log \rho_2(x_2) dx_2 + \int \rho(x) \log \frac{\rho(x)}{\rho_1(x_1)\rho_2(x_2)} dx \\ &= \int \rho(x) \log \rho_1(x_1) dx + \int \rho(x) \log \rho_2(x_2) dx + \int \rho(x) \log \frac{\rho(x)}{\rho_1(x_1)\rho_2(x_2)} dx \\ &= \int \rho(x) \log \rho(x) dx \\ &= H(\mu). \end{aligned}$$

But by Fact 7, the relative entropy term is non-negative.  $\square$

### E.4 Lemma 2

*Proof.* Upper semi-continuity follows directly from Fatou's lemma.  $\Omega$  is proper because its domain is nonempty and evidently  $\Omega$  does not attain the value  $+\infty$ .

Now let  $\theta \in [0, 1]$  and  $A_1, A_2 \in \text{dom } \Omega$ . Then

$$\begin{aligned}
& \Omega[\theta A_1 + (1 - \theta)A_2] \\
&= -\log \int_{\mathbb{R}^n} \left( e^{-\frac{1}{2}x^T A_1 x - U(x)} \right)^\theta \left( e^{-\frac{1}{2}x^T A_2 x - U(x)} \right)^{1-\theta} dx \\
&\geq -\log \left[ \left( \int_{\mathbb{R}^n} e^{-\frac{1}{2}x^T A_1 x - U(x)} dx \right)^\theta \left( \int_{\mathbb{R}^n} e^{-\frac{1}{2}x^T A_2 x - U(x)} dx \right)^{1-\theta} \right] \\
&= \theta \Omega[A_1] + (1 - \theta) \Omega[A_2],
\end{aligned}$$

where we have used Hölder's inequality in the second step. This establishes concavity. Strict concavity on  $\text{dom } \Omega$  follows from the following fact: Hölder's inequality holds with equality in this scenario if and only if  $e^{-\frac{1}{2}x^T A_1 x - U(x)} = e^{-\frac{1}{2}x^T A_2 x - U(x)}$  for all  $x$ , i.e., if and only if  $A_1 = A_2$ .

Lastly observe that since  $\text{dom } \Omega$  is an open set, for any  $A \in \text{dom } \Omega$ ,

$$\int_{\mathbb{R}^n} e^{\delta x^2} e^{-\frac{1}{2}x^T A x - U(x)} dx < +\infty$$

for some  $\delta > 0$ . Now, for any polynomial  $P$ , there exists a constant  $C$  such that for all  $A'$  in a sufficiently small neighborhood of  $A$ ,

$$P(x) e^{-\frac{1}{2}x^T A' x - U(x)} \leq C e^{\delta x^2} e^{-\frac{1}{2}x^T A x - U(x)}.$$

Since derivatives of all orders of the integrand in (2.2) are of the form

$$P(x) e^{-\frac{1}{2}x^T A x - U(x)},$$

differentiation under the integral is justified, and the smoothness result follows.  $\square$

## E.5 Lemma 16

*Proof.* First assume  $A \in \text{dom } \Omega$ , so  $Z[A] < +\infty$ . Let  $\mu \in \mathcal{M}_2$  and define  $f(x) := \frac{1}{2}x^T A x + U(x)$ . For any  $f$  such that  $e^{-f}$  is integrable, define  $\nu_f$  to be the probability measure with density proportional to  $e^{-f}$ . Then provided  $\mu \ll \lambda$ ,

$$\begin{aligned}
\int f d\mu - H(\mu) &= \Omega[A] - \int \log \left( \frac{1}{Z[A]} e^{-f} \right) d\mu - H(\mu) \\
&= \Omega[A] + \int \log \left( \frac{d\mu}{d\lambda} \right) - \log \left( \frac{d\nu_f}{d\lambda} \right) d\mu \\
&= \Omega[A] + \int \log \frac{d\mu}{d\nu_f} d\mu \\
&= \Omega[A] - H_{\nu_f}(\mu) \geq \Omega[A].
\end{aligned} \tag{E.1}$$

Since  $\mu \in \mathcal{M}_2$ , we have  $H(\mu) < +\infty$  as discussed in Remark 17. Careful observation reveals that manipulations are valid in the sense of the extended real numbers even when  $\int f \, d\mu = +\infty$ . Moreover,  $\mu \not\ll \lambda$  if and only if  $\mu \not\ll \nu_f$ , in which case both sides of (E.1) are  $+\infty$ . Therefore (E.1) holds for all  $\mu \in \mathcal{M}_2$ .

For  $A \in \text{dom } \Omega$ , (E.1) establishes the ‘ $\leq$ ’ direction of (3.3). For  $A \notin \text{dom } \Omega$ ,  $\Omega[A] = -\infty$ , so this direction is immediate.

Next suppose that  $A \in \text{dom } \Omega$ . Since  $\text{dom } \Omega$  is open, it follows that  $\nu_f \in \mathcal{M}_2$ . From (E.1) and the inequality  $-H_{\nu_f}(\mu) \geq 0$  (which holds with equality if and only if  $\mu = \nu_f$ ), it follows that (3.3) holds. Moreover, that the infimum in (3.3) is uniquely attained at  $\mu = \nu_f$ , i.e., at  $d\mu(x) = \frac{1}{Z[A]} e^{-\frac{1}{2}x^T A x - U(x)} \, dx$ .  $\square$

## E.6 Lemma 19

*Proof.* By definition  $\mathcal{F}[G] = -\infty$  whenever  $G \in \mathcal{S}^N \setminus \mathcal{S}_+^N$ . Now we show that also  $\mathcal{F}[G] = -\infty$  for  $G$  on the boundary  $\partial\mathcal{S}_+^N$ . This follows from the fact that for such  $G$ , any  $\mu \in \mathcal{G}^{-1}(G)$  is supported on a subspace of  $\mathbb{R}^n$  of positive codimension, i.e., not absolutely continuous with respect to the Lebesgue measure, and therefore  $H(\mu) = -\infty$ . Moreover, since such  $\mu$  is in  $\mathcal{M}_2$ , we have (via the weak growth condition) that  $\int U \, d\mu \in (-\infty, \infty]$ , so the expression within the supremum of (3.2) is  $-\infty$  for all  $\mu \in \mathcal{G}^{-1}(G)$ .

Meanwhile, for  $G \in \mathcal{S}_{++}^N$ , one can see that  $\mathcal{F}[G] > -\infty$  by considering  $\mu$  to be mean-zero with a compactly supported smooth density, linearly transformed to have the appropriate covariance  $G$ . For such  $\mu$ , both terms in the supremum are finite.

Moreover, for  $G \in \mathcal{S}_{++}^N$  we also have that  $\mathcal{F}[G] < +\infty$ . Indeed, for  $\mu \in \mathcal{G}^{-1}(G)$ , by Lemma 8 we have  $H(\mu) \leq \frac{1}{2} \log [(2\pi e)^n \det G]$ . Since  $\int U \, d\mu \geq -C_U(1 + \text{Tr } G)$ , we have a finite upper bound on the expression in the supremum in (3.2), which finishes the proof.  $\square$

## E.7 Lemma 20

*Proof.* Let  $G_1, G_2 \in \mathcal{S}_{++}^n$ ,  $\theta \in [0, 1]$ , and  $\varepsilon > 0$ . Furthermore let  $\mu_1, \mu_2 \in \mathcal{M}_2$  such that  $\mu_i \in \mathcal{G}^{-1}(G_i)$  and  $\Psi[\mu_i] \geq \mathcal{F}[G_i] - \varepsilon/2$ . Then, noting that  $\theta\mu_1 + (1 - \theta)\mu_2 \in \mathcal{G}^{-1}(\theta G_1 + (1 - \theta)G_2)$ , we observe

$$\begin{aligned} \mathcal{F}[\theta G_1 + (1 - \theta)G_2] &= \sup_{\mu \in \mathcal{G}^{-1}(\theta G_1 + (1 - \theta)G_2)} \Psi[\mu] \\ &\geq \Psi[\theta\mu_1 + (1 - \theta)\mu_2] \\ &\geq \theta\Psi[\mu_1] + (1 - \theta)\Psi[\mu_2] \\ &\geq \theta\mathcal{F}[G_1] + (1 - \theta)\mathcal{F}[G_2] - \varepsilon, \end{aligned}$$

where the penultimate step employs convexity of  $\Psi$ . Since  $\varepsilon$  was arbitrary, we have established concavity.



The fact that  $\mathcal{F}$  is proper follows from Lemma 19. Since  $\mathcal{F}$  is concave, by Theorem 14 it is continuous on  $\text{int dom } \mathcal{F}$ , which is in fact all of  $\text{dom } \mathcal{F}$  by the weak growth assumption. Thus we only need to check upper semi-continuity at points  $G$  outside of  $\text{dom } \mathcal{F}$ . At  $G \notin \overline{\text{dom } \mathcal{F}} = \mathcal{S}_+^n$ , upper semi-continuity is trivial because  $\mathcal{F} \equiv -\infty$  on a neighborhood of  $G$ . Therefore let  $G \in \partial \mathcal{S}_{++}^n$  and suppose that  $G_k \in \mathcal{S}_{++}^n$  such that  $G_k \rightarrow G$  as  $k \rightarrow \infty$ . We need to show that  $\limsup_{k \rightarrow \infty} \mathcal{F}[G_k] = -\infty$ . Throwing out all  $G_k \notin \mathcal{S}_{++}^n$  from the sequence cannot increase the limit superior, so we can just assume that  $G_k \in \mathcal{S}_{++}^n$  for all  $k$ . Since  $G \in \partial \mathcal{S}_{++}^n$ , we have  $\det G = 0$ , and therefore  $\det G_k \rightarrow 0$ . By Lemma 19 we have

$$\mathcal{F}[G_k] \leq \frac{1}{2} \log [(2\pi e)^n \det G_k] + C_U(1 + \text{Tr } G_k).$$

Since the right-hand side of this inequality goes to  $-\infty$  as  $k \rightarrow \infty$ , the proof is complete.  $\square$

## E.8 Lemma 21

*Proof.* Observe that (1)  $\Omega$  and  $\mathcal{F}$  are upper semi-continuous, proper concave functions (by Lemmas 2 and 20), (2)  $\mathcal{F} = \Omega^*$  and  $\Omega = \mathcal{F}^*$ , and (3) both  $\text{dom } \Omega$  and  $\text{dom } \mathcal{F} = \mathcal{S}_{++}^n$  are open. Then the strict concavity and differentiability of  $\mathcal{F}$  on  $\text{dom } \mathcal{F} = \mathcal{S}_{++}^n$  follow directly from Theorem 22.

Now we turn to proving  $C^\infty$ -smoothness. Though infinite-order differentiability is not typically discussed in convex analysis, it can be obtained from infinite-order differentiability and strict convexity of the convex conjugate via the implicit function theorem. Indeed, define the smooth function  $h : \mathcal{S}_{++}^n \times \text{dom } \Omega \rightarrow \mathcal{S}^n$  by

$$h(G, A) = \nabla \Omega[A] - G.$$

Then  $Dh = (-I_{\mathcal{S}^n} \mid \nabla^2 \Omega)$ , and since  $\Omega$  is smooth and strictly concave, the right block is invertible for all  $A, G$ . Fix some  $G' \in \mathcal{S}_{++}^n$ , and let  $A' = \nabla \mathcal{F}[G'] \in \text{dom } \Omega$ , so  $h(G', A') = 0$ . Then the implicit function theorem gives the existence of a smooth function  $\phi$  on a neighborhood  $\mathcal{V} \subset \mathcal{S}_{++}^n$  of  $G'$  such that  $h(G, \phi(G)) = 0$  for all  $G \in \mathcal{V}$ . But this means precisely that  $\phi = \nabla \mathcal{F}$ , hence in particular  $\nabla \mathcal{F}$  is smooth at  $G'$ .  $\square$

## E.9 Lemma 36

*Proof.* Write

$$Z[A, \varepsilon U] = \int e^{-\frac{1}{2}x^T A x - \varepsilon U(x)} dx.$$

We want to show that as  $\varepsilon \rightarrow 0^+$ ,  $Z[\cdot, \varepsilon U]$  epi-converges (see Definition 24) to  $Z[\cdot, \varepsilon U]$ . If so, then  $-\Omega[\cdot, \varepsilon U]$  epi-converges  $-\Omega[\cdot, 0]$ , and Theorems 26 and 25 yield in particular that  $\mathcal{F}[\cdot, \varepsilon U] \rightarrow \mathcal{F}[\cdot, 0]$  pointwise on  $\mathcal{S}_{++}^n$  as  $\varepsilon \rightarrow 0^+$ . Then

by Theorem 23 we have the pointwise convergence of the gradients on  $\mathcal{S}_{++}^n$ , i.e.,  $A[G, \varepsilon U] \rightarrow A[G, 0] = G^{-1}$  as  $\varepsilon \rightarrow 0^+$  for  $G \in \mathcal{S}_{++}^n$ .

Thus it remains to show that  $Z[\cdot, \varepsilon U]$  epi-converges to  $Z[\cdot, \varepsilon U]$ . The first of the conditions in Definition 24 follows immediately from Fatou's lemma, so we need only show that for any  $A \in \mathcal{S}^N$ , there exists a sequence  $A_\varepsilon \rightarrow A$  such that

$$\limsup_{\varepsilon \rightarrow 0^+} Z[A_\varepsilon, \varepsilon U] \leq Z_\varepsilon[A, 0]$$

In particular, it suffices to show that

$$\limsup_{\varepsilon \rightarrow 0^+} Z[A, \varepsilon U] \leq Z_\varepsilon[A, 0]. \quad (\text{E.2})$$

For  $A \notin \mathcal{S}_{++}^n$ , the righthand side is  $+\infty$ , so the inequality holds trivially.

Thus assume  $A \in \mathcal{S}_{++}^n$ . By the weak growth condition, we can write  $U(x) = \tilde{U}(x) - \lambda - \lambda \|x\|^2$ , where  $C > 0$  and  $\tilde{U} \geq 0$ . Then

$$Z[A, \varepsilon U] = \int e^{\varepsilon \lambda} e^{-\frac{1}{2} x^T (A - \varepsilon \lambda) x - \varepsilon \tilde{U}(x)} dx \leq \int e^{\varepsilon \lambda} e^{-\frac{1}{2} x^T (A - \varepsilon \lambda) x} dx,$$

and evidently the righthand side converges to  $Z[A, 0]$  by dominated convergence.  $\square$

## E.10 Lemma 37

*Proof.* Let  $G \in \mathcal{S}_{++}^n$ . Recall Eq. (E.2) from the proof of Lemma 36. From this inequality, it follows that there exists  $\tau > 0$  and an open neighborhood  $\mathcal{N}$  of  $G^{-1}$  in  $\mathcal{S}_{++}^n$  such that  $A \in \text{dom } \Omega[\cdot, \varepsilon U]$  for all  $(\varepsilon, A) \in (0, \tau) \times \mathcal{N}$ .

Now consider  $\hat{\varepsilon} > 0$  sufficiently small so that  $\hat{\varepsilon} < \tau$  and  $\hat{A} := A_G(\hat{\varepsilon}) \in \mathcal{N}$  (possible by Lemma 36). Define the smooth function  $h : (0, \tau) \times \mathcal{N} \rightarrow \mathcal{S}^N$  by

$$h(\varepsilon, A) = \nabla_A \Omega[A, \varepsilon U] - G.$$

Then  $Dh(\varepsilon, A) = (* \mid \nabla_A^2 \Omega[A, \varepsilon U])$ , and since  $\Omega[\cdot, \varepsilon U]$  is smooth and strictly concave, the right block is invertible for all  $\varepsilon, A$ . Moreover, we have  $h(\hat{\varepsilon}, \hat{A}) = 0$  by construction. Then the implicit function theorem gives the existence of a smooth function  $\phi$  on a neighborhood  $I$  of  $\hat{\varepsilon}$  such that  $h(\varepsilon, \phi(\varepsilon)) = 0$  for all  $\varepsilon \in I$ . But this means precisely that  $\phi = A_G$ . The implicit function theorem then also says that

$$A'_G(\varepsilon) = -(\nabla_A^2 \Omega[A_G(\varepsilon), \varepsilon U])^{-1} \frac{\partial h}{\partial \varepsilon}(\varepsilon, A_G(\varepsilon)) \quad (\text{E.3})$$

for all  $\varepsilon \in I$ , where  $A'_G$  denotes the ordinary derivative of the function  $A_G$  of a single variable. In particular Eq. (E.3) holds at  $\varepsilon = \hat{\varepsilon}$ . But since  $\hat{\varepsilon}$  was arbitrary (beyond being taken sufficiently small), it follows that Eq. (E.3) simply holds for all  $\varepsilon > 0$  sufficiently small.

We want to show that all derivatives of  $A_G : (0, \infty) \rightarrow \mathcal{S}^N$  extend continuously to  $[0, \infty)$ . Starting with  $A'_G$ , we can examine these functions by taking further derivatives on the righthand side of Eq. (E.3). The result will be an expression involving integrals of the form

$$\int P(x, U(x)) e^{-\frac{1}{2}x^T A_G(\varepsilon)x - \varepsilon U(x)} dx,$$

where  $P$  is some polynomial, and it suffices to show that such integrals converge to their desired limits

$$\int P(x, U(x)) e^{-\frac{1}{2}x^T G^{-1}x} dx.$$

The argument is by dominated convergence. First observe that from the at-most-exponential growth assumption (Assumption 5), it follows that there exist  $a, b > 0$  such that  $|P(x, U(x))| \leq ae^{b\|x\|}$  for all  $x$ . As in the proof of Lemma 36, write  $U(x) = \tilde{U}(x) - \lambda - \lambda\|x\|^2$ , where  $C > 0$  and  $\tilde{U} \geq 0$ . Then

$$\begin{aligned} |P(x, U(x)) e^{-\frac{1}{2}x^T A_G(\varepsilon)x - \varepsilon U(x)}| &\leq |P(x, U(x))| e^{\varepsilon\lambda} e^{-\frac{1}{2}x^T (A_G(\varepsilon) - \varepsilon\lambda)x - \varepsilon\tilde{U}(x)} \\ &\leq ae^{b\|x\|} e^{\varepsilon\lambda} e^{-\frac{1}{2}x^T (A_G(\varepsilon) - \varepsilon\lambda)x}. \end{aligned}$$

Then for all  $\varepsilon > 0$  small enough such that  $\varepsilon < 1$  and  $A_G(\varepsilon) - \varepsilon\lambda \succ \frac{1}{2}G^{-1}$ , we see that the absolute value of the integrand is bounded uniformly by

$$ae^{b\|x\|} e^{\varepsilon\lambda} e^{-\frac{1}{4}x^T G^{-1}x},$$

which is integrable. This completes the dominated convergence argument, and we conclude that all derivatives of  $A_G$  extend continuously to  $[0, \infty)$ .

Next we aim to use the preceding to show that all derivatives of  $\Phi_G$  and  $\Sigma_G$  also extend continuously to  $[0, \infty)$ .

To this end, recall the Dyson equation

$$\Sigma_G = A_G - G^{-1},$$

which requires that the desired extension property of  $\Sigma_G$  is equivalent to that of  $A_G$ , which we have already proved.

Now for any  $\varepsilon > 0$ , we have

$$\begin{aligned} \Phi_G(\varepsilon) &= 2\mathcal{F}[G, \varepsilon U] - \text{Tr} \log G - N \log(2\pi e) \\ &= \text{Tr}[A_G(\varepsilon)G] - 2\Omega[A_G(\varepsilon), \varepsilon U] - \text{Tr} \log G - N \log(2\pi e) \end{aligned}$$

by Legendre duality, from which it follows from our extension property for  $A_G$ , together with the arguments used to establish it, that all derivatives of  $\Phi_G$  extend continuously to  $[0, \infty)$ .  $\square$

### E.11 Lemma 44

*Proof.* Based on Eqs. (4.6) and (4.7), we want to show that  $G[A^{(M)}(\varepsilon), U_\varepsilon^{(M)}] \sim G[A^{(M)}(\varepsilon), \varepsilon U]$ . As a first step, we aim to show that

$$Z[A^{(M)}(\varepsilon), U_\varepsilon^{(M)}] \sim Z[A^{(M)}(\varepsilon), \varepsilon U].$$

Indeed, we can write

$$\begin{aligned} & Z[A^{(M)}(\varepsilon), \varepsilon U] - Z[A^{(M)}(\varepsilon), U_\varepsilon^{(M)}] \\ &= \int e^{-\frac{1}{2}x^T A^{(M)}(\varepsilon)x - \varepsilon U(x)} \left( 1 - e^{-\frac{1}{2}x^T [\Sigma_G(\varepsilon) - \Sigma_G^{(\leq M)}(\varepsilon)]x} \right) dx. \end{aligned} \quad (\text{E.4})$$

We can choose  $C$  such that

$$-C\varepsilon^{M+1} \preceq \Sigma_G(\varepsilon) - \Sigma_G^{(\leq M)}(\varepsilon) \preceq C\varepsilon^{M+1}$$

for all  $\varepsilon > 0$  sufficiently small.

Now let  $R(\varepsilon) = \varepsilon^{-p/2}$  for  $p \in (0, 1)$ . We split the integral in (E.4) into a part over  $B_{R(\varepsilon)}(0)$  and another part over the complement. The integrand is dominated by  $e^{-\delta x^T x}$  for some  $\delta$  uniform in  $\varepsilon$ , the integral of which over the complement of  $B_{R(\varepsilon)}(0)$  decays super-algebraically as  $\varepsilon \rightarrow 0$ , so we can neglect this contribution.

Meanwhile, for  $x \in B_{R(\varepsilon)}(0)$ , we have

$$\left| x^T \left[ \Sigma_G(\varepsilon) - \Sigma_G^{(\leq M)}(\varepsilon) \right] x \right| \leq C\varepsilon^{M+1-p},$$

hence there exists  $C'$  such that

$$\left| 1 - e^{-\frac{1}{2}x^T [\Sigma_G(\varepsilon) - \Sigma_G^{(\leq M)}(\varepsilon)]x} \right| \leq C'\varepsilon^{M+1-p}$$

for all  $x \in B_{R(\varepsilon)}(0)$ . Combining with (E.4) and dominated convergence, we have established  $Z[A^{(M)}(\varepsilon), U_\varepsilon^{(M)}] \sim Z[A^{(M)}(\varepsilon), \varepsilon U]$ .

This result, together, together with analogous arguments applied to integrals of the form

$$\int x_i x_j e^{-\frac{1}{2}x^T A^{(M)}(\varepsilon)x - \varepsilon U(x)} \left( 1 - e^{-\frac{1}{2}x^T [\Sigma_G(\varepsilon) - \Sigma_G^{(\leq M)}(\varepsilon)]x} \right) dx,$$

yields  $G[A^{(M)}(\varepsilon), U_\varepsilon^{(M)}] \sim G[A^{(M)}(\varepsilon), \varepsilon U]$ . □

### E.12 Lemma 50

*Proof.* For convenience, we define

$$\mathcal{F}_c[G] := \sup_{\mu \in \mathcal{G}^{-1}(G) \cap \mathcal{M}_c} \left[ H(\mu) - \int U d\mu \right].$$

Evidently  $\mathcal{F}_c \leq \mathcal{F}$  and  $\mathcal{F}_c[G] = -\infty$  if  $G \notin \mathcal{S}_{++}^n$ , so we can restrict attention to  $G \in \mathcal{S}_{++}^n$ .

Fix  $\varepsilon > 0$ . Let  $G \in \mathcal{S}_{++}^n$ , so  $\mathcal{F}[G]$  is finite, and let  $\mu \in \mathcal{M}_2$  such that

$$H(\mu) - \int U \, d\mu \geq \mathcal{F}[G] - \varepsilon/2.$$

In particular,  $H(\mu) \neq -\infty$ , so  $d\mu = \rho \, dx$  for some density  $\rho$ . Then consider the measure  $\mu_R \in \mathcal{M}_c(R)$  given by density  $\rho_R := Z_R^{-1} \cdot \rho \cdot \chi_R$ , where  $\chi_R$  is the indicator function for  $B_R(0)$  and  $Z_R = \int_{B_R(0)} \rho \, dx$ . By monotone convergence,  $Z_R \rightarrow 1$ .

Unfortunately we cannot expect  $\mathcal{G}(\mu_R) = G$ , but we do have  $\mathcal{G}(\mu_R) \rightarrow G$  (following from dominated convergence, together with the finite second moments of  $\mu$ ). We then want to modify  $\mu_R$  (keeping its support compact) to construct a nearby measure with the correct second moments.

To this end let  $G_R = \tau_R[G - \mathcal{G}(\mu_R)] + \mathcal{G}(\mu_R)$ , where  $\tau_R > 1$  is chosen so that  $\tau_R \rightarrow +\infty$  and the eigenvalues of  $G_R$  remain uniformly bounded away from zero and infinity (possible since  $\mathcal{G}(\mu_R) \rightarrow G$ ). Note that we have  $G = \tau_R^{-1}G_R + (1 - \tau_R^{-1})\mathcal{G}(\mu_R)$ .

Now let  $\pi \in \mathcal{M}_2$  be any compactly supported measure with a density and finite entropy, and let  $\pi_R = T_R\#\pi$ , where  $T_R$  is a linear transformation chosen so that  $\mathcal{G}(\pi_R) = G_R$ . Since the eigenvalues of  $G_R$  are uniformly bounded away from zero and infinity, the  $T_R$  can be chosen to have determinants uniformly bounded away from zero and infinity (which guarantees that the  $|H(\pi_R)|$  are uniformly bounded), and  $\pi_R$  can be taken to have uniformly bounded support. Then finally we can define a measure  $\nu_R := \tau_R^{-1}\pi_R + (1 - \tau_R^{-1})\mu_R$ , so  $\mathcal{G}(\nu_R) = G$  and  $\nu_R$  is compactly supported.

For the proof it suffices to show that

$$H(\nu_R) - \int U \, d\nu_R \rightarrow H(\mu) - \int U \, d\mu \tag{E.5}$$

as  $R \rightarrow \infty$ .

By the weak growth condition (Definition 3), we can choose a constant  $C$  such that  $\tilde{U}$  defined by  $\tilde{U}(x) := C(1 + \|x\|^2) + U(x)$  satisfies  $\tilde{U}(x) \geq \|x\|^2$ . Now

$$\int (1 + \|x\|^2) \, d\mu_R \rightarrow \int (1 + \|x\|^2) \, d\mu < +\infty$$

by monotone convergence together with the fact that  $Z_R \rightarrow 1$ . Furthermore

$$\tau_R^{-1} \int (1 + \|x\|^2) \, d\pi_R \rightarrow 0,$$

so in fact

$$\int (1 + \|x\|^2) \, d\nu_R \rightarrow \int (1 + \|x\|^2) \, d\mu < +\infty$$

Therefore, without loss of generality, we can prove E.5 under the assumption that  $U(x) \geq \|x\|^2$ . But then  $\int U \, d\mu_R \rightarrow \int U \, d\mu$  by monotone convergence, and

$\tau_R^{-1} \int U \, d\pi_R \rightarrow 0$  since the  $\pi_R$  have uniformly bounded support, so in fact  $\int U \, d\nu_R \rightarrow \int U \, d\mu$ .

Then we need only show that  $H(\nu_R) \rightarrow H(\mu)$ . Here one verifies from the construction that  $\nu_R$  converges weakly to  $\mu$ , and moreover the second moments of  $\nu_R, \mu$  are uniformly bounded, so by Lemma 9, we have  $\limsup_R H(\nu_R) \leq H(\mu)$ .

But by the concavity of the entropy, we have  $H(\nu_R) \geq \tau_R^{-1} H(\pi_R) + (1 - \tau_R^{-1}) H(\mu_R)$ . Now recall that the  $|H(\pi_R)|$  are uniformly bounded in  $R$ , so  $\tau_R^{-1} H(\pi_R) \rightarrow 0$ . Thus the statement  $\liminf_R H(\nu_R) \geq H(\mu)$  (and hence also  $H(\nu_R) \rightarrow H(\mu)$ ) will follow if we can establish  $H(\mu_R) \rightarrow H(\mu)$ .

Now

$$H(\mu_R) = \log(Z_R) - Z_R^{-1} \int_{B_R(0)} \rho \log \rho \, dx.$$

But we know  $Z_R \rightarrow 1$ , so we need only show that

$$\int_{B_R(0)} \rho \log \rho \, dx \rightarrow \int \rho \log \rho \, dx.$$

From Lemma 8, the negative part of  $\rho \log \rho$  is integrable. But then the fact that  $H(\mu) > -\infty$  precisely means that the positive part of  $\rho \log \rho$  is integrable, i.e.,  $\rho \log \rho$  is absolutely integrable. Then the desired fact follows from dominated convergence.  $\square$

## F The noninteracting partition function

We want to compute  $Z_0[A] = Z[A; 0]$  for  $A \in \mathbf{H}^d$ . Fix some  $A \in \mathbf{H}^d$ , and diagonalize  $A = U^\dagger \Lambda U$ , where  $U$  is unitary and  $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_d)$ . Then  $\hat{H}_0[A] = \frac{1}{2} c^\dagger U^\dagger \Lambda U c$ . Define  $\tilde{c}_j = \sum_k U_{jk} c_k$ , so  $\tilde{c}_i^\dagger = \sum_k [U^\dagger]_{ki} c_k^\dagger$ , and let  $\tilde{c} = (\tilde{c}_1, \dots, \tilde{c}_d)^\top$ . Then  $\hat{H}_0[A] = \frac{1}{2} \sum_i \lambda_i \tilde{n}_i$ , where  $\tilde{n}_i := \tilde{c}_i^\dagger \tilde{c}_i$ . It can be readily verified that the operators  $\tilde{c}_i$  and  $\tilde{c}_i^\dagger$  defined via such a *canonical transformation* satisfy the same anticommutation relations as the  $c_i$  and  $c_i^\dagger$ , hence can be viewed as new annihilation and creation operators. Hence in particular the number operators  $\tilde{n}_i$  all commute, and

$$\begin{aligned} Z_0[A] &= \text{Tr} \left[ e^{-\hat{H}_0[A]} \right] \\ &= \text{Tr} \left[ e^{-\sum_k \lambda_k \tilde{n}_k} \right] \\ &= \text{Tr} \left[ \prod_k e^{-\lambda_k \tilde{n}_k} \right]. \end{aligned}$$

Now since  $\tilde{n}_k$  is idempotent,

$$\begin{aligned} e^{-\lambda_k \tilde{n}_k} &= \sum_{m=0}^{\infty} \frac{1}{m!} (-\lambda_k)^m (\tilde{n}_k)^m \\ &= 1 + \left( \sum_{m=1}^{\infty} \frac{1}{m!} (-\lambda_k)^m \right) \tilde{n}_k \\ &= 1 + (e^{-\lambda_k} - 1) \tilde{n}_k. \end{aligned}$$

Then with  $|\tilde{\mathbf{n}}\rangle$  denoting, for  $\mathbf{n} \in \{0, 1\}^d$ , the occupation number basis with respect to the new creation and annihilation operators, we have

$$\begin{aligned} Z_0[A] &= \text{Tr} \left[ \prod_k \left( 1 + [e^{-\lambda_k} - 1] \tilde{n}_k \right) \right] \\ &= \sum_{\mathbf{n} \in \{0,1\}^d} \langle \tilde{\mathbf{n}} | \prod_k \left( 1 + [e^{-\lambda_k} - 1] \tilde{n}_k \right) | \tilde{\mathbf{n}} \rangle \\ &= \sum_{\mathbf{n} \in \{0,1\}^d} \prod_k \left( 1 + [e^{-\lambda_k} - 1] n_k \right) \\ &= \prod_k \sum_{n \in \{0,1\}} \left( 1 + [e^{-\lambda_k} - 1] n \right) \\ &= \prod_k \left( 1 + e^{-\lambda_k} \right) \\ &= \det \left( 1 + e^{-A} \right), \end{aligned}$$

from which it follows that

$$\Omega_0[A] = -\log \det (1 + e^{-A}) = -\text{Tr} [\log (1 + e^{-A})],$$

so

$$D_0[A] = \nabla \Omega_0[A] = \frac{e^{-A}}{1 + e^{-A}} = (1 + e^A)^{-1}.$$

Therefore the inverse map  $A_0[D]$  is given by

$$A_0[D] = \log (D^{-1} - 1)$$

for  $D \in \text{int} \mathcal{D}_d^{(1)}$ .

Now the noninteracting density operator is then given by

$$\rho_0[A] = \frac{1}{Z_0[A]} e^{-c^\dagger A c} = \frac{1}{Z_0[A]} e^{-\sum_k \lambda_k \tilde{n}_k} = \prod_{k=1}^d \frac{1 + [e^{-\lambda_k} - 1] \widehat{n}_k}{1 + e^{-\lambda_k}}.$$

Note that even in the case in which  $\lambda_k = \pm\infty$ ,  $\rho_0[A]$  is then well-defined, and in fact  $\gamma(\rho_0[A])$  can be obtained as a limit of  $\gamma(\rho_0[A^{(m)}])$  for  $A^{(m)} = U^\dagger \Lambda^{(m)} U$  for  $\Lambda^{(m)}$  finite and diagonal such that  $\Lambda^{(m)} \rightarrow \Lambda$ .

Thus if  $D \in \partial \mathcal{D}_d^{(1)}$  and one takes  $D^{(m)} \in \text{int} \mathcal{D}_d^{(1)}$  simultaneously diagonalizable with  $D$  such that  $D^{(m)} \rightarrow D$ , then  $\gamma(\rho_0[A_0[D^{(m)}]]) = D^{(m)} \rightarrow D$ , but also  $\gamma(\rho_0[A_0[D^{(m)}]]) \rightarrow \gamma(\rho_0[A])$ , where  $A = \log(D^{-1} - 1)$  has eigenvalues possibly  $\pm\infty$ .



## G The Kadanoff-Baym contour

Here we briefly discuss one main non-equilibrium setting of interest, known as the *Kadanoff-Baym* formalism. One considers an initial time  $t_0$  and a final time  $t_1$ , with  $t_1 > t_0$ , and for  $t \in [t_0, t_1]$ ,  $\hat{H}(t)$  denotes the Hamiltonian at time  $t$ . This Hamiltonian determines the evolution, starting at time  $t_0$ , of a prepared grand canonical ensemble defined by a density operator  $\rho$ , i.e., a positive semi-definite operator on the Fock space of unit trace. Assuming, for simplicity, strict positive definiteness, we can write

$$\rho = \frac{1}{\text{Tr}[e^{-\beta\bar{H}}]} e^{-\beta\bar{H}}$$

for some Hamiltonian  $\bar{H}$  and inverse temperature  $\beta$ . Of course, this form leaves freedom in choosing  $\beta$ , but it is good to think of  $\beta$  as a free parameter. Often  $\bar{H}$  may be thought of as  $\hat{H}(t_0) - \mu\hat{N}$ , but this need not be the case. To ensure that Assumption 8 holds, it will suffice to assume that  $\text{Tr}[e^{-\beta\bar{H} + \varepsilon\hat{N}}] < +\infty$  for some  $\varepsilon > 0$  sufficiently small. This condition is analogous to the condition  $\mu \in \text{int dom } Z$  discussed in Appendix 5.2 for the equilibrium finite-temperature ensemble. Assuming the condition, let  $\hat{O}_N$  denote the restriction of  $e^{-\beta\bar{H}}$  to the  $N$ -particle subspace. Then it follows that  $\text{Tr}[\hat{O}_N]$  decays exponentially in  $N$ , hence  $\|\hat{O}_N\|_2$  does as well.

Here the contour is the Kadanoff-Baym contour  $\mathcal{C}^{\text{KB}}$ , specified by the path  $\gamma^{\text{KB}}$ , which can be written as a concatenation

$$\gamma^{\text{KB}} = \gamma^- + \gamma^+ + \gamma^{\text{M}}.$$

Here  $\gamma^- : (0, t_1 - t_0) \rightarrow \mathbb{C}$  is defined by  $s \mapsto s + t_0$ ,  $\gamma^+ : (0, t_1 - t_0) \rightarrow \mathbb{C}$  is defined by  $s \mapsto t_1 - s$ , and  $\gamma^{\text{M}} : (0, \beta) \rightarrow \mathbb{C}$  is defined by  $s \mapsto t_0 - is$ . Accordingly we define subcontours,  $\mathcal{C}_\pm$  and  $\mathcal{C}_{\text{KB}}$ . The concatenation  $\gamma^{\text{KB}}$  is viewed as a function  $(s_0, s_1) \rightarrow \mathbb{C}$ , where  $s_0 = 0$  and  $s_1 = 2(t_1 - t_0) + \beta$ .

We have already defined the contour Hamiltonian  $\hat{H}(z)$  for  $z \in \mathcal{C}_\pm$ . To complete the specification of our ensemble we stipulate that  $\hat{H}(z) = \bar{H}$  for  $z \in \mathcal{C}_{\text{M}}$ . For contour times  $s, s' < t_1 - t_0$ , the contour-ordered Green's function  $G(s, s')$  recovers the appropriate notion of the real-time-ordered non-equilibrium Green's function; similarly, appropriate notions of advanced and retarded Green's functions can be recovered from the contour-ordered Green's function. However, only the contour-ordered Green's function admits a favorable perturbation theory, and this remarkable fact is one motivation for considering it. See [100] for further details. In this work we additionally see that the contour-ordered setting is also the natural setting in which to recover a sparsity result for the self-energy of impurity problems in the non-equilibrium setting.

Now one can readily check that the partition function is given by  $Z = \text{Tr}[e^{-\beta\bar{H}}] > 0$  (so Assumption 9 is satisfied). Now we verify Assumption 8. For  $s' \leq s \leq s_1 - \beta$ , note that  $U(s, s')$  is unitary, hence bounded. Moreover, for  $s_1 - \beta \leq s' \leq s$ , we

have  $U(s, s') = e^{-(s-s')\overline{H}}$ , which is trace class (by our assumption), hence bounded. It follows that for any  $s_0 \leq s' \leq s \leq s_1$ , the operator  $U(s, s')$  is bounded. In fact,  $U(s_1, s_1 - \beta) = e^{-\beta\overline{H}}$ , and as mentioned above, the operator norm of this operator restricted to the  $N$ -particle subspace decays exponentially in  $N$ . Thus Assumption 8 is satisfied.