UNIVERSITY OF CALIFORNIA,
IRVINE

Stochastic control methods in non-equilibrium thermodynamics
Fundamental bounds on dissipation and power

DISSERTATION

submitted in partial satisfaction of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

in Mechanical and Aerospace Engineering

by

Rui Fu

Dissertation Committee:
Professor Tryphon T. Georgiou, Chair
Professor Solmaz S. Kia
Professor Athanasios Sideris

2020

DEDICATION

To my family

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

First of all, I would like to express my deepest gratitude and thanks to my advisor Professor Tryphon Georgiou for his valuable guidance and continuous support during my PhD study. He is the person I have ever seen who loves research so much from the heart. His passion on research, attitude to the detail, and encouragement to students really impressed me. He is always willing to spend time explaining very basic things that confuse me, and telling me how these small things matter. He is also very happy to work and discuss the detailed computation together. What's more, he spent a lot of time to teach me how to do a good presentation with endless patience. Now I enjoyed research itself more and more. I could not imagine how lucky I am to have a so nice, kind, and supportive advisor like him for my PhD study.

I would like to express my sincere gratitude to my committee members Professor Solmaz Kia, professor Katya Krupchyk, Professor Athanasios Sideris, and Professor Haithem Taha for their help and insightful comments on my research.

I would like to express my sincere thanks to Dr. Amirhossein Taghvaei, a postdoctoral fellow in our research team. Thanks for his guidance, and endless patience to work with me. I am grateful for his kind heart to encourage me when I lacked confidence. This dissertation could not be completed without his help and support.

Many thanks to my research colleagues, Zahra Askarzadeh (such a good time working day and night with you and Abhishek for the paper [1,2]), Yongxin Chen, Olga Movilla, Abhishek Halder, Amirhossein Karimi, Anqi Dong, and Fariba Ariaei. My thanks also to all my friends, especially Yan Li, Cheng Chen, Rong Kou, Bo Zheng, Yan Zhao, and my badminton friends, for making life at UC Irvine a great, wonderful and enjoyable experience.

Last but not least, I would express my sincere thanks to my family: my parents, brother and sister-in-law for their unconditional love and support through all the years. I am very thankful to the special person in my life, Ke Zhang. Thank you for all you do for me, and all the love and support you give me.

Chapter 3 is taken, in full, from the paper that is published as "Maximal power output of a stochastic thermodynamic engine" by Rui Fu, Amirhossein Taghvaei, Yongxin Chen and Tryphon Georgiou in Automatica, 2020 [3]; as well as, "Maximal power output of a stochastic thermodynamic engine" by Rui Fu, Olga Movilla, Amirhossein Taghvaei, Yongxin Chen and Tryphon Georgiou in the 2020 American Control Conference. The dissertation author is the primary investigator and author of these papers.

Chapter 4 is taken, in part, from the paper published as "Harvesting energy from a periodic heat bath" by Rui Fu, Olga Movilla, Amirhossein Taghvaei, Yongxin Chen and Tryphon Georgiou in the 59th IEEE Conference of Decision and Control, 2020. The dissertation author is the primary investigator and author of this paper.

Chapter 5 is taken, in part, from the paper "Thermodynamic bounds on power under an

arbitrary temperature profile" by Rui Fu, Olga Movilla, Amirhossein Taghvaei, Yongxin Chen and Tryphon Georgiou, that is prepared to submit to Physical Review Letters, 2020. The dissertation author is the primary investigator and author of this paper.

VITA

Rui Fu

EDUCATION

PhD student, in Mechanical and Aerospace Engineering **2016-2020**
University of California, Irvine *Irvine, USA*

M.S., in Mechanical and Aerospace Engineering **2019**
University of California, Irvine *Irvine, USA*

M.S., in Applied Mathematics **2016**
Northwestern Polytechnical University *Xi'an, China*

B.S., in Applied Mathematics **2013**
Xuchang University *Xuchang, China*

RESEARCH EXPERIENCE

Graduate Research Assistant **2017–2020**
University of California, Irvine *Irvine, USA*

REFEREED JOURNAL PUBLICATIONS

Stability theory of stochastic models in Opinion Dynamics **2019**
IEEE Transactions on Automatic Control

Maximal power output of stochastic thermodynamic engine **2020**
Automatica

Thermodynamic bounds on power under an arbitrary temperature profile (Part I) **2020**
Physical review letters (to be submitted)

Thermodynamic bounds on power under an arbitrary temperature profile (Part II) **2020**
Physical review letters (to be submitted)

REFEREED CONFERENCE PUBLICATIONS

Opinion Dynamics over Influence Networks **2019**
The 2019 American Control Conference (ACC)

Maximal power output of a stochastic Thermodynamic engine **2020**
The 2020 American Control Conference (ACC)

Harvest energy from a periodic heat bath **2020**
the 59th IEEE Conference of Decision and Control, 2020 (CDC)

ABSTRACT OF THE DISSERTATION

Stochastic control methods in non-equilibrium thermodynamics
Fundamental bounds on dissipation and power

By

Rui Fu

Doctor of Philosophy in Mechanical and Aerospace Engineering

University of California, Irvine, 2020

Professor Tryphon T. Georgiou, Chair

Classical thermodynamics is aimed at quantifying the efficiency of thermodynamic engines by bounding the maximal amount of mechanical energy produced compared to the amount of heat required under quasi-static operation. While this was accomplished early on, by Carnot and Clausius, the more practical problem of quantifying limits of output power remained elusive due to the fact that quasi-static processes require infinitely slow cycling, resulting in a vanishing power. Recent insights, drawn from stochastic thermodynamics, bridge the gap between theory and practice by presenting fresh approaches that lead to general laws applicable to the non-equilibrium system. Remarkably, the problem of minimizing dissipation over a finite time window can be expressed as a stochastic control problem leading to physically meaningful expressions for the dissipation cost in thermodynamic engines.

Building on the framework of *stochastic thermodynamics*, we derive bounds on the maximal power that can be drawn by cycling an overdamped ensemble of particles via a time-varying potential while alternating contact with heat baths of different temperature (T_c cold, and T_h hot). The previous work focused on the setting where the potential is quadratic and the distributions are Gaussian. Our analysis relaxes this assumption and asks about fundamental bounds on the power under arbitrary distribution. Specifically, first by casting

the optimization problem of the power into the Benamou-Brenier formulated Optimal mass transport problem, we show that the power output is bounded by the Fisher information of the boundary distribution, which can become unbounded with complicated and irregular control. However, it is unreasonable to expect technological solutions to such demands, and therefore, a constraint on the complexity of the potential seems meaningful. Assuming a suitable bound M on the spatial gradient of the controlling potential, we show that the maximal achievable power is bounded by $\frac{M}{8}(\frac{T_h}{T_c}-1)$. Moreover, we show that this bound can be reached to within a factor of $(\frac{T_h}{T_c}-1)/(\frac{T_h}{T_c}+1)$ by operating the cyclic thermodynamic process with a quadratic potential.

While the majority of works in the literature are concerned with thermodynamic engines operating in the setting of Carnot's cyclic contact with alternating heat baths, the natural processes in living organisms do not follow this setting. Instead, it is the periodic fluctuations in chemical concentrations in conjunction with the variability of electrochemical potentials that provide the universal source of cellular energy. Thus, energy exchange is often mediated by continuous processes and energy differentials, whereas the Carnot cycle reflects the switching mechanics of an idealized engine. We herein propose to consider thermodynamic processes driven by a heat bath with periodic and continuous temperature profile, and study questions of maximal power and efficiency at maximal power. Our results state that the maximal power satisfies a bound proportional to the average fluctuations in the temperature profile. Moreover, we show the surprising result that the efficiency at maximal power does not depend on the specific temperature profile, only the maximum and minimum.

Chapter 1

Introduction

1.1 Background and Motivation

Thermodynamics is the branch of physics which is concerned with the relation between heat and other forms of energy. Historically, it was born out of the quest to quantify the maximal efficiency of heat engines, i.e., the maximal ratio of the total work output over the total heat input to a thermodynamic system. This was accomplished in the celebrated work of Carnot [4, 5] where, assuming that transitions take place infinitely slowly (*quasi-static* operation), it was shown that the maximal efficiency possible is $\eta_C = 1 - T_c/T_h$ (*Carnot efficiency*). In this expression, T_h and T_c are the absolute temperatures of two heat reservoirs, hot and cold respectively, with which the heat engine makes contact with during phases of a periodic operation known as *Carnot cycle*.

Carnot's result provides the absolute theoretical limit for the efficiency of a heat engine, but provides no insight on the amount of power output that can be achieved. Specifically, in order to reach Carnot efficiency, the period of the Carnot cycle must tend to infinity, resulting in vanishing total power output. Whereas, to achieve non-vanishing power output,

a thermodynamic process must take place in finite time, and thereby, away from *equilibrium* [6–8]. To this end, the framework of stochastic thermodynamics [9–23] has been developed in recent years, to allow quantifying work in non-equilibrium thermodynamic transitions. It is rooted in probabilistic models in the form of stochastic differential equations to specify the behaviour of particles in a thermodynamic ensemble. Manipulation of the ensemble is effected by a confining potential that serves as a *control input*. This potential, together with a heat reservoir, couples the (*canonical*) ensemble to the environment. The transferred work and heat can then be computed at the level of individual particles and averaged over the ensemble. Important goals of the theory have been to assess the amount of work needed for *bit-erasure in finite time* [24–26] and hence computation, i.e., a finite-time Landauer bound, as well as assessing the efficiency of thermodynamic engines operating at maximal power.

The question of efficiency at maximal power was studied independently by Chambadal [27], Novikov [28] and Curzon and Ahlborn [29] based on a certain "endoreversible" assumption to reflect finite-time heat transfer. They derived the bound $\eta_{CA} = 1 - \sqrt{T_c/T_h} = 1 - \sqrt{1-\eta_C}$ for efficiency, where the T_h and T_c designate temperatures of a hot and cold heat reservoir respectively, at maximal power $k(\sqrt{T_h} - \sqrt{T_c})^2$, with k being the heat conductance. Then, based on the low-dissipation assumption, the efficiency at maximum power has been obtained in [30, 31], which is reduced to the Curzon-Ahlborn efficiency for symmetric coupling to the cold and hot bath. Subsequent works, for example, Chen and Yan [32], investigated maximal power and the efficiency at maximal power when the heat transfer is proportional to the difference of inverse temperatures instead of Newton's cooling Law, which gives $\eta_{CY} = \eta_C/(2 - \gamma\eta_C)$. Here $\gamma = 1/(1 + \sqrt{\beta/\alpha})$, and α, β represent the heat transfer coefficient between the working substance and the cold or the cold reservoir, respectively.

The work mentioned so far for maximal power are all about the heat engine in the macroscopic scale. The existence of these results is of fundamental importance. Yet, these authors did not address the crucial question of how to actually reach this maximal power and effi-

ciency, as they did not provide information about optimal engine cycles. With the development of stochastic thermodynamics during the last twenty years, fresh approaches have rebuilt general laws that can be applicable to a system driven out of equilibrium. These approaches provide new insights to analyze and quantify the power extracted and the efficiency at maximal power by studying the thermodynamic system in the microscopic scale. In contrast to the macroscopic heat engines considered in conventional endoreversible thermodynamics, thermal fluctuations play an important role in most biologically relevant systems and dynamics. In this regime, the mathematical model for thermodynamic engines must incorporate the effect of fluctuation and thus allow also for backward steps even in a directed motion [33–35]. A notable contribution in this regard, which is closely related to the subject of this thesis, is [35], concerned with a class of Brownian Carnot heat engines that consists of two adiabatic and two isothermal steps modeled by the overdamped Langevin equation, and studies the problem of maximal power and efficiency at maximal power under the assumption of Gaussian boundary distributions and quadratic potential. The experimental realization of the Brownian Carnot engine with a single microscopic particle as a working substance that allows to transform the heat transferred from thermal fluctuations into mechanical work, is reported in [36]. The similar problem in low dissipation regime has been studied in [31] and [37]. Analysis of thermodynamic engines in a quantum mechanical setting appears in [38].

1.2 Objectives and Contributions

The contributions of this dissertation lie in providing insights and quantitative bounds for the maximal power that a heat engine can achieve in the framework of stochastic thermodynamics, control protocols to attain these bounds and more broadly drawing a link between thermodynamics and control. Specifically this dissertation investigates the optimal perfor-

mance of the periodically driven thermodynamic heat engine in three different settings. The objectives and contributions of each chapter are as follows.

- Chapter 3: Maximal power output of a stochastic Carnot-like thermodynamic engine

  We study a thermodynamic engine modeled by an over-damped Langevin equation and operating through a Carnot-like cycle. In this, we consider the problem of maximizing power and efficiency at maximal power, and seek to determine the optimal control law that maximizes the power in finite time transitions. We relax the assumptions considered in earlier literature on Gaussian boundary distributions being Gaussian and the potential quadratic. We herein propose a physically motivated constraint on the control effort, and study the respective optimal performance on power. A physically reasonable upper bound of the power is derived, and given in the form of the ratio of temperatures of the two heat baths.

  The contributions are summarized as follows:

  - We make a connection between the maximal power output problem in finite time transition of a stochastic thermodynamic engine and the Benamou-Brenier formulated optimal mass transport problem.
  - We obtain an upper bound for the power which is proportional to the Fisher information of the boundary distributions.
  - With a suitable constraint on the potential, we derive a lower and an upper bound for the maximal power that can be extracted.

- Chapter 4: Maximal power output of a sinusoidally driven thermodynamic engine

  We study a thermodynamic engine driven by a sinusoidal temperature profile and modeled by the over-damped Langevin equation. We consider the problem of maximizing power and efficiency at maximal power, and seek to determine the control input that maximizes the power in finite time transitions. First, we consider a sinusoidal control

input and use the Fourier representations method to quantify power output of the engine, and then maximize the power with respect to the phase and amplitude of the control input. Second, we consider maximizing power of the heat engine with sinusoidal temperature profile and a general control input in the linear response regime. The result we get shows that the optimal protocol for the sinusoidal temperature driven heat engine is also sinusoidal but of different amplitude and phase.

The contributions are summarized as follows:

- We formulate the problem of optimal performance for a sinusoidally driven heat engine as an optimal control problem.
- We obtain an explicit form for the optimal control law, and show that the optimal protocol of a sinusoidally driven heat engine is also sinusoidal function.
- We derive the optimal power output in the linear response regime, and numerically verify our theoretical results with a numeral method based on Fourier representation.

- Chapter 5: Thermodynamic bounds on power under arbitrary temperature profile
  We study a thermodynamic engine driven by a heat bath with continuously and periodically changing temperature profile, modeled by the over-damped Langevin equation. We consider the problem of maximizing power and efficiency at maximal power and seek to determine the control input that maximizes power in finite time transitions.

  The contributions are summarized as follows:

  - We connect the problem of the optimal performance to a control problem.
  - We obtain a upper bound on the power that any periodically and continuously driven heat engine can achieve, which is proportional to the average fluctuations in the temperature profile.

– We show that the efficiency at maximal power does not depend on the temperature profile.

Chapter 2

The history of thermodynamics

2.1 Classical thermodynamics

Classical thermodynamics describes the states of thermodynamic systems at near equilibrium, and models the exchanged of energy, work and heat based on the Laws of thermodynamics as its heart.

2.1.1 Laws of thermodynamics

For the First law of thermodynamics,

$$\Delta E = \Delta Q - W, \tag{2.1}$$

where ΔE denotes the change in the internal energy, Q denotes the quantity of the heat that is supplied to the system, and W denotes the amount of the work done by the system. It is a version of conservation of energy, states that the total energy of an isolated system is

constant. Energy can be transferred from one form to another, but can neither be destroyed or created. For the Second law of thermodynamics,

$$\Delta S \geq 0, \tag{2.2}$$

where S denotes the entropy of the system. It states that the total entropy production can never decrease. It can remain constant in the ideal case where the system is in thermodynamic equilibrium, or is going through a reversible process. In all processes that occur, including spontaneous processes, the total entropy of the system and its surroundings increases and the process is irreversible in the thermodynamic sense.

2.1.2 Carnot heat engine

Usually the second law of thermodynamics is rationalized by considering processes that involve transferring heat into work by engines, which return to their original thermodynamic states after completion of one cycle. Carnot heat engine is a theoretical engine going through a reversible Carnot cycle. The Carnot cycle is an ideal thermodynamic cycle proposed by the French physicist Sadi Carnot in 1824, who gives a theoretical upper bound on the efficiency that any classical thermodynamic engine can achieve when it operates between two heat baths with different temperatures during the conversion of heat into work, or conversely, the efficiency of a refrigerator that creates the temperature difference with surroundings by applying work to the system. The Carnot cycle consists of four phases as follows [39]:

- Isothermal expansion: Heat is transferred from the hot reservoir with constant temperature T_h. The gas particle starts to expand, and the work is done on the surroundings by pushing up the piston, which is shown in Figure 2.1. During this process, even though the pressure of the gas drops down, the temperature inside the piston keeps constant as it is continuously contacted with the hot reservoir, from which the heat

Q_h is absorbed, leading to the increase of the entropy of the gas $\Delta S = \frac{Q_h}{T_h}$.

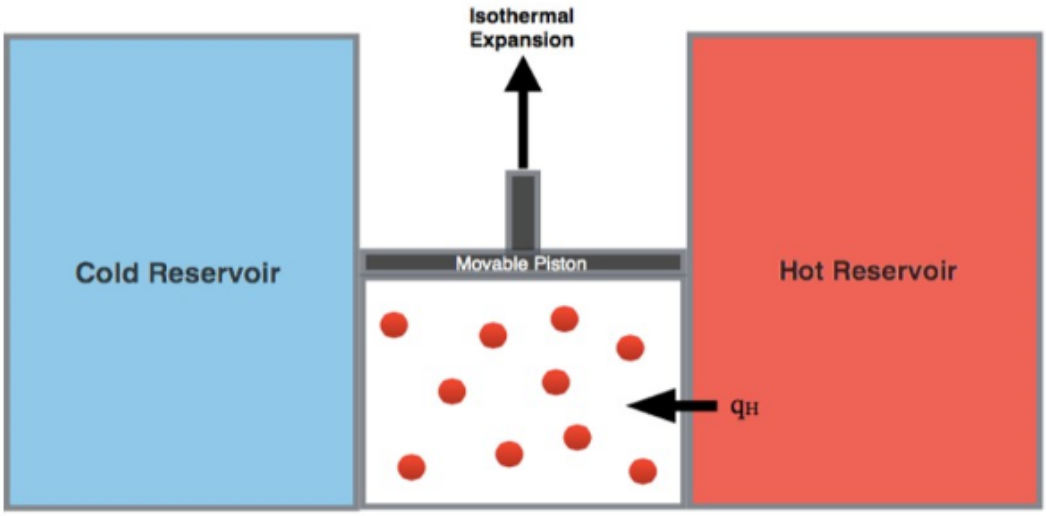

Figure 2.1 – Isothermal expansion of the Carnot cycle.

- Isentropic (reversible and adiabatic) expansion: During this process, the expansion is continued by reducing the pressure of the gas, shown in Figure 2.2. The work is done on the surroundings with rising up the piston. There is no heat transfer between the system and the surroundings because the gas is thermally insulated from the reservoir and the engine is also assumed to be frictionless, leading to the temperature of the gas cooling down to $T_c (< T_h)$.

- Isothermal Compression: Heat is transferred to the cold reservoir with constant temperature T_c. The work is done by the surrounding on the gas with pushing down the piston, leading to the increase of the pressure of the gas, which is shown in Figure 2.3. During this process, even though the pressure of the gas goes up, the temperature

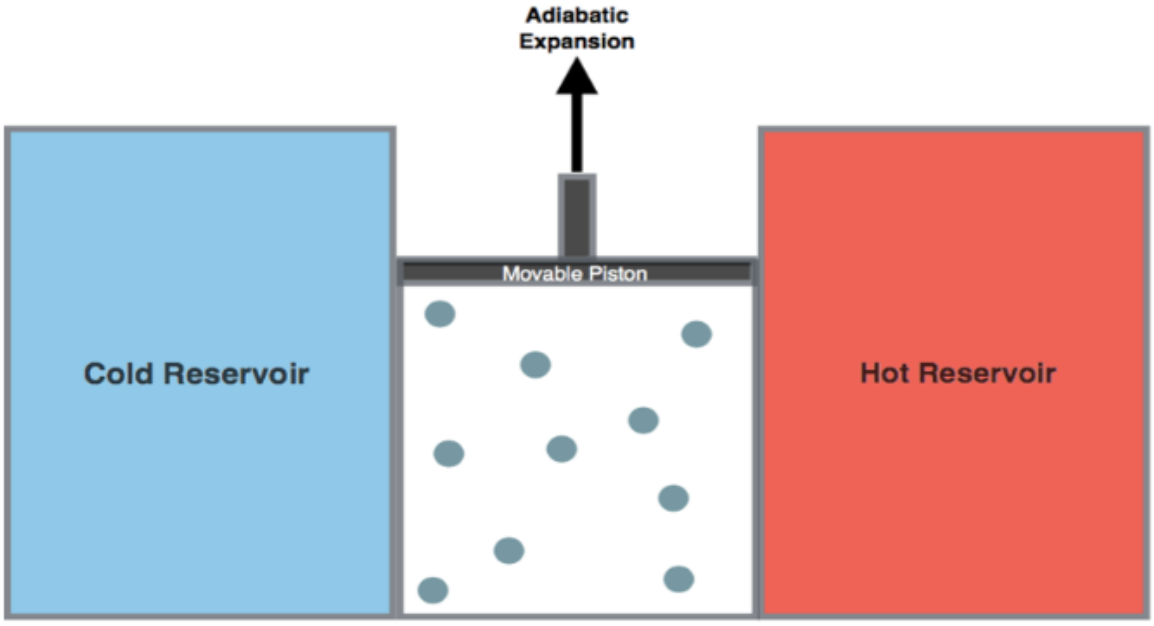

Figure 2.2 – Isentropic expansion of the Carnot cycle.

inside the piston keeps constant as it is continuously contacted with the cold reservoir, by which the heat Q_c is absorbed, leading to the decreased entropy of the gas $\Delta S = \frac{Q_c}{T_c}$.

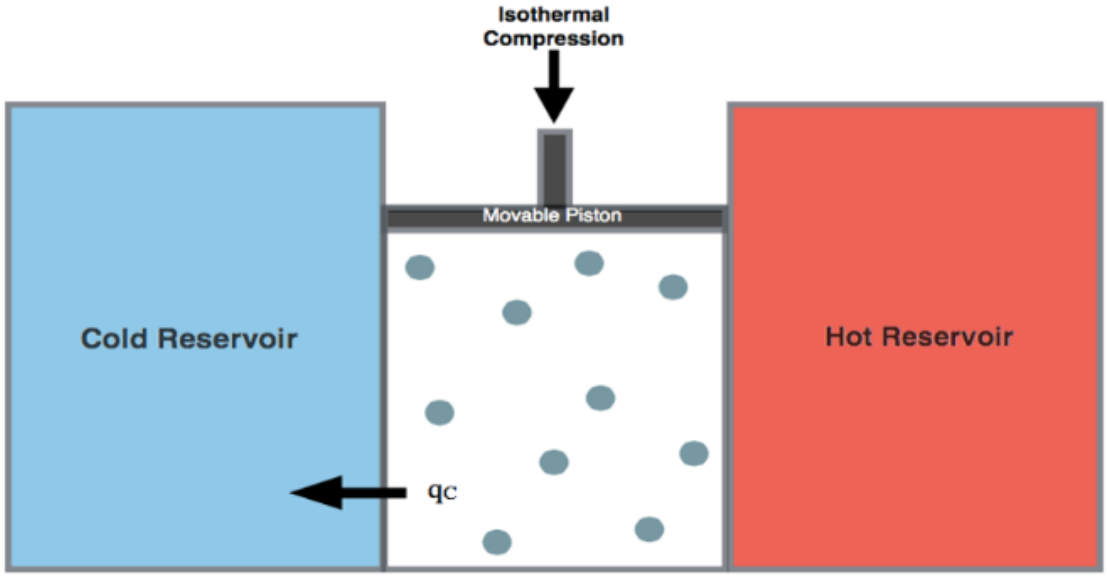

Figure 2.3 – Isothermal compression of the Carnot cycle.

- Isentropic (reversible and adiabatic) compression: During this process, the compression is continued by pushing down the piston, the pressure of the gas is increased, shown in Figure 2.4. The work is done by the surroundings to the gas by pushing down the piston. There is no heat transfer between the system and the surroundings because the gas is thermally insulated from the reservoir and the engine is also assumed to be frictionless, leading to the temperature of the gas heating up to T_1.

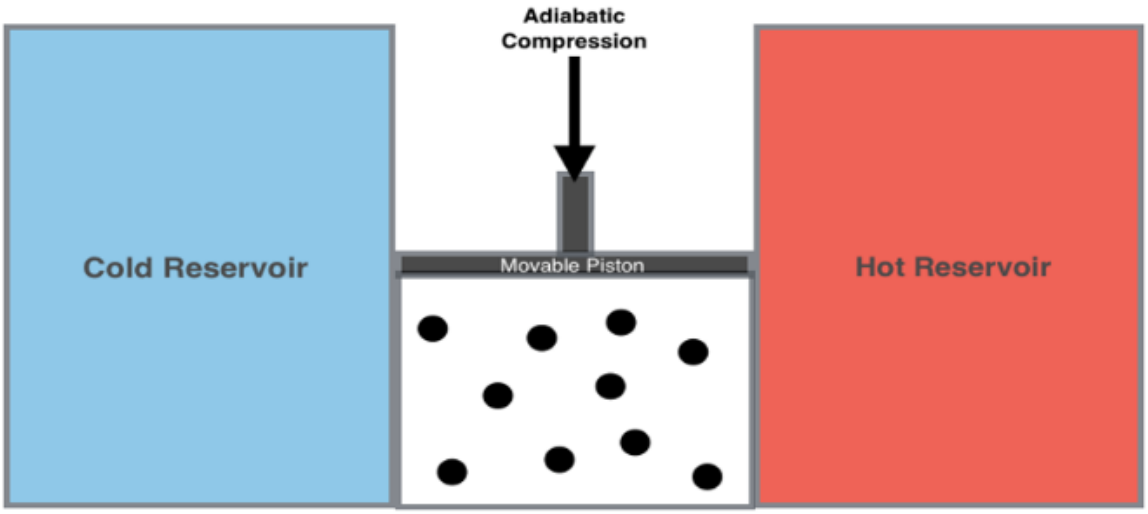

Figure 2.4 – Isentropic compression of the Carnot cycle.

The pressure-volume (P-V) graph for the Carnot cycle is also shown in Figure 2.5.

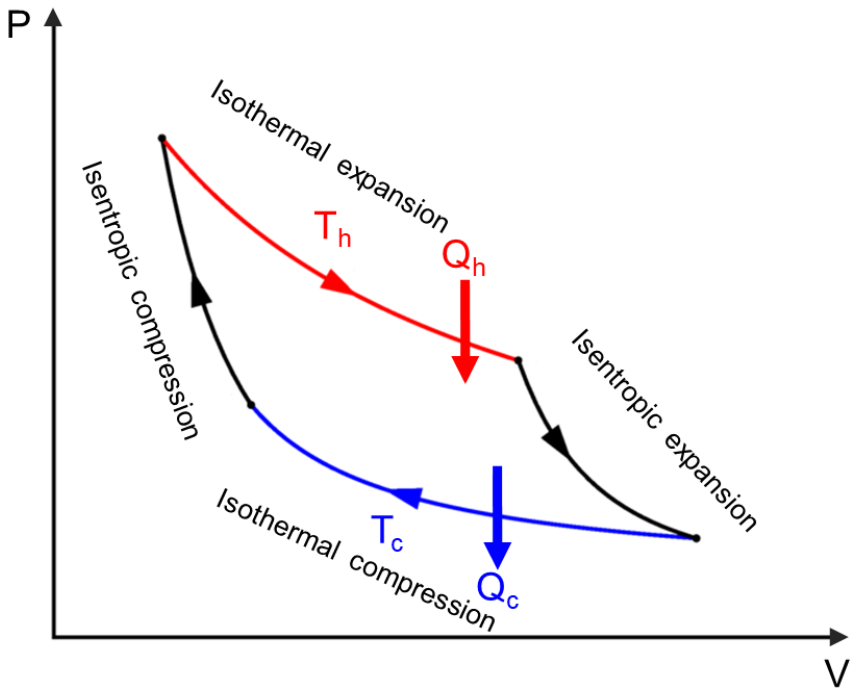

Figure 2.5 – Pressure-volume graph for the Carnot cycle.

The efficiency for the heat engine going through the Carnot cycle is defined as Carnot efficiency given as:

$$\eta_C = 1 - \frac{T_c}{T_h}, \tag{2.3}$$

which gives the maximal efficiency that the heat engine can achieve when it operates between two temperatures T_h and $T_c (< T_h)$. However, in order to achieve the Carnot efficiency, the process should be quasi-static, all the states should be at near-equilibrium, which gives the vanishing power.

2.2 Stochastic thermodynamics

Stochastic thermodynamics provides a new framework to describe small systems like colloids or biomolecules driven far away from equilibrium but still contacted with a heat bath. The laws for classical thermodynamics including the first-law like conservation of energy involving exchanged heat, applied work and entropy production entering refinements of the second law can be meaningfully defined on the level of individual trajectories [9, 18], which yields the respective ensemble quantities after averaging. Various exact relations involving the distribution of such quantities like integral and detailed fluctuation theorems for total entropy production and the Jarzynski relation follow from such an approach based on Langevin dynamics.

2.2.1 Fluctuation theorem

The fluctuation theorem (FT) [40] relates to quantifying the relationship between the probability that the entropy increases and decreases during a fixed amount of time for a system currently not at thermodynamic equilibrium. From the second law of thermodynamics, we know the entropy of an isolated system should keep increasing until it reaches the equilibrium. However, we realized that the second law is only a statistical conclusion after the discovery of statistical mechanics, indicating that there may exist some nonzero probability that the entropy of an isolated system spontaneously decrease, which is exactly the fluctuation theorem quantifies.

The fluctuation theorem states that for systems away from equilibrium during a finite interval $[0, t]$, the ratio between the probability for $\bar{\Sigma}_t$ taking on a value A and the opposite one $-A$ will be exponential in At with $\bar{\Sigma}_t$ defining the probability distribution of the time-averaged

irreversible entropy production

$$\frac{\Pr(\bar{\Sigma}_t = A)}{\Pr(\bar{\Sigma}_t = -A)} = e^{At}. \tag{2.4}$$

The fluctuation theorem provides a clear mathematical expression for the probability that entropy flows in a direction that is opposite to the one dictated by the second law of thermodynamics.

2.2.2 Jarzynski equality

In thermodynamics, the work done on the system W is connected to the free energy difference ΔF between two states through the following inequality

$$W \geq \Delta F, \tag{2.5}$$

where the equality holds only for the quasi-static process. Jarzynski equality [41, 42] sheds light on the equality relation between the free energy difference and the precise amount of work required for finite-time transitions,

$$e^{-\frac{\Delta F}{k_B T}} = \mathbb{E}\left\{e^{-\frac{W}{k_B T}}\right\}, \tag{2.6}$$

with $\mathbb{E}\{.\}$ denotes the expectation on the path space of system trajectories, k_B is the Boltzmann constant, and T represents the temperature of the heat bath. It keeps valid no matter how fast the process is.

Chapter 3

Maximal power output of a stochastic Carnot-like thermodynamic engine

3.1 Introduction

The present work focuses on maximizing power in general, relaxing the assumption of Gaussian boundary distributions and quadratic potential, within the context of stochastic thermodynamics [10, 11], which is a *stochastic control problem*. Our analysis is based on an overdamped Langevin model for thermodynamic processes (with damping coefficient γ), and explores advantages and pitfalls of selecting arbitrary control input, i.e., confining potential, for steering thermodynamic ensembles through cyclic operation while alternating contact between available heat reservoirs. It is noted that without physically motivated constraints on the actuation potential, the power output can become unbounded. The salient feature of actuation (time-varying potential $U(t, x)$, with t denoting time and $x \in \mathbb{R}^d$ the spacial coordinate) that draws increasing amounts of power is its ability to drive the thermodynamic ensemble to a state of very low entropy. Indeed, the magnitude of the spatial gradient of the

potential $\nabla_x U(t,x)$ plays a key role. Thus, it is reasonable on physical grounds to suitably constrain this mode of "control" actuation, that is responsible for energy exchange between the ensemble and the environment. The present work puts forth and motivates the bound[1] (equation (3.61))

$$\frac{1}{\gamma}\int_{\mathbb{R}^d} \|\nabla_x U(t,x)\|^2 \rho(t,x)\,\mathrm{d}x \leq M,$$

where ρ denotes the thermodynamic state, as a suitable such constraint, and under this assumption it is shown that a maximal amount of power output that can be extracted by cyclic operation of a Carnot-like engine is

$$\frac{M}{8}(\frac{T_h}{T_c}-1)\left(\frac{\frac{T_h}{T_c}-1}{\frac{T_h}{T_c}+1}\right) \leq P_{\max} \leq \frac{M}{8}(\frac{T_h}{T_c}-1).$$

That is, the upper bound $\frac{M}{8}(\frac{T_h}{T_c}-1)$ on power output only depends on M and the temperature of the two heat baths[2]. Moreover, this bound can be attained within a factor of $(\frac{T_h}{T_c}-1)/(\frac{T_h}{T_c}+1)$, which depends only on the ratio of temperatures of the two heat baths as well.

The exposition proceeds as follows. Section 3.2 details the stochastic model, thermodynamic ensembles and the heat/energy exchange mechanism. Section 3.3 is a brief overview of optimal transport theory, on which the main results are based. Section 3.4 explores a connection between the second law of thermodynamics and the Wasserstein geometry of optimal mass transport that underlies the mechanism of energy dissipation in thermodynamic transitions. Section 3.5 returns to the concept of a cyclicly operated thermodynamic engine and expresses the optimal efficiency and power output as functions of the operating protocol (solution of a stochastic control problem that dictates the choice of control time-varying potential), tem-

[1]Interestingly, this can also be expressed in information theoretic terms, as a bound on the Fisher information of thermodynamic states.

[2]In general power output is an *extensive* quantity, as it depends on the size of the thermodynamic ensemble/engine. However, in our treatment, the ensemble is described by a probability distribution (normalized). Hence, the bounds appear as "intensive."

perature of heat reservoirs, timing of the cyclic operation, and thermodynamic states at the end of phases of the Carnot-like cycle. Section 3.6 contains the main results regarding seeking maximal power output. Specifically, Section 3.6.1 explains optimal scheduling times, Section 3.6.2 highlights questions that arise based on physical grounds for Gaussian thermodynamic states, Sections 3.6.3 and 3.6.4 discuss optimal thermodynamic states at the two ends of the Carnot-like cycle, and Sections 3.6.5 and 3.6.6 derive bounds on maximal achievable power with or without constraint on the controlling potential. A concluding remarks section recaps and points to future research directions and open problems.

3.2 Stochastic thermodynamic models

We begin by describing the basic model for a *thermodynamic ensemble* used in this work. This consists of a large collection of Brownian particles that interact with a *heat bath* in the form of a stochastic excitation and driven under the influence of an *external (time varying) potential* between end-point states. The dynamics of individual particles are expressed in the form of stochastic differential equations.

3.2.1 Langevin dynamics

The (under-damped) Langevin equations

$$\mathrm{d}X_t = \frac{p_t}{m}\mathrm{d}t \tag{3.1a}$$

$$\mathrm{d}p_t = -\nabla_x U(t, X_t)\mathrm{d}t - \frac{\gamma p_t}{m}\mathrm{d}t + \sqrt{2\gamma k_B T(t)}\mathrm{d}B_t, \tag{3.1b}$$

represent a standard model for molecular systems interacting with a thermal environment. Throughout, $X_t \in \mathbb{R}^d$ denotes the location of a particle and p_t denotes its momentum at

time t, $U(t,x)$ denotes a time-varying potential for $x \in \mathbb{R}^d$, m is the mass of the particle, γ is the viscosity coefficient, k_B is the Boltzmann constant, $T(t)$ denotes the temperature of the heat bath at time t, and B_t denotes a standard $\mathbb{R}^d$-valued Brownian motion.

In this paper, we consider only the case where inertial effects in the Langevin equation (3.1b) are negligible for the time resolution of interest. Specifically, for temporal resolution $\Delta t \gg \frac{m}{\gamma}$ and small particle size, the dynamics reduce to the *over-damped Langevin equation*

$$\mathrm{d}X_t = -\frac{1}{\gamma}\nabla_x U(t, X_t)\mathrm{d}t + \sqrt{\frac{2k_B T(t)}{\gamma}}\mathrm{d}B_t. \tag{3.2}$$

Intuitively, equation (3.2) is obtained from (3.1b) by setting $\mathrm{d}p_t = 0$ and replacing $\frac{p_t}{m}\mathrm{d}t = \mathrm{d}X_t$. For a more detailed explanation see [10, page 20].

Thus, we view $\{X_t\}_{t\geq 0}$ as a diffusion process. The state of the thermodynamic ensemble is identified with the probability density of X_t, denoted by $\rho(t,x)$, which satisfies the Fokker-Planck equation

$$\frac{\partial \rho}{\partial t} - \frac{1}{\gamma}\nabla_x \cdot [(\nabla_x U + k_B T \nabla_x \log \rho)\rho\,] = 0. \tag{3.3}$$

Remark 3.2.1. *The under-damped Langevin equation* (3.1) *is the most common dynamical model for a particle immersed in a heat bath [10]; alternative models can be based on e.g., a Poisson process for the thermal excitation, space-dependent viscosity coefficient, and possibly nonlinear effects of an interaction potential. It has been used to model e.g., colloidal particles in a laser trap, enzymes and molecular motors in single molecule assays, and so on [11]. In the present work, we follow recent literature [13, 35, 43–45] where, besides Brownian excitation, the focus on constant viscosity coefficient γ.*

3.2.2 Heat, work, and the first law

The evolution of the thermodynamic ensemble under the influence of the time-varying thermal environment and the time-varying potential $U(t, x)$, leads to exchange of heat and work, respectively. Heat and work can be defined at the level of a single particle as explained below.

The energy exchange between an individual particle and the thermal environment represents *heat*. This exchange is effected by forces exerted on the particle due to viscosity ($-\gamma\frac{\mathrm{d}X_t}{\mathrm{d}t}$) and due to the random thermal excitation ($\sqrt{2\gamma k_B T}\frac{\mathrm{d}B_t}{\mathrm{d}t}$). It is formally expressed as the product of force and displacement

$$(-\gamma\frac{\mathrm{d}X_t}{\mathrm{d}t} + \sqrt{2\gamma k_B T}\frac{\mathrm{d}B_t}{\mathrm{d}t}) \circ \mathrm{d}X_t$$

in Stratonovich form. Using (3.2), formally,

$$-\gamma\frac{\mathrm{d}X_t}{\mathrm{d}t} + \sqrt{2\gamma k_B T}\frac{\mathrm{d}B_t}{\mathrm{d}t} = \nabla_x U(t, X_t),$$

which leads to the expression $đQ = \nabla_x U(t, X_t) \circ \mathrm{d}X_t$ for the heat; see [10, Chapter 4.1] for a more detailed exposition. Then, bringing in the Itô correction, we arrive at

$$\begin{aligned} đQ = -\frac{1}{\gamma}\|\nabla_x U(t, X_t)\|^2 \mathrm{d}t + \Delta_x U(t, X_t)\frac{k_B T(t)}{\gamma}\mathrm{d}t \\ + \nabla_x U(t, X_t)\sqrt{\frac{2k_B T(t)}{\gamma}}\mathrm{d}B_t. \end{aligned}$$

Note that we use $đ$, as in the case of not-perfect differentials, to emphasize that $\int đQ$ depends on the path and not just on end-point conditions.

The *work* transferred to the particle by a change in the actuating potential is taken as[3]

$$đW = \frac{\partial U}{\partial t}(t, X_t)\mathrm{d}t. \tag{3.4}$$

Thence, since the internal energy is simply the value of the potential, the *first law of thermodynamics*, $\mathrm{d}U(t, X_t) = đQ + đW$, holds.

Accordingly, for a thermodynamic *ensemble* at a state $\rho(t, x)$, the heat and work differentials are

$$đ\mathcal{Q} = \left[\int_{\mathbb{R}^d} \left(-\frac{1}{\gamma}\|\nabla_x U\|^2 + \Delta_x U \frac{k_B T}{\gamma}\right) \rho \,\mathrm{d}x\right] \mathrm{d}t \tag{3.5a}$$

$$đ\mathcal{W} = \left[\int_{\mathbb{R}^d} \frac{\partial U}{\partial t} \rho \,\mathrm{d}x\right] \mathrm{d}t, \tag{3.5b}$$

leading to the first law for the ensemble $\mathrm{d}\mathcal{E}(\rho, U) = đ\mathcal{Q} + đ\mathcal{W}$, where the internal energy is

$$\mathcal{E}(\rho, U) = \int_{\mathbb{R}^d} U\rho \,\mathrm{d}x, \tag{3.5c}$$

and depends on ρ, U, whereas $\mathcal{Q}, \mathcal{W}$ depend on the path.

3.2.3 Summary notation

As usual, $\mathbb{R}^d$ denotes the d-dimensional Euclidean space, for $d \in \mathbb{N}$, with $\langle x, y\rangle$ and $\|x\| = \sqrt{\langle x, x\rangle}$ denoting the respective inner product and norm, for $x, y \in \mathbb{R}^d$. Throughout the chapter, the stochastic differential equations are stated in Itô form, unless the Stratonovich integration notation $\circ$ is used explicitly. The Gaussian distribution with mean m and covariance Σ is denoted by $N(m, \Sigma)$. For convenience we provide the following Table of the various quantities, including the corresponding units in SI format: Newton (N), seconds (s),

[3]This particular formula for the work has been the subject of considerable debate [46], [47], [48], [49].

Table 3.1 – Symbols and corresponding units

Definition	Notation	Units
time	t	s
position of particle	X_t	m
Boltzmann constant	k_B	Nm
damping coefficient	γ	Ns/m
potential	$U(t,x)$	Nm
temperature	T	^{o}K
Brownian motion	B_t	$s^{\frac{1}{2}}$
density in $\mathbb{R}^d$	$\rho(t,x)$	m^{-d}
velocity field in $\mathbb{R}^d$	$v(t,x)$	m/s
Wasserstein metric, length	$\mathrm{W}_2(\cdot,\cdot)$, $\ell_{\rho[t_i,t_f]}$	m
entropy	$\mathcal{S}(\rho)$	Nm
work (particle/ensemble)	$W, \mathcal{W}$	Nm
heat (particle/ensemble)	$Q, \mathcal{Q}$	Nm
energy (particle/ensemble)	$U, \mathcal{E}$	Nm
free energy	$\mathcal{F}$	Nm
bound in (3.61)	M	Nm/s
power	P	Nm/s

meter (m), absolute temperature in degrees Kelvin (^{o}K).

3.3 A brief excursion into optimal mass transport

As it turns out, dissipation in Langevin models (3.2) is closely linked to the path that a thermodynamic ensemble traverses. This path is seen as a trajectory in the space probability distributions and its length, that quantifies dissipation, is metrized by the so-called Wasserstein metric. Thus, we now embark on a brief excursion into the basics of optimal mass transport so as to provide context for needed results in Wasserstein geometry –the pertinence of the Wasserstein metric to thermodynamics has been recognized in [11, 50–54].

We denote by $\mathcal{P}_2(\mathbb{R}^d)$ the space of probability distributions with finite second-order moment.

We utilize the notation $\mathrm{d}\mu$ to signify a probability measure while we write $\mathrm{d}\mu(x) = \rho(x)dx$ to signify that $\mathrm{d}\mu$ is absolutely continuous with respect to the Lebesgue measure $\mathrm{d}x$ with ρ the corresponding probability density.

For $\mathrm{d}\mu_0, \mathrm{d}\mu_1 \in \mathcal{P}_2(\mathbb{R}^d)$, the 2-Wasserstein distance

$$\mathrm{W}_2(\mu_0, \mu_1) := \sqrt{\inf_{\pi \in \Pi(\mu_0, \mu_1)} \int_{\mathbb{R}^d \times \mathbb{R}^d} \|x - y\|^2 \mathrm{d}\pi(x, y)}, \tag{3.6}$$

where $\Pi(\mu_0, \mu_1)$ denotes the set of probability measures on the product space $\mathbb{R}^d \times \mathbb{R}^d$ with μ_0, μ_1 as marginals, is a *bona fide* metric on the space of distributions. This metric, in fact, induces a Riemannian-like structure as we explain below.

The expression $\int \|x - y\|^2 \mathrm{d}\pi(x, y)$ above is a relaxation of the transportation cost

$$\int_{\mathbb{R}^d} \|x - \Psi(x)\|^2 \mathrm{d}\mu_0(x)$$

in *Monge's* problem [55] to be minimized over maps Ψ that transfer the "mass" distribution μ_0 into μ_1, i.e., such that $\int_A d\mu_1 = \int_{\Psi^{-1}(A)} d\mu_0$ over measurable sets A. This relation is denoted by $\Psi \sharp \mu_0 = \mu_1$. In case the two measures admit densities, it can be expressed via the change of variables formula

$$\det(\nabla_x \Psi(x)) \rho_1(\Psi(x)) = \rho_0(x),$$

for the respective ρ_i's ($i \in \{0, 1\}$). In fact, in this case where both measures admit densities, the support of the optimal π in (the convex problem) (3.6) coincides with the graph of the unique minimizing map $\Psi : \mathbb{R}^d \to \mathbb{R}^d$ for Monge's problem. Further, the optimal Ψ is the gradient of a convex function ψ on $\mathbb{R}^d$ [55, Ch.5], i.e., $\Psi = \nabla_x \psi$. Interesting, being a gradient "vector field," Ψ is curl-free, which in itself characterizes optimality.

We now sketch how $\mathcal{P}_2(\mathbb{R}^d)$ can be equipped with a Riemannian-like structure, while we refer

to [56] for a rigorous exposition. For brevity and for notational convenience, we are only concerned with distributions that admit densities and use the simplified notation $\rho \in \mathcal{P}_2(\mathbb{R}^d)$.

Consider an "infinitessimal" perturbation $\rho + \delta \in \mathcal{P}_2(\mathbb{R}^d)$ and the solution ϕ to the Poisson equation

$$\nabla_x \cdot (\rho \nabla_x \phi) = -\delta.$$

The map $\Psi = \mathrm{Id} + \nabla_x \phi$, where Id denotes the identity map, optimally transports ρ into $\rho + \delta$ as $(\mathrm{Id} + \nabla_x \phi) \sharp \rho \simeq \rho + \delta$. Alternatively, $v = \nabla_x \phi$ can be viewed as a velocity field effecting transport as $\rho - \nabla_x \cdot (\rho v) = \rho + \delta$.

Thus, the correspondence $\delta \to \nabla_x \phi$ identifies tangent directions δ on $\mathcal{P}_2(\mathbb{R}^d)$, i.e., rates of change $\frac{\partial \rho}{\partial t} = \delta$ about a given density ρ, with an (optimal) corresponding velocity field $\nabla_x \phi$. Hence, it is natural to consider the (twice) average "kinetic energy" to define a metric on tangent directions on $\mathcal{P}_2(\mathbb{R}^d)$. Specifically, if $\delta_i = \frac{\partial \rho}{\partial t_i}$, for $i \in \{1, 2\}$, represent two tangent directions at ρ, we define the inner product

$$\langle \frac{\partial \rho}{\partial t_1}, \frac{\partial \rho}{\partial t_2} \rangle_{\mathrm{W}} := \int_{\mathbb{R}^d} \langle \nabla_x \phi_1, \nabla_x \phi_2 \rangle \rho \, \mathrm{d}x, \tag{3.7}$$

where the ϕ_i's solve $\nabla_x \cdot (\rho \nabla_x \phi_i) = -\frac{\partial \rho}{\partial t_i}$. The associated norm is

$$\| \frac{\partial \rho}{\partial t} \|_{\mathrm{W}} := \sqrt{\langle \frac{\partial \rho}{\partial t}, \frac{\partial \rho}{\partial t} \rangle_{\mathrm{W}}}.$$

Consider $\rho_{[t_i, t_f]} := \{\rho(t, \cdot) \in \mathcal{P}_2(\mathbb{R}^d) | t \in [t_i, t_f]\}$ as a curve (path) in $\mathcal{P}_2(\mathbb{R}^d)$. Two quantities of interest are its length,

$$\ell_{\rho_{[t_i, t_f]}} := \int_{t_i}^{t_f} \| \frac{\partial \rho}{\partial t} \|_{\mathrm{W}} \mathrm{d}t, \tag{3.8}$$

and the *kinetic energy* integral (*action*) along the path

$$\mathcal{A}_{\rho_{[t_i,t_f]}} := \int_{t_i}^{t_f} \|\frac{\partial \rho}{\partial t}\|_{\mathrm{W}}^2 \mathrm{d}t \tag{3.9}$$

(modulo a factor of $\frac{1}{2}$). It can be seen that

$$\ell_{\rho_{[t_i,t_f]}} = \min \sqrt{(t_f - t_i)\mathcal{A}_{\rho_{[t_i,t_f]}}},$$

over time-parametrizations of the path, with the minimum corresponding to constant velocity. Moreover, the minimal path-length between two end-points ρ_{t_i} and ρ_{t_f} turns out to be precisely $\mathrm{W}_2(\rho_{t_i}, \rho_{t_f})$, and thus, $\mathcal{P}_2(\mathbb{R}^d)$ is a length space, [56], [55, Chapter 8].

We conclude with an important inequality linking the Wasserstein metric to information functionals. Consider a reference probability distribution $\mathrm{d}\mathfrak{m} = e^{-V}\mathrm{d}x \in \mathcal{P}_2(\mathbb{R}^d)$, with $V(x)$ having Hessian $\nabla_x^2 V \geq \kappa I$ for $\kappa \in \mathbb{R}$, and $d\mu = \rho d\mathfrak{m}$ also in $\mathcal{P}_2(\mathbb{R}^d)$. The *relative entropy* and *Fisher information* functionals, respectively, are defined by

$$H(\mu|\mathfrak{m}) := \int_{\mathbb{R}^d} \rho \log(\rho)\, \mathrm{d}\mathfrak{m}, \tag{3.10a}$$

$$I(\mu|\mathfrak{m}) := \int_{\mathbb{R}^d} \|\nabla_x \log(\rho)\|^2 \rho\, \mathrm{d}\mathfrak{m}. \tag{3.10b}$$

These are linked to the Wasserstein distance via the following HWI* inequality [57, 58],

$$H(\mu_1|\mathfrak{m}) - H(\mu_2|\mathfrak{m}) \leq \mathrm{W}_2(\mu_1,\mu_2)\sqrt{I(\mu_1|\mathfrak{m})} - \frac{\kappa^2}{2}\mathrm{W}_2^2(\mu_1,\mu_2), \quad \forall \mu_1,\mu_2 \in \mathcal{P}_2(\mathbb{R}^d). \tag{3.11}$$

3.4 The second law, dissipation, and Wasserstein geometry

Next, we discuss the *second law of thermodynamics* in the context of an ensemble of particles obeying over-damped Langevin dynamics (3.2) for a heat bath with constant temperature $T(t) = T$. The classical formulation of the law amounts to the inequality

$$\mathcal{W} - \Delta\mathcal{F} \geq 0, \tag{3.12}$$

where $\mathcal{W} = \int_{t_i}^{t_f} đ\mathcal{W}$ is the work transferred to the ensemble over a time interval (t_i, t_f), and $\Delta\mathcal{F}$ is the change in the free energy[4]

$$\mathcal{F}(\rho, U) = \mathcal{E}(\rho, U) - T\mathcal{S}(\rho) \tag{3.13}$$

between the two end-point states, see [12, 59]. Here,

$$\mathcal{S}(\rho) = -k_B \int_{\mathbb{R}^d} \log(\rho)\, \rho \, \mathrm{d}x \tag{3.14}$$

denotes the entropy of the state ρ, and U the potential. Inequality (3.12) becomes equality for quasi-static (reversible) thermodynamic transitions. In general, for irreversible transitions, the gap in (3.12) quantifies dissipation. Interestingly, alternative formulations that shed light into irreversible transitions have recently been discovered. A most remarkable identity was discovered by Jarzynski in the late 90's [41] to hold for irreversible thermodynamic

[4]The free energy represents the amount of energy that can be delivered at temperature T with fixed potential U. However, a rather revealing re-write of the free energy is as the relative entropy (KL-divergence) between the current state ρ and the Gibbs distribution $\rho_{\text{Gibbs}}(x) = e^{-\beta U(x)}/Z$, with $\beta = 1/k_B T$ and $Z = \int_{\mathbb{R}^d} e^{-\beta U(x)} \mathrm{d}x$ the partition function. Specifically, $\mathcal{F}(\rho, U) = \beta^{-1} \int_{\mathbb{R}^d} \log(\frac{\rho(x)}{\rho_{\text{Gibbs}}(x)})\rho(x)\mathrm{d}x - \beta^{-1}\log(Z)$.

transitions between work and free energy, in the form,

$$\mathbb{E}\left\{e^{-\beta W}\right\} - e^{-\beta \Delta \mathcal{F}_{eq}} = 0,$$

where the expectation is taken over the probability law on paths, $W = \int đW$ represents the work along trajectories of individual particles, and $\Delta \mathcal{F}_{eq} = -\beta^{-1} \log(\frac{Z_{t_f}}{Z_{t_i}})$ signifies the difference of the equilibrium free energy $-\beta^{-1} \log(Z_t)$ at the two end-points in time $t \in \{t_i, t_f\}$. Here, $Z_t = \int_{\mathbb{R}^d} e^{-\beta U(t,x)} \mathrm{d}x$ where, as usual, $\beta = 1/k_B T$. In Jarzynski's original derivation [41, 60] of the Jarzynski equality, the notions of work and heat are in alignment with the ones used in this paper, though [60] considers more general stochastic dynamics satisfying one type of detailed balance condition [60, Section 1]. Interestingly, the Jarzynski equality holds even for an alternative notion of work, see e.g., [61].

While the Jarzynski relation establishes equality between the above functional of the work and free energy differences, it does not allow quantifying the actual expected work performed on the ensemble. An alternative identity that quantifies explicitly the gap in (3.12) holds for irreversible thermodynamic transitions. This identity is (cf. Theorem 3.1)

$$\mathcal{W} - \Delta \mathcal{F} = \underbrace{\gamma \int_{t_i}^{t_f} \left\| \frac{\partial \rho}{\partial t} \right\|_{\mathrm{W}}^2 \mathrm{d}t}_{\text{dissipation}}, \tag{3.15}$$

which is γ times $\mathcal{A}_{\rho_{[t_i,t_f]}}$, the action integral along the time-parametrized path traversed.

$$\mathcal{W} - \Delta \mathcal{F} = \frac{\gamma}{t_f - t_i} \mathrm{W}_2(\rho_{t_i}, \rho_{t_f})^2 \tag{3.16}$$

quantifies the least amount of work needed for transition between specified end-point thermodynamic states, or the maximal work that can be drawn. We recap the key points below.

Theorem 3.1. *Consider thermodynamic transitions between states ρ_{t_i}, ρ_{t_f}, under constant temperature T and a time-varying potential U for the overdamped Langevin model* (3.2).

Then,

$$\mathcal{W} - \Delta\mathcal{F} \geq \frac{\gamma}{t_f - t_i} \ell^2_{\rho_{[t_i,t_f]}}. \tag{3.17}$$

Relation (3.17) *holds with equality for a path of the thermodynamic ensemble chosen to be a constant speed* W_2*-geodesic, effected by a suitable potential, a choice that corresponds to minimal dissipation.*

Proof. We first derive (3.15), cf. [62, 63] for similar computations with time independent potential. Consider

$$\begin{aligned}\frac{\mathrm{d}\mathcal{F}}{\mathrm{d}t}(\rho, U) &= \frac{\mathrm{d}}{\mathrm{d}t}\mathcal{E}(\rho, U) - T\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{S}(\rho) \\ &= \int_{\mathbb{R}^d} \frac{\partial U}{\partial t}\rho dx + \int_{\mathbb{R}^d} (U + k_B T(1 + \log\rho)) \frac{\partial \rho}{\partial t} dx.\end{aligned}$$

Using the Fokker-Planck equation (3.3), the second term

$$\begin{aligned}&\int_{\mathbb{R}^d} (U + k_B T(1 + \log\rho)) \frac{1}{\gamma} \nabla_x \cdot [(\nabla_x U + k_B T \nabla_x \log\rho)\rho] \,\mathrm{d}x \\ =& -\frac{1}{\gamma}\int_{\mathbb{R}^d} \|\nabla_x U + k_B T \nabla_x \log\rho\|^2 \rho \,\mathrm{d}x \\ =& -\gamma \int_{\mathbb{R}^d} \|v\|^2 \rho \,\mathrm{d}x,\end{aligned}$$

where the first equality follows using integration by parts (under standard assumptions on the decay rate of ρ at infinity), while the second equality is a re-write using[5]

$$v := -\frac{1}{\gamma}(\nabla_x U + k_B T \nabla_x \log\rho). \tag{3.18}$$

[5]We note that v is known as Nelson's current velocity [64].

Thus, $\frac{\mathrm{d}\mathcal{F}}{\mathrm{d}t}(\rho, U) = \int_{\mathbb{R}^d} \frac{\partial U}{\partial t} \rho \, \mathrm{d}x - \gamma \int_{\mathbb{R}^d} \|v\|^2 \rho \, \mathrm{d}x$. Integrating over $[t_i, t_f]$ yields

$$\Delta \mathcal{F} = \mathcal{W} - \gamma \int_{t_i}^{t_f} \int_{\mathbb{R}^d} \|v\|^2 \rho \, \mathrm{d}x \, \mathrm{d}t, \tag{3.19}$$

where v is the gradient of $\phi = -\frac{1}{\gamma}(U + k_B T \log \rho)$ and satisfies the continuity equation $\nabla \cdot (\rho \nabla \phi) = \frac{\partial \rho}{\partial t}$ as claimed. This establishes (3.15).

The inequality (3.17) follows from the fact that the W_2-length of the path $\rho_{[t_i,t_f]}$ (i.e., as a curve in $\mathcal{P}_2$), is given by (3.8). Specifically, provided $\int_{\mathbb{R}^d} \|v\|^2 \rho \, \mathrm{d}x = \alpha^2$ remains constant along the path (i.e., for $t \in [t_i, t_f]$),

$$\alpha = \frac{1}{t_f - t_i} \ell_{\rho_{[t_i,t_f]}}.$$

and the claim follows. If on the other hand the kinetic energy varies with time, then the path $\rho(t, \cdot)$, time-reparametrized by

$$\tilde{t}(t) := \frac{\ell_{\rho_{[t_i,t]}}}{\ell_{\rho_{[t_i,t_f]}}} (t_f - t_i) + t_i$$

will be traversed via a velocity field

$$\tilde{v}(\tilde{t}(t)) = \frac{v(t)}{\|v(t)\|_\rho} \frac{\ell_{\rho_{[t_i,t_f]}}}{t_f - t_i}.$$

Knowing $\tilde{v}$, a new potential $\tilde{U}$ can be computed so that $\tilde{v}(\tilde{t}, \cdot) = \nabla_x \tilde{U}(\tilde{t}, \cdot) + k_B T \nabla_x \log(\rho(\tilde{t}, \cdot))$. Finally, equality in (3.17) holds when taking $\rho_{[t_i,t_f]}$ to be a geodesic [65]. □

Remark 3.4.1. *Early work by Jordan et al. [50], pointing out that the gradient flow of the free energy in* W_2 *is the Fokker-Planck equation, set the stage for understanding the role of the Wasserstein geometry in quantifying dissipation. This fact was recognized in [11, 51, 52] and more recently developed in [53, 54].*

3.5 Cyclic operation of engines

We consider two types of thermodynamic transitions, isothermal and adiabatic. The first corresponds to a situation where the system remains in contact with a heat bath of constant temperature T while a time-varying potential steers its thermodynamic state $\rho(t,.)$ from an *initial* $\rho(t_i,\cdot)$ to a *final* $\rho(t_f,\cdot)$. The adiabatic transition amounts to abrupt changes in both, the temperature of the heat bath as well as the shape of the potential, that are fast enough not to have any measurable effect on the state $\rho(t,.)$ and, as a consequence, to the entropy of the ensemble. We evaluate next the energy and work budgets in the corresponding actuation protocols.

3.5.1 Isothermal transition

We consider transition between states ρ_{t_i} and ρ_{t_f} for the ensemble modeled by (3.2), over a time interval $[t_i, t_f]$, under the time-varying potential $U(t, X_t)$ and in contact with a heat bath of temperature T. Using the relationship (3.15) between work, free energy, and the dissipation, and the first law, we have the following identity relating thermodynamic quantities in isothermal transitions

$$\mathcal{W} = \Delta\mathcal{E} - T\Delta\mathcal{S} + \mathcal{W}_{\text{irr}} \tag{3.20a}$$

$$\mathcal{Q} = T\Delta\mathcal{S} - \mathcal{W}_{\text{irr}} \tag{3.20b}$$

with the *irreversible* $\mathcal{W}_{\text{irr}}$ that represents dissipation attaining its minimal value

$$\frac{\gamma}{t_f - t_i}\mathrm{W}_2(\rho_{t_i}, \rho_{t_f})^2 \tag{3.20c}$$

by the choice of actuation $\nabla_x U(t,\cdot)$ in (3.18) with v the optimal velocity field minimizing dissipation in (3.15) (item iii) in Theorem 3.1).

It is important to note that the minimizing v can be obtained by solving a convex reformulation of (3.15) in terms of the density $\rho(t,\cdot)$ and the momentum field $\mathbf{p}(t,\cdot) = v(t,\cdot)\rho(t,\cdot)$, in the form

$$\min_{\mathbf{p}(t,\cdot),\rho(t,\cdot)} \int_{t_i}^{t_f} \int_{\mathbb{R}^d} \frac{\|\mathbf{p}\|^2}{\rho} \mathrm{d}x\mathrm{d}t \tag{3.21a}$$

$$\text{subject to } \frac{\partial \rho}{\partial t} + \nabla_x \cdot \mathbf{p} = 0 \tag{3.21b}$$

$$\text{and } \rho(t_i,\cdot), \rho(t_f,\cdot) \text{ specified.} \tag{3.21c}$$

Then, $v = \mathbf{p}/\rho$, see [66, Section 4] and [55, p. 241].

3.5.2 Adiabatic transition

We now consider transition between ρ_{t_i} and ρ_{t_f} for the ensemble modeled by (3.2), over a time interval $[t_i, t_f]$, under abrupt changes in the potential $U(t,\cdot)$ and the temperature T of the heat bath.

The transition takes place over an infinitesimally short time interval about time t (with t^-/t^+ indicating the left/right limits, respectively). Thus, the temperature T of the heat bath jumps between values $T(t^-)$ and $T(t^+)$ while, at the same time, the controlling potential switches from $U(t^-,\cdot)$ to $U(t^+,\cdot)$.

The energy budget of the transition no longer contains irreversible losses, as the right hand side of (3.15) vanishes. Moreover, the entropy of the ensemble remains constant. Thus, the

work input into the system equal to change in internal energy,

$$\mathcal{W} = \int_{\mathbb{R}^d} (U(t^+, x) - U(t^-, x))\rho(t, x)\mathrm{d}x = \Delta\mathcal{E}, \tag{3.22a}$$

and therefore no heat transfer takes place, and therefore,

$$\mathcal{Q} = 0. \tag{3.22b}$$

3.5.3 Finite-time Carnot cycle

We are now in position to consider a complete *Carnot-like thermodynamic cycle* where the ensemble is steered between two states ρ_a and ρ_b during isothermal expansion (from ρ_a to ρ_b) and contraction (from ρ_b to ρ_a) phases, separated by adiabatic transitions. Periodic operation about such a scheduling is sought as a means to extract work from a heat bath. A schematic in Figure 3.1 depicts the phases of the cyclic operation. These four phases are described in detail next.

1) Isothermal process in temperature T_h ("hot"): The first step is an isothermal expansion over the time interval $(0, t_1)$ in contact with a heat bath of temperature $T = T_h$. Change in the potential steers the ensemble from a starting state ρ_a to a terminal state ρ_b. As in (3.20),

$$\mathcal{W}^{(1)} = \Delta\mathcal{E}^{(1)} - T_h\Delta\mathcal{S}^{(1)} + \mathcal{W}_{\mathrm{irr}}^{(1)} \tag{3.23a}$$

$$\mathcal{Q}^{(1)} = T_h\Delta\mathcal{S}^{(1)} - \mathcal{W}_{\mathrm{irr}}^{(1)} \tag{3.23b}$$

where the superscript enumerates the phase in the cycle, and the minimal work loss $\mathcal{W}_{\mathrm{irr}}^{(1)}$

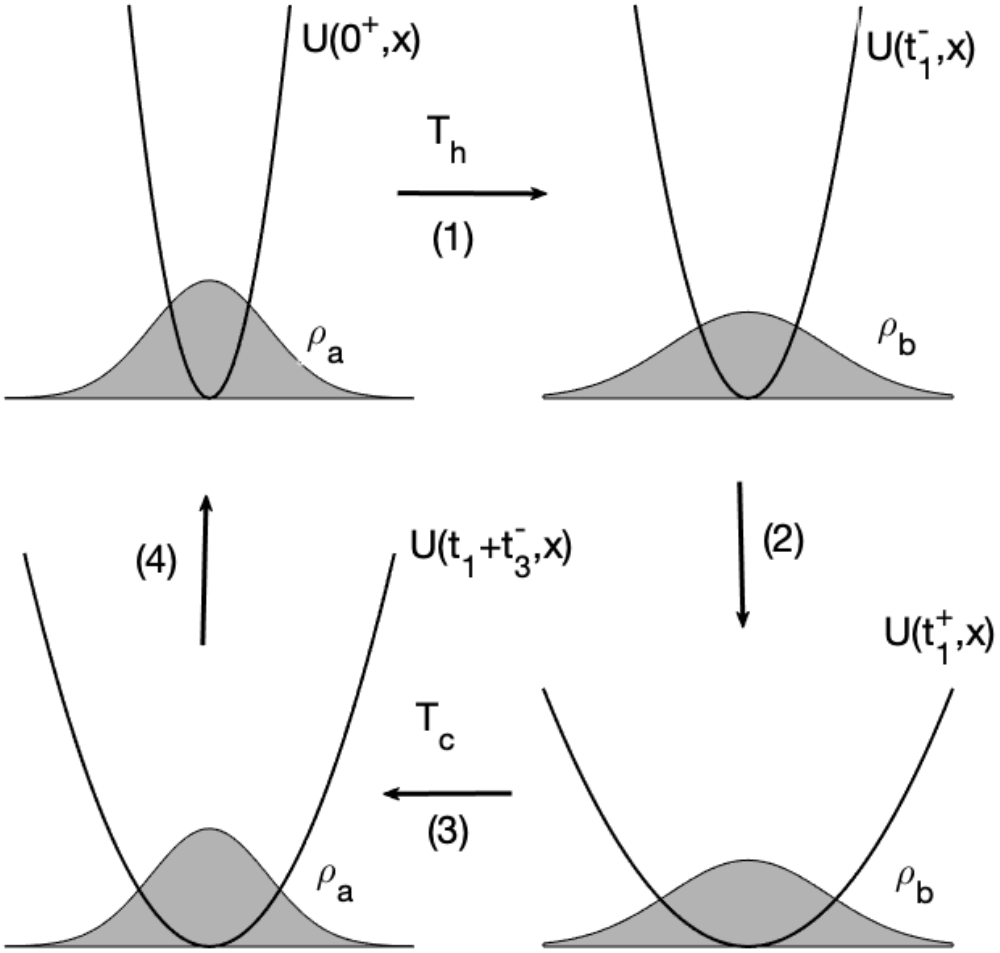

Figure 3.1 – Carnot-like cycle of a stochastic model for a heat engine (with $d = 1$): the operation cycles clockwise through two isothermal transitions (1) and (3), and two adiabatic transitions (2) and (4). During the isothermal transitions having duration t_1 and t_3, the ensemble is in contact with a "hot" reservoir of temperature T_h, and a "cold" one of temperature T_c, respectively. The adiabatic transitions are considered to be instantaneous, i.e., $t_2 = t_4 = 0$. The marginal densities are ρ_a and ρ_b.

depends only on the end-point states as it equals

$$\mathcal{W}_{\mathrm{irr}}^{(1)} = \frac{\gamma}{t_1}\mathrm{W}_2(\rho_a, \rho_b)^2. \tag{3.23c}$$

2) Adiabatic process: The second phase of the cycle is an adiabatic transition at time $t = t_1$, over an infinitesimal interval (of duration"$t_2 = 0$"), bringing the ensemble in contact with a heat bath of temperature T_c ("cold"). As in (3.22),

$$\mathcal{W}^{(2)} = \Delta\mathcal{E}^{(2)} \tag{3.24a}$$

$$\mathcal{Q}^{(2)} = 0 \tag{3.24b}$$

while the state remains at ρ_b.

3) Isothermal process in temperature T_c ("cold"): The third step is an Isothermal contraction over the time interval (t_1, t_1+t_3) while in contact with a heat bath of temperature T_c. Actuation in the form of the time-varying potential causes the state of the ensemble to return to ρ_a back from starting at ρ_b. Once again, as in (3.20),

$$\mathcal{W}^{(3)} = \Delta\mathcal{E}^{(3)} - T_c\Delta\mathcal{S}^{(3)} + \mathcal{W}^{(3)}_{\text{irr}} \tag{3.25a}$$

$$\mathcal{Q}^{(3)} = T_c\Delta\mathcal{S}^{(3)} - \mathcal{W}^{(3)}_{\text{irr}} \tag{3.25b}$$

$$\mathcal{W}^{(3)}_{\text{irr}} = \frac{\gamma}{t_3}\mathrm{W}_2(\rho_a, \rho_b)^2. \tag{3.25c}$$

4) Adiabatic process: Finally, an adiabatic transition over an interval of infinitesimal duration ("$t_4 = 0$") returns the ensemble to be in contact with a heat reservoir of temperature T_h for a total period of the cycle $t_{\text{period}} = t_1 + t_3$. The state of the ensemble remains at ρ_a, to begin the cycle again. As before, in (3.22),

$$\mathcal{W}^{(4)} = \Delta\mathcal{E}^{(4)} \tag{3.26a}$$

$$\mathcal{Q}^{(4)} = 0 \tag{3.26b}$$

3.5.4 Thermodynamic efficiency & power delivered

For a cyclic process the total change in internal energy

$$\sum_{i=1}^{4} \Delta\mathcal{E}^{(i)} = 0.$$

On the other hand, the entropy doesn't change during the adiabatic transitions

$$\Delta\mathcal{S}^{(i)} = 0, \text{ for } i = 2, 4,$$

while, since it depends only on the end-point states

$$\Delta S^{(1)} = -\Delta S^{(3)} = \mathcal{S}(\rho_b) - \mathcal{S}(\rho_a) =: \Delta\mathcal{S}.$$

As a result, the total work output is

$$-\mathcal{W} = -\left(\sum_{i=1}^{4} \Delta\mathcal{E}^{(i)} - \sum_{i=1}^{4} T_i \Delta\mathcal{S}^{(i)} + \sum_{i=1}^{4} \mathcal{W}_{irr}^{(i)}\right) = (T_h - T_c)\Delta\mathcal{S} - \mathcal{W}_{\text{irr}}^{(1)} - \mathcal{W}_{\text{irr}}^{(3)}. \quad (3.27)$$

Thus, assuming optimality of the choice of the potential to minimize $\mathcal{W}_{\text{irr}}$ in each transition, we conclude that the total work output possible is

$$-\mathcal{W} = (T_h - T_c)\Delta\mathcal{S} - \gamma(\frac{1}{t_1} + \frac{1}{t_3})\mathrm{W}_2(\rho_a, \rho_b)^2. \quad (3.28)$$

Since $T_h > T_c$, naturally, a necessary condition for positive work output is that $\Delta\mathcal{S} := \mathcal{S}(\rho_b) - \mathcal{S}(\rho_a) > 0$ which dictates that phase 1 is an isothermal expansion and phase 3, an isothermal contraction.[6]

The thermodynamic efficiency of an engine is the ratio of work extracted over the heat

[6]The opposite would be true if we sought to operate the cycle for refrigeration purposes.

dissipated,

$$\eta = \frac{-\mathcal{W}}{\mathcal{Q}_h}, \tag{3.29}$$

where the heat input during isothermal expansion is

$$\mathcal{Q}_h = \Delta\mathcal{Q}^{(1)} = T_h\Delta\mathcal{S} - \mathcal{W}_{\text{irr}}.$$

Once again assuming optimality ($\mathcal{W}_{\text{irr}} = \frac{\gamma}{t_1}\mathrm{W}_2(\rho_a, \rho_b)^2$), the bound on the efficiency is seen to be

$$\eta = \frac{(T_h - T_c)\Delta\mathcal{S} - \gamma(\frac{1}{t_1} + \frac{1}{t_3})\mathrm{W}_2(\rho_a, \rho_b)^2}{T_h\Delta\mathcal{S} - \gamma\frac{1}{t_1}\mathrm{W}_2(\rho_a, \rho_b)^2}. \tag{3.30}$$

When the period of the cyclic process tends to infinity (and hence, $t_1, t_3 \to \infty$), tends to the Carnot limit for quasistatic (infinitely slow) transitions $\eta_{\mathrm{C}} = 1 - \frac{T_c}{T_h}$.

Periodic operation, over a finite period $t_1 + t_3$ (since $t_2 = t_4 = 0$), delivers

$$P = -\frac{\mathcal{W}}{(t_1 + t_3)} = \frac{(T_h - T_c)\Delta\mathcal{S} - \gamma(\frac{1}{t_1} + \frac{1}{t_3})\mathrm{W}_2(\rho_a, \rho_b)^2}{t_1 + t_3}$$

units of power. Note that the power output is zero when Carnot efficiency is achieved, because the total duration $t_1 + t_3 \to \infty$. In the sequel, we focus on assessing bounds on available power.

3.6 Fundamental limits to power

Our main interest is in assessing the maximal amount of power that can be drawn by a thermodynamic engine operating between heat baths with temperatures T_h and $T_c < T_h$,

i.e.,"hot" and "cold", respectively. In the present work we draw conclusions based on the basic model in (3.2) via analysis of the thermodynamic cycle that was presented in Section 3.5.

Consider the expression in (5.2) for the power that can be drawn via a cyclic operation as discussed. Preparation of the ensemble, and actuation during the cycle, allow a number of choices. Specifically, the power depends on the period $t_1 + t_3$, the times of the two isothermal phases t_1, t_3 individually, as well as the end-point states (distributions) ρ_a, ρ_b. The latter choice impacts both, the Wasserstein distance $\mathrm{W}_2(\rho_a, \rho_b)$ as well as the change in entropy $\Delta\mathcal{S}$. We will explore systematically the various options.

3.6.1 Optimizing the time scheduling

Optimizing the maximal power delivered during cyclic operation

$$P = \frac{1}{t_1 + t_3}(T_h - T_c)\Delta\mathcal{S} - \frac{\gamma}{t_1 t_3}\mathrm{W}_2(\rho_a, \rho_b)^2,$$

with respect to choices for t_1, t_3, with $\mathrm{W}_2(\rho_a, \rho_b)$, T_h, T_c and $\Delta\mathcal{S}$ kept fixed, gives that

$$t_1 = t_3 = \frac{4\gamma \mathrm{W}_2(\rho_a, \rho_b)^2}{(T_h - T_c)\Delta\mathcal{S}}, \tag{3.31}$$

and therefore that the period for the cycle is

$$t_{\text{cycle}} := t_1 + t_3 = \frac{8\gamma \mathrm{W}_2(\rho_a, \rho_b)^2}{(T_h - T_c)\Delta\mathcal{S}}. \tag{3.32}$$

If instead we specify the period of the cycle t_{cycle}, and optimize with respect to the breakdown

between t_1 and t_3, we once again obtain that the durations of the two phases are equal

$$t_1 = t_3 = \frac{t_{\text{cycle}}}{2}. \tag{3.33}$$

Remark 3.6.1 (Efficiency at maximum power)**.** *The thermodynamic efficiency* (3.29) *of the engine, when it is operating at optimal transition times* (3.31) *that maximize the power, is equal to*

$$\eta_{SS} = \frac{2(T_h - T_c)}{3T_h + T_c} = \frac{\eta_C}{2 - \frac{\eta_C}{2}} \tag{3.34}$$

This result appeared in [31], and [35] for the case of Gaussian marginals ρ_a, ρ_b and potential $U(t, x)$ that is quadratic in x. Our derivation establishes (3.34) *in a general setting.*

Using the expression (3.32), the total power delivered

$$P = \frac{(T_h - T_c)^2}{16\gamma} \left(\frac{\Delta \mathcal{S}}{\mathrm{W}_2(\rho_a, \rho_b)} \right)^2. \tag{3.35}$$

But as we will see in Section 3.6.2, optimizing the power for ρ_a, ρ_b leads to the non-physical conclusion of a vanishingly small t_{cycle} .

3.6.2 The caveat of optimal t_{cycle}: Gaussian states ρ_a, ρ_b

The case where the two marginal distributions/states are Gaussian allows for closed-form expressions for $\Delta \mathcal{S}$ and their Wasserstein distance. Indeed, if ρ_a, ρ_b are Gaussian distributions with zero mean and variances Σ_a, Σ_b, respectively, then

$$\mathrm{W}_2(\rho_a, \rho_b)^2 = \mathrm{tr}\big(\Sigma_a + \Sigma_b - 2(\Sigma_a^{1/2} \Sigma_b \Sigma_a^{1/2})^{1/2}\big) \tag{3.36a}$$

and

$$\Delta\mathcal{S} = \mathcal{S}(\rho_b) - \mathcal{S}(\rho_a) = \frac{1}{2} k_B \log\det(\Sigma_b \Sigma_a^{-1}). \tag{3.36b}$$

Evidently, these allow deriving explicit expressions for the available power in terms of the respective variances.

Specializing to the case of scalar processes with σ_i ($i \in \{a, b\}$) the corresponding standard deviation, i.e., $\Sigma_i = \sigma_i^2$, and period t_{cycle} for the thermodynamic cycle as in (3.32), we obtain that the maximal power available, as a function of σ_a and σ_b, is given by

$$P(\sigma_a, \sigma_b) = \frac{k_B^2 (T_h - T_c)^2}{16\gamma} \left(\frac{\log \frac{\sigma_b}{\sigma_a}}{\sigma_b - \sigma_a} \right)^2. \tag{3.37}$$

The corresponding heat uptaken from the hot reservoir and the work extracted during one cycle are

$$\mathcal{Q}^{(1)} = \mathcal{Q}_h = \frac{1}{4} k_B (3T_h + T_c) \log \frac{\sigma_b}{\sigma_a}$$

and

$$-\mathcal{W} = \frac{1}{2} k_B (T_h - T_c) \log \frac{\sigma_b}{\sigma_a},$$

respectively.

The maximum of the power $P(\sigma_a, \sigma_b)$ over either σ_a, or σ_b, takes place when $\sigma_a = \sigma_b$. But at this limiting condition, although

$$\max_{\sigma_b} P(\sigma_a, \sigma_b) = \frac{k_B^2 (T_h - T_c)^2}{16\gamma\sigma_a^2} \tag{3.38a}$$

and the rate with which heat is drawn is

$$\lim_{\sigma_b \to \sigma_a} \frac{\mathcal{Q}_h}{t_{\text{cycle}}} = \frac{k_B^2(3T_h + T_c)(T_h - T_c)}{32\gamma\sigma_a^2},$$

the limiting values of $-\Delta\mathcal{W}$, $\mathcal{Q}_h$ over a cycle vanish, as does the period t_{cycle} of the cycle. Thus we are led to a non-physical situation of a vanishingly small period for the thermodynamic cycle.

A similar issue in the context of power in quantum engines is brought up in [38]. In the setting herein, in addition, it is seen that taking

$$\sigma_a \to 0$$

and operating with a vanishingly small period for the cycle, leads to infinite power. Once again, bringing up a non-practical situation that is questionable on physical grounds. In the sequel we focus on t_{cycle} being finite.

3.6.3 Optimizing the thermodynamic state ρ_b

Henceforth we fix the period t_{cycle} as well as the duration of the isothermal phases according to (3.33). The power delivered, as a function of the ρ_i's ($i \in \{a, b\}$), is

$$\frac{(T_h - T_c)}{t_{\text{cycle}}}(\mathcal{S}(\rho_b) - \mathcal{S}(\rho_a)) - \frac{4\gamma}{t_{\text{cycle}}^2}\mathrm{W}_2(\rho_a, \rho_b)^2. \tag{3.39}$$

We now consider the problem to maximize power over choice of ρ_b, with ρ_a specified. This problem reduces to finding a suitable minimizer of

$$\min_{\rho_b}\{\mathrm{W}_2(\rho_a, \rho_b)^2 - h\mathcal{S}(\rho_b)\} \tag{3.40}$$

for $h = \frac{t_{\text{cycle}}(T_h - T_c)}{4\gamma}$. Throughout we assume that states have finite second-order moments. As noted earlier, the space of probability distributions (measures, in general) with finite second-order moments $\mathcal{P}_2(\mathbb{R}^d)$ is metrized by the Wasserstein metric $\mathrm{W}_2(\cdot,\cdot)$ and, as can easily be verified, the expression

$$\mathrm{W}_2(\rho_a, \rho_b)^2 - h\mathcal{S}(\rho_b) \tag{3.41}$$

is strictly convex, which leads to the following statement.

Proposition 3.6.1. *Assuming that T_h, T_c as well as t_{cycle} and an initial state $\rho_a \in \mathcal{P}_2(\mathbb{R}^d)$ are specified, there exists a unique minimizer ρ_b of* (3.40)*.*

Proof. Equation (3.40) is similar to one step in the so-called JKO-scheme (also, proximal projection) that displays the heat equation as the gradient flow of the Shannon entropy [50]. While $\mathrm{W}_2(\rho_a, \rho_b)^2 - h\mathcal{S}(\rho_b)$ is strictly convex, it is not automatically bounded from below. Thus, a rather extensive and technical argument is needed to show existence and uniqueness of a minimizer. This is detailed in [50, Proposition 4.1]. □

We conclude this section with two statements. The first establishes implicit conditions for optimality of ρ_b, in maximizing the expression in (3.39) (equivalently, minimizing (3.41)). For ease of referencing we view the expression in (3.39) as a function of ρ_b, namely,

$$f(\rho_b) := \frac{(T_h - T_c)}{t_{\text{cycle}}}(\mathcal{S}(\rho_b) - \mathcal{S}(\rho_a)) - \frac{4\gamma}{t_{\text{cycle}}^2}\mathrm{W}_2(\rho_a, \rho_b)^2. \tag{3.42}$$

The following lemma provides stationarity conditions for $f(\rho_b)$ that, albeit, are implicit in that they involve the optimal transport map from ρ_a and ρ_b that minimizes quadratic transportation cost [55, Ch. 5]. We first highlight stationarity conditions that characterize the minimizer of $f(\cdot)$ in (3.42).

Lemma 3.6.1. *Consider two probability densities ρ_a, ρ_b^* in $\mathcal{P}_2(\mathbb{R}^d)$, where ρ_b^* is the unique maximizer of $f(\rho_b)$, and let $\nabla_x\psi$, for a convex function ψ on $\mathbb{R}^d$, be such that $\nabla_x\psi\sharp\rho_a = \rho_b^*$. The following (stationarity) condition holds*

$$k_B(T_h - T_c)\nabla_x \log \rho_b^*(y) - \frac{8\gamma}{t_{\text{cycle}}}\left((\nabla_x\psi)^{-1} - \text{Id}\right)(y) = 0, \tag{3.43}$$

where Id *denotes the identity map.*

Proof. Consider an arbitrary smooth vector field ξ on $\mathbb{R}^d$ with bounded support, and $\Psi_s : \mathbb{R}^d \to \mathbb{R}^d$ defined by

$$\frac{\partial}{\partial s}\Psi_s(x) = \xi(\Psi_s(x)), \quad \Psi_0 = \text{Id},$$

for $x \in \mathbb{R}^d$ and $s \geq 0$. If $\rho_s := \Psi_s\sharp\rho_b^*$, we claim that

$$\lim_{s\to 0}\frac{1}{s}(f(\rho_s) - f(\rho_b^*)) \geq \int_{\mathbb{R}^d}\langle D_f(x), \xi(x)\rangle\rho_b^*(x)\mathrm{d}x, \tag{3.44}$$

where, for $\Delta T := T_h - T_c$,

$$D_f(x) = -\frac{k_B\Delta T}{t_{\text{cycle}}}\nabla\log(\rho_b^*(x)) + \frac{8\gamma}{t_{\text{cycle}}^2}(\nabla\psi^{-1}(x) - x).$$

Assuming the claim is true (to be shown shortly), then, because ρ_b^* is the maximizer, $f(\rho_s) \leq f(\rho_b^*)$. Therefore

$$\int\langle D_f(x), \xi(x)\rangle\rho_b^*(x)\mathrm{d}x \leq \lim_{s\to 0}\frac{f(\rho_s) - f(\rho_b^*)}{s} \leq 0.$$

Hence, by symmetry $\xi \to -\xi$,

$$\int\langle D_f(x), \xi(x)\rangle\rho_b^*(x)\mathrm{d}x = 0. \tag{3.45}$$

This is true for all vector fields $\xi \in C_0^\infty(\mathbb{R}^d, \mathbb{R}^d)$. As a result, $D_f(x) = 0$, concluding (3.43) and the lemma.

It remains to prove (3.44). By definition,

$$f(\rho_s) - f(\rho_b^*) = \frac{\Delta T}{t_{\text{cycle}}}(\mathcal{S}(\rho_s) - \mathcal{S}(\rho_b^*)) - \frac{4\gamma}{t_{\text{cycle}}^2}(\mathrm{W}_2(\rho_a, \rho_s)^2 - \mathrm{W}_2(\rho_a, \rho_b^*)^2).$$

The entropy term

$$\begin{aligned}\mathcal{S}(\rho_s) =& - k_B \int \log(\rho_s(x))\rho_s(x)\mathrm{d}x = -k_B \int \log(\rho_s(\Psi_s(x)))\rho_b^*(x)\mathrm{d}x \\ =& - k_B \int \log(\frac{\rho_b^*((x))}{\det(\nabla\Psi_s(x))})\rho_b^*(x)\mathrm{d}x = S(\rho_b^*) + k_B \int \log(\det(\nabla\Psi_s(x)))\rho_b^*(x)\mathrm{d}x.\end{aligned}$$

Therefore

$$\begin{aligned}\lim_{s\to 0}\frac{1}{s}(S(\rho_s) - S(\rho_b^*)) =& \lim_{s\to 0}\frac{k_B}{s}\int \log(\det(\nabla\Psi_s(x)))\rho_b^*(x)\mathrm{d}x = k_B\int \nabla\cdot\xi(x)\rho_b^*(x)\mathrm{d}x \\ =& - k_B \int \langle \xi(x), \nabla\log(\rho_b^*(x))\rangle\rho_b^*(x)\mathrm{d}x.\end{aligned}$$

The Wasserstein term

$$\begin{aligned}\mathrm{W}_2(\rho_a, \rho_s)^2 - \mathrm{W}_2(\rho_a, \rho_b^*)^2 \le& \int \|\nabla\psi^{-1}(x) - \Psi_s(x)\|^2\rho_b^*(x)\mathrm{d}x - \int \|\nabla\psi^{-1}(x) - x\|^2\rho_b^*(x)\mathrm{d}x \\ =& \int \langle x - \Psi_s(x), 2\nabla\psi^{-1}(x) - x - \Psi_s(x)\rangle\rho_b^*(x)\mathrm{d}x.\end{aligned}$$

Therefore

$$\lim_{s\to 0}\frac{1}{s}\left[\mathrm{W}_2(\rho_a, \rho_s)^2 - \mathrm{W}_2(\rho_a, \rho_b^*)^2\right] \le -2\int \langle \xi(x), \nabla\psi^{-1}(x) - x\rangle\rho_b^*(x)\mathrm{d}x.$$

Using the two expressions, the one for derivative of the entropy and the other for the Wasserstein distance, the claim follows.

□

The lemma, which is of independent interest, is used in the proof of the following proposition concluding the section. The proposition states that, for scalar distributions for simplicity, if ρ_a is Gaussian, then so is ρ_b. As a consequence the optimal actuation protocol is based on a time-varying potential $U(t, x)$ that is quadratic in x.

Proposition 3.6.2. *If ρ_a is a one-dimensional Gaussian distribution with zero mean and variance σ_a^2, then ρ_b^* is also Gaussian with zero mean and variance σ_b^2, where*

$$\sigma_b = \frac{1+\sqrt{1+c}}{2}\sigma_a, \tag{3.46}$$

and $c = \frac{k_B(T_h - T_c)t_{\text{cycle}}}{2\gamma\sigma_a^2}$.

Proof. According to Proposition 3.6.1, the maximizer is unique. Therefore, it is sufficient to show that the Gaussian distribution $N(0, \sigma_b^2)$, where σ_b^2 is given by (3.46), satisfies the optimality condition (3.43). When ρ_a, ρ_b^* are Gaussian, $\nabla\psi^{-1}(y) = \frac{\sigma_a}{\sigma_b}y$. Hence, the optimality condition reads

$$\begin{aligned}
&\frac{k_B t_{\text{cycle}}\Delta T}{8\gamma}\nabla \log \rho_b^*(y) - y + \nabla\psi^{-1}(y) = \frac{k_B t_{\text{cycle}}\Delta T}{8\gamma}\frac{y}{\sigma_b^2} - (1 - \frac{\sigma_a}{\sigma_b})y \\
=&(\frac{k_B t_{\text{cycle}}\Delta T}{8\gamma\sigma_b^2} - 1 + \frac{\sigma_a}{\sigma_b})y = 0, \quad \forall\, y \in \mathbb{R},
\end{aligned}$$

which is satisfied when σ_b is according to (3.46).

□

Remark 3.6.2. *In earlier works, it is commonly assumed that the marginal distributions ρ_a, ρ_b are Gaussian and the potential function $U(t, x)$ is quadratic in x. Proposition 3.6.2 justifies this assumption to some extent: if ρ_a is specified to be Gaussian, the optimal ρ_b and the optimal potential function that achieve the maximum power, are Gaussian and quadratic,*

respectively. However, as we will see in Section 3.6.4, if instead ρ_b is specified as Gaussian distribution, the optimal ρ_a is not Gaussian. Gaussian distributions turn out instead to be local minimizers *of the power under certain conditions (see discussion following Remark 3.6.3).*

3.6.4 Optimizing the thermodynamic state ρ_a

We now consider the dependence of the maximal power on ρ_a, i.e., on the thermodynamic state at which the ensemble begins its expansive phase. As we will see, the situation is not symmetric to the conclusions drawn in Section 3.6.3 with regard to ρ_b and, without further assumptions, an optimal ρ_a does not exist. Interestingly, on closer inspection, the source of this conundrum is the unreasonably high demands on the magnitude of $\nabla_x U$ for the controlling potential $U(t, x)$. The insights gained lead to the framework for maximal power in the follow up section.

For simplicity, and without any loss of generality for the purposes of this section, we assume that ρ_b is specified to be a zero-mean Gaussian distribution with standard deviation σ_b. In view of (3.39), a choice of ρ_a that is close to a Dirac delta distribution allows arbitrarily large negative values for the entropy, i.e., $\mathcal{S}(\rho_a) \simeq -\infty$, and hence infinite power.

Thus, it is natural to impose a lower bound on the entropy of ρ_a, or simply fix $-\infty < s_a = \mathcal{S}(\rho_a) < \mathcal{S}(\rho_b)$. But in this case, and once more in view of (3.39), maximal power would be drawn by minimizing $\mathrm{W}_2(\rho_a, \rho_b)$ over probability densities ρ_a with entropy s_a. We claim that

$$\inf_{\rho_a}\{\mathrm{W}_2(\rho_a, \rho_b) \mid \mathcal{S}(\rho_a) = s_a > -\infty\} = 0. \tag{3.47}$$

To see this note that

$$\inf_{\rho_a} \mathrm{W}_2(\rho_a, \rho_b) = 0$$

by taking ρ_a to approximate an increasingly fine train of suitably scaled Dirac deltas, i.e.,

$$\rho_a(x) \approx \sum_{i \in \mathbb{Z}} \rho_i \delta_{x_i}(x)$$

where $\rho_i = \int_{x_i}^{x_{i+1}} \rho_b(x)\mathrm{d}x$ and x_i ($i \in \mathbb{Z}$) equispaced. The latter is a singular distribution which, however, can be approximated arbitrarily closely in W_2 by a probability density with any given entropy. Such a density can be produced by approximating Dirac deltas by a piecewise constant function with finite support.

The optimization problem (3.47) is inherently related to the continuity of the entropy functional with respect to the Wasserstein distance. For a rigorous treatment of the problem, see [67], where it is shown that unless certain regularity assumptions are in place for ρ_a and ρ_b, the infimum in (3.47) is zero.

Remark 3.6.3 (Gaussian is not optimal for ρ_a)**.** *The preceeding arguments show that a Gaussian distribution is not the optimal choice for ρ_a with respect to maximizing power, even when ρ_b is Gaussian, unless additional constraints are introduced.*

Since the Gaussian distribution maximizes entropy when mean and variance are specified, it is natural to explore constraints on the mean and variance of ρ_a for the purposes of maximizing power. Without loss of generality, the mean can be assumed to be zero and the variance specified to be $\sigma_a^2 < \sigma_b^2$. First-order and second order optimality analysis for the power output (3.39), at $\rho_a = N(0, \sigma_a^2)$ is carried out. It turns out that, although $N(0, \sigma_a^2)$ satisfies the first-order optimality condition, it does not satisfy the second-order optimality condition. In fact, $N(0, \sigma_a^2)$ is a local minimizer when $\sigma_a < \sigma_b < k_B(T_h - T_c)t_{\mathrm{cycle}}/(8\gamma\sigma_a)$.

The analysis is given as follows, and aims to highlight that the conjecture that a Gaussian ρ_a is optimal fails. This, in fact, is not surprising. Maximizing the power over ρ_a is equivalent to minimizing the entropy of ρ_a. Minimizing entropy under fixed variance constraint does not lead to Gaussian distributions as two Dirac delta distribution with the desired mean and variance achieve negative infinity entropy.

Proof. Let $\mathcal{A}_{0,\sigma^2}$ denote the set of absolutely continuous distributions with mean 0 and variance σ^2 :

$$\mathcal{A}_{0,\sigma^2} := \left\{ \rho \in \mathcal{P}_{2,\mathrm{ac}}(\mathbb{R}^d); \int x\rho(x)\mathrm{d}x = 0, \int x^2\rho(x)\mathrm{d}x = \sigma^2 \right\}$$

and consider the functional

$$g(\rho_a) = -\frac{T_h - T_c}{t_{\mathrm{cycle}}} S(\rho_a) - \frac{4\gamma}{t_{\mathrm{cycle}}^2} \mathrm{W}_2(\rho_a, \rho_b)^2 \tag{3.48}$$

that represents the portion of power given in (3.39) that depends on ρ_a.

(i) We first show that $\max_{\rho_a \in \mathcal{A}_{0,\sigma_a^2}} g(\rho_a)$ is unbounded, and hence that the maximizer does not exist. Consider a sequence of density functions $\{\mu_n\}_{n\in\mathbb{N}}$ according to

$$\mu_n = \frac{1}{2}\mu_n^{(1)} + \frac{1}{2}\mu_n^{(2)},$$

where

$$\begin{aligned} \mu_n^{(1)} &= N(\sqrt{\sigma_a^2 - \frac{1}{n^2}}, \frac{1}{n^2}) \\ \mu_n^{(2)} &= N(-\sqrt{\sigma_a^2 - \frac{1}{n^2}}, \frac{1}{n^2}). \end{aligned} \tag{3.49}$$

It is easy to verify that $\mu_n \in \mathcal{A}_{0,\sigma_a^2}$. The goal is to show that $\lim_{n\to\infty} g(\mu_n) = \infty$. The

entropy $\mathcal{S}(\mu_n)$ is bounded from above as follows,

$$\mathcal{S}(\mu_n) \leq \frac{1}{2}(k_B \log(2) + \mathcal{S}(\mu_n^{(1)})) + \frac{1}{2}(k_B \log(2) + \mathcal{S}(\mu_n^{(2)})) = k_B \log 2\sqrt{2\pi e} - k_B \log n.$$

The first inequality follows from a respective bound on the entropy of Gaussian mixtures [68, Theorem 3].

The Wasserstein distance is bounded by

$$\mathrm{W}_2(\mu_n, \rho_b)^2 \leq \frac{1}{2}\mathrm{W}_2(\mu_n^{(1)}, \rho_b)^2 + \frac{1}{2}\mathrm{W}_2(\mu_n^{(2)}, \rho_b)^2 = \sigma_a^2 + \sigma_b^2 - 2\frac{\sigma_b}{n}, \tag{3.50}$$

where the convexity of the functional $\mathrm{W}_2(\cdot, \rho_b)^2$ is used [69, Eq.(2.12)]. Combining the two bounds for the entropy and Wasserstein distance yields

$$\begin{aligned} g(\mu_n) &= -\frac{\Delta T}{t_{\mathrm{cycle}}}\mathcal{S}(\mu_n) - \frac{4}{t_{\mathrm{cycle}}^2}\mathrm{W}_2(\mu_n, \rho_b)^2 \\ &\geq \frac{k_B \Delta T}{t_{\mathrm{cycle}}}\left(-\log 2\sqrt{2\pi e} + \log n\right) - \frac{4}{t_{\mathrm{cycle}}^2}\left(\sigma_a^2 + \sigma_b^2 - 2\frac{\sigma_b}{n}\right). \end{aligned}$$

Taking the limit $n \to \infty$ proves $\lim_{n\to\infty} g(\mu_n)$ is unbounded, and hence that there is no maximizer. Next, in (ii) and (iii) we show that the Gaussian distribution is instead a local minimizer, under certain conditions, and hence that it is the opposite of distributions that we seek.

(ii) We now perform first-order optimality analysis for

$$\max_{\rho_a \in \mathcal{A}_{0,\sigma_a^2}} g(\rho_a)$$

at $\rho_a = N(0, \sigma_a^2)$. Consider a smooth vector field with bounded support $\xi \in C_0^\infty(\mathbb{R}, \mathbb{R})$

such that

$$\int_{\mathbb{R}^d} \xi(x)\rho_a(x)dx = 0. \tag{3.51}$$

Then, as before, define the flow $\Psi_s : \mathbb{R} \to \mathbb{R}$ generated by ξ according to

$$\frac{\partial}{\partial s}\Psi_s(x) = \xi(\Psi_s(x)), \quad \Psi_0 = \mathrm{Id},$$

for $x \in \mathbb{R}$ and $s \in \mathbb{R}$. Now define $\tilde{\rho}_s := \Psi_s \sharp \rho_a$. Because of (3.51), the mean of the distribution $\tilde{\rho}_s$ remains constant at 0. However, the variance changes from σ_a^2, and $\tilde{\rho}_s$ is not inside $\mathcal{A}_{0,\sigma_a^2}$.

In order to keep the variance constant at σ_a^2, we project $\tilde{\rho}_s$ into $\mathcal{A}_{0,\sigma_a^2}$ by setting $\rho_s = G_s \sharp \tilde{\rho}_s$, where $G_s(x) := r(s)x$ and

$$r(s) = \frac{\sigma_a}{\sqrt{\int_{\mathbb{R}^d} |y|^2 \tilde{\rho}_s(y)\mathrm{d}y}} = \frac{\sigma_a}{\sqrt{\int_{\mathbb{R}^d} |\Psi_s(x)|^2 \rho_a(x)\mathrm{d}x}}.$$

By definition of G_s, the variance of ρ_s is equal to σ_a^2, hence $\rho_s \in \mathcal{A}_{0,\sigma_a^2}$.

Following the same procedure as in the proof of Lemma 3.6.1, the first-order optimality condition is

$$\int \langle D_g(x), v(x)\rangle \rho_a(x)\mathrm{d}x = 0, \tag{3.52}$$

where

$$v(x) = \frac{\mathrm{d}}{\mathrm{d}s} G_s(\Psi_s(x))|_{s=0} = \xi(x) - \frac{x}{\sigma_a^2}\int \xi(z) z \rho_a(z)\mathrm{d}z, \tag{3.53}$$

and

$$D_g(x) = \frac{k_B \Delta T}{t_{\text{cycle}}} \nabla \log(\rho_a^*(x)) + \frac{8\gamma}{t_{\text{cycle}}^2}(\nabla\psi(x) - x)$$

where $\nabla\psi$ is the optimal transport map from ρ_a to ρ_b. For the setting where ρ_a and ρ_b are $N(0, \sigma_a^2)$ and $N(0, \sigma_b^2)$, respectively, $\nabla\psi(x) = \frac{\sigma_b}{\sigma_a}x$ and

$$D_g(x) = -\frac{k_B \Delta T}{t_{\text{cycle}}}x + \frac{8\gamma}{t_{\text{cycle}}^2}(\frac{\sigma_b}{\sigma_a} - 1)x. \tag{3.54}$$

Letting $\alpha := -\frac{k_B\Delta T}{t_{\text{cycle}}} + \frac{8\gamma}{t_{\text{cycle}}^2}(\frac{\sigma_b}{\sigma_a} - 1)$ and inserting (3.54) into (3.52) yields $\alpha \int xv(x)\rho_a(x)\mathrm{d}x = 0$, which is satisfied because of (3.53). Hence, ρ_a being $N(0, \sigma_a^2)$ satisfies first-order optimality condition.

(iii) We follow up by carrying out second-order analysis. The objective is to show that the limit

$$\begin{aligned}&\lim_{s\to 0} \frac{1}{s^2}\left(g(\rho_s) + g(\rho_{-s}) - 2g(\rho_a)\right)\\ &= -\frac{\Delta T}{t_{\text{cycle}}} \lim_{s\to 0} \frac{\mathcal{S}(\rho_s) + \mathcal{S}(\rho_{-s}) - 2\mathcal{S}(\rho_a)}{s^2} - \frac{4\gamma}{t_{\text{cycle}}^2} \lim_{s\to 0} \frac{\mathrm{W}_2(\rho_s, \rho_b)^2 + \mathrm{W}_2(\rho_{-s}, \rho_b)^2 - 2\mathrm{W}_2(\rho_a, \rho_b)^2}{s^2}\end{aligned}$$

can be strictly positive. Assume $\xi(x) = \nabla\eta(x)$ for some η, and define

$$\zeta(x) = \eta(x) - \frac{x^2}{2\sigma_a^2}\int_{\mathbb{R}} \langle z, \nabla\eta(z)\rangle \rho_a(z)\mathrm{d}z,$$

For the second order derivative of the entropy, we use the existing results from [69, Eq. (2.30), Eq.(3.37)], where it is shown that

$$\frac{\mathrm{d}^2}{\mathrm{d}s^2} S(\rho_s)\bigg|_{s=0} = -k_B \int \left(\|\nabla^2\zeta\|_F^2 + \frac{1}{\sigma_a^2}\|\nabla\zeta\|^2\right)\rho_a \mathrm{d}x \tag{3.55}$$

Next, we consider the second order derivative of the Wasserstein distance. Since $\nabla\psi\sharp\rho_a = \rho_b$ and $(G_s \circ \Psi_s)\sharp\rho_a = \rho_s$, we have

$$\mathrm{W}_2(\rho_b, \rho_s)^2 \leq \int_{\mathbb{R}} \|\nabla\psi(x) - G_s(\Psi_s(x))\|^2 \rho_a(x)\mathrm{d}x.$$

As a result

$$\begin{aligned}
&\lim_{s\to 0} \frac{\mathrm{W}_2(\rho_s, \rho_b)^2 + \mathrm{W}_2(\rho_{-s}, \rho_b)^2 - 2\mathrm{W}_2(\rho_a, \rho_b)^2}{s^2} \\
\leq &\lim_{s\to 0} \frac{1}{s^2}\bigg[\int \|\nabla\psi(x) - G_s(\Psi_s(x))\|^2 \rho_a(x)\mathrm{d}x + \int \|\nabla\psi(x) - G_{-s}(\Psi_{-s}(x))\|^2 \rho_a(x)\mathrm{d}x \\
&\qquad - 2\int \|\nabla\psi(x) - x\|^2 \rho_a(x)\mathrm{d}x\bigg] \\
= &\lim_{s\to 0} \frac{1}{s^2}\bigg[\int \left(\|G_s(\Psi_s(x))\|^2 + \|G_{-s}(\Psi_{-s}(x))\|^2\right) \rho_a(x)\mathrm{d}x \\
&- 2\int \left(\|x\|^2 + \langle \Omega_s(x), \nabla\psi(x)\rangle\right) \rho_a(x)\mathrm{d}x\bigg],
\end{aligned}$$

where $\Omega_s(x) = G_s(\Psi_s(x)) + G_{-s}(\Psi_{-s}(x)) - 2x$. The first three terms cancel out, because the variance is constant. Therefore, the limit simplifies to

$$\begin{aligned}
&\lim_{s\to 0} \frac{\mathrm{W}_2(\rho_s, \rho_b)^2 + \mathrm{W}_2(\rho_{-s}, \rho_b)^2 - 2\mathrm{W}_2(\rho_a, \rho_b)^2}{s^2} \\
\leq &-2\int_{\mathbb{R}^d} \lim_{s\to 0} \langle \frac{\Omega_s(x)}{s^2}, \nabla\psi(x)\rangle \rho_a(x)\mathrm{d}x \\
= &-2\int_{\mathbb{R}^d} \langle \frac{\partial^2 G_s(\Psi_s(x))}{\partial s^2}\bigg|_{s=0}, \nabla\psi(x)\rangle \rho_a(x)\mathrm{d}x \\
= &-2\frac{\sigma_b}{\sigma_a}\int_{\mathbb{R}^d} \langle \frac{\partial^2 G_s(\Psi_s(x))}{\partial s^2}(x)\bigg|_{s=0}, x\rangle \rho_a(x)\mathrm{d}x,
\end{aligned} \tag{3.56}$$

where $\nabla\psi(x) = \frac{\sigma_b}{\sigma_a}x$ is used in the last step. Next we compute $\frac{\partial^2 G_s(\Psi_s(x))}{\partial s^2}(x)|_{s=0}$. Differentiating once gives

$$\frac{\partial G_s(\Psi_s(x))}{\partial s} = \frac{\partial}{\partial s}(r(s)\Psi_s(x)) = r(s)\nabla\eta\left(\Psi_s(x)\right) + \dot{r}(s)\Psi_s(x).$$

Differentiating twice and evaluating at $s = 0$ gives

$$\left.\frac{\partial^2 G_s(\Psi_s(x))}{\partial s^2}\right|_{s=0} = 2\dot{r}(0)\nabla\eta(x) + r(0)\nabla^2\eta(x)\nabla\eta(x) + \ddot{r}(0)x.$$

Inserting this expression into (3.56) gives

$$\begin{aligned}&\lim_{s\to 0}\frac{\mathrm{W}_2(\rho_s,\rho_b)^2 + \mathrm{W}_2(\rho_{-s},\rho_b)^2 - 2\mathrm{W}_2(\rho_a,\rho_b)^2}{s^2}\\ \leq& -2\frac{\sigma_b}{\sigma_a}\left[2\dot{r}(0)\int\langle x,\nabla\eta(x)\rangle\rho_a(x)\mathrm{d}x + \int\langle x,\nabla^2\eta(x)\nabla\eta(x)\rangle\rho_a(x)\mathrm{d}x + \sigma_a^2\ddot{r}(0)\right].\end{aligned} \tag{3.57}$$

Inserting the derivatives of $r(s)$,

$$\begin{aligned}\dot{r}(0) =& -\frac{1}{\sigma_a^2}\int\langle\nabla\eta(x),x\rangle\rho_a(x)\mathrm{d}x,\\ \ddot{r}(0) =& \frac{3}{\sigma_a^4}\left(\int\langle\nabla\eta(x),x\rangle\rho_a(x)\mathrm{d}x\right)^2 - \frac{1}{\sigma_a^2}\int(\|\nabla\eta(x)\|^2 + \langle x,\nabla^2\eta(x)\nabla\eta(x)\rangle)\rho_a(x)\mathrm{d}x,\end{aligned}$$

gives

$$\begin{aligned}&\lim_{s\to 0}\frac{\mathrm{W}_2(\rho_s,\rho_b)^2 + \mathrm{W}_2(\rho_{-s},\rho_b)^2 - 2\mathrm{W}_2(\rho_a,\rho_b)^2}{s^2}\\ \leq& \frac{-2\sigma_b}{\sigma_a}\left[\frac{1}{\sigma_a^2}\left(\int\langle x,\nabla\eta(x)\rangle\rho_a\mathrm{d}x\right)^2 - \int\|\nabla\eta\|^2\rho_a\mathrm{d}x\right] = 2\frac{\sigma_b}{\sigma_a}\int\|\nabla\zeta(x)\|^2\rho_a(x)\mathrm{d}x.\end{aligned} \tag{3.58}$$

Using (3.55) and (3.58), we conclude that

$$\lim_{s\to 0}\frac{1}{s^2}\left(g(\rho_s) + g(\rho_{-s}) - 2g(\rho_a)\right) \geq \left(\frac{k_B\Delta T}{t_{\mathrm{cycle}}\sigma_a^2} - \frac{8\gamma\sigma_b}{t_{\mathrm{cycle}}^2\sigma_a}\right)\int\|\nabla\zeta\|^2\rho_a\mathrm{d}x.$$

Hence, when $\sigma_b \in (\sigma_a, \frac{k_B\Delta T t_{\mathrm{cycle}}}{8\gamma\sigma_a}]$, the second-order variation is positive and $\rho_a = N(0,\sigma_a^2)$ is a local minimizer.

□

3.6.5 Maximal power with arbitrary potential

In this section, we show that the power output of a thermodynamic engine, under any choice of potential $U(t,x)$ cannot exceed a bound that involves the Fisher information of the marginal state ρ_a.

Proposition 3.6.3. *Under the standing assumptions on the Carnot-like cycle, the power output* (3.39)*, is bounded by*

$$P \leq \frac{k_B^2(T_h - T_c)^2}{16\gamma} I(\rho_a \mathrm{d}x|\mathrm{d}x). \tag{3.59}$$

Proof. First, we will show that It is a consequence of the HWI* inequality (3.11);
The proof follows by expressing the HWI* inequality (3.11) for a Gaussian reference measure $\mathrm{d}\mathfrak{m}_g = (2\pi\sigma^2)^{-\frac{d}{2}} e^{-\frac{\|x\|^2}{2\sigma^2}} dx$ with constant $\kappa = \frac{1}{\sigma^2}$ and taking the limit as $\sigma \to \infty$. The relative entropy with respect to Gaussian measure is

$$H(\mu|\mathfrak{m}_g) = \int \log(\frac{\mathrm{d}\mu}{\mathrm{d}x})\mathrm{d}\mu - \int \log(\frac{\mathrm{d}\mathfrak{m}_g}{\mathrm{d}x})\mathrm{d}\mu = -k_B^{-1}\mathcal{S}(\frac{\mathrm{d}\mu}{\mathrm{d}x}) + \frac{\sigma_{\mu_a}^2}{2\sigma^2} + \frac{d}{2}\log(2\pi\sigma^2).$$

where $\sigma_\mu^2 := \int \|x\|^2 \mathrm{d}\mu$. Therefore, the left hand side of (3.11)

$$H(\mu_a|\mathfrak{m}_g) - H(\mu_b|\mathfrak{m}_g) = k_B^{-1}(\mathcal{S}(\rho_b) - \mathcal{S}(\rho_a)) + \frac{\sigma_{\mu_a}^2 - \sigma_{\mu_b}^2}{2\sigma^2},$$

with $\rho_a = \mathrm{d}\mu_a/\mathrm{d}x, \rho_b = \mathrm{d}\mu_b/\mathrm{d}x$, and $\sigma_{\mu_a}^2, \sigma_{\mu_b}^2$ are the corresponding variances. On the right hand side, the Fisher information term becomes

$$\begin{aligned} I(\mu_a|\mathfrak{m}_g) &= \int \|\nabla \log(\frac{\mathrm{d}\mu_a}{\mathrm{d}\mathfrak{m}_g})\|^2 \mathrm{d}\mu_a = \int \|\nabla \log(\frac{\mathrm{d}\mu_a}{\mathrm{d}x}) - \nabla \log(\frac{\mathrm{d}\mathfrak{m}_g}{\mathrm{d}x})\|^2 \mathrm{d}\mu_a \\ &= \int \|\nabla \log(\frac{\mathrm{d}\mu_a}{\mathrm{d}x})\|^2 \mathrm{d}\mu_a - 2\int \langle \nabla \log(\frac{\mathrm{d}\mu_a}{\mathrm{d}x}), \frac{-x}{\sigma^2}\rangle \mathrm{d}\mu_a + \int \|\frac{-x}{\sigma^2}\|^2 \mathrm{d}\mu_a \\ &= I(\mu_a|dx) - \frac{2d}{\sigma^2} + \frac{\sigma_{\mu_a}^2}{\sigma^4}. \end{aligned}$$

Thus, taking the limit $\sigma \to \infty$, (3.60) follows. Second

$$\mathcal{S}(\rho_b) - \mathcal{S}(\rho_a) \leq k_B \mathrm{W}_2(\rho_a, \rho_b) \sqrt{I(\rho_a \mathrm{d}x|\mathrm{d}x)}. \tag{3.60}$$

Using the formula for power (3.39), we have

$$\begin{aligned} P &\leq \frac{(T_h - T_c)\Delta\mathcal{S}}{t_{\text{cycle}}} - \frac{4\gamma}{t_{\text{cycle}}^2} \frac{\Delta\mathcal{S}^2}{k_B^2 I(\rho_a \mathrm{d}x|\mathrm{d}x)} \\ &= -\frac{4\gamma}{t_{\text{cycle}}^2} \frac{\left(\Delta\mathcal{S} - \frac{t_{\text{cycle}} k_B^2 (T_h - T_c)}{8\gamma} I(\rho_a \mathrm{d}x|\mathrm{d}x)\right)^2}{k_B^2 I(\rho_a \mathrm{d}x|\mathrm{d}x)} + \frac{k_B^2 (T_h - T_c)^2}{16\gamma} I(\rho_a \mathrm{d}x|\mathrm{d}x) \\ &\leq \frac{k_B^2 (T_h - T_c)^2}{16\gamma} I(\rho_a \mathrm{d}x|\mathrm{d}x), \end{aligned}$$

concluding the bound in (3.59). □

We point out that the bound in (3.59) becomes tight when t_{cycle} takes the optimal value (3.32) and $\rho_b \to \rho_a$. Specifically, if $\rho_a = N(0, \sigma_a^2)$ and $\rho_b = N(0, \sigma_b^2)$ are Gaussian distributions and t_{cycle} takes the optimal value (3.32), then as $\sigma_b \to \sigma_a$ the power output is given by (3.38a), which coincides with (3.59), since $I(\rho_a \mathrm{d}x|\mathrm{d}x) = \frac{1}{\sigma_a^2}$.

3.6.6 Maximal power under constrained potential

While a lower bound on $\mathcal{S}(\rho_a)$ readily implies an upper bound on the available power, achieving such a bound in general requires a cyclic operation involving an irregular and complicated potential function $U(t, x)$ to bring back the ensemble to ρ_a at end of each cycle. It is unreasonable to expect technological solutions to such demands, and therefore, a constraint on the complexity of the potential function seems meaningful. To this end, we

propose the constraint

$$\frac{1}{\gamma}\int_{\mathbb{R}^d}\|\nabla_x U(t,x)\|^2\rho(t,x)\,\mathrm{d}x \le M \tag{3.61}$$

for all $t \in (0, t_{\text{cycle}})$. Thus, we analyze the maximum power (3.39) that can be extracted from a thermodynamic engine, under the constraint (3.61).

Theorem 3.6.1. *Consider a thermodynamic ensemble, undergoing a Carnot cycle as described in Section 3.5, governed with the over-damped Langevin equation* (3.2)*. Then, the maximum power P that can be extracted from the cycle, over all marginal probability distributions ρ_a and ρ_b, the cycle period t_{cycle}, and all potential functions $U(t,x)$ that respect the bound* (3.61)*, satisfies*

$$\frac{M}{8}(\frac{T_h}{T_c}-1)\frac{\frac{T_h}{T_c}-1}{\frac{T_h}{T_c}+1} \le P_{\max} \le \frac{M}{8}(\frac{T_h}{T_c}-1) \tag{3.62}$$

Proof. The proof for the upper-bound follows from bounding the entropy difference $S(\rho_b) - S(\rho_a)$ under the constraint (3.61). During the isothermal transition in contact with the cold bath with temperature T_c,

$$\begin{aligned}\mathcal{S}(\rho_b)-\mathcal{S}(\rho_a) =& -\int_{\frac{t_{\text{cycle}}}{2}}^{t_{\text{cycle}}}\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{S}(\rho(t,\cdot))\,\mathrm{d}t \\ =&\frac{-k_B}{\gamma}\int_{\frac{t_{\text{cycle}}}{2}}^{t_{\text{cycle}}}(\langle\nabla_x\log\rho,\nabla_x U\rangle_\rho + k_BT_c\|\nabla_x\log\rho\|_\rho^2)\mathrm{d}t,\end{aligned}$$

where the notation $\langle\nabla_x f,\nabla_x g\rangle_\rho := \int_{\mathbb{R}^d}\langle\nabla_x f,\nabla_x g\rangle\rho\mathrm{d}x$ and $\|\nabla_x f\|_\rho = \sqrt{\langle\nabla_x f,\nabla_x f\rangle_\rho}$ is used. By the Cauchy-Schwartz inequality and constraint (3.61),

$$-\langle\nabla_x\log\rho,\nabla_x U\rangle_\rho \le \|\nabla_x U\|_\rho\|\nabla_x\log\rho\|_\rho \le \sqrt{\gamma M}\|\nabla_x\log\rho\|_\rho.$$

Hence,

$$\begin{aligned}\mathcal{S}(\rho_b) - \mathcal{S}(\rho_a) \leq & \frac{k_B}{\gamma} \int_{\frac{t_{\text{cycle}}}{2}}^{t_{\text{cycle}}} \left(\sqrt{\gamma M} \|\nabla_x \log \rho\|_\rho - k_B T_c \|\nabla_x \log \rho\|_\rho^2 \right) \mathrm{d}t \\ \leq & \frac{k_B}{\gamma} \int_{\frac{t_{\text{cycle}}}{2}}^{t_{\text{cycle}}} \frac{\gamma M}{4 k_B T_c} \mathrm{d}t = \frac{M}{8 T_c} t_{\text{cycle}}.\end{aligned}$$

This concludes the bound $\Delta\mathcal{S} \leq \frac{M}{T_c} \frac{t_{\text{cycle}}}{8}$ on the entropy difference, which yields to upper-bound on the power output:

$$P \leq \frac{(T_h - T_c)}{t_{\text{cycle}}} \Delta\mathcal{S} - \frac{1}{t_{\text{cycle}}} W_{\text{irr}} \leq \frac{M(T_h - T_c)}{8 T_c} \tag{3.63}$$

where $W_{\text{irr}} \geq 0$ is used.

Next, we prove the lower-bound by describing a setting so that the power is equal to the lower bound. Assume the marginal distributions ρ_a and ρ_b are Gaussian $N(0, \sigma_a^2)$ and $N(0, \sigma_b^2)$ respectively, and the potential function $U(t, x) = \frac{1}{2} a_t x^2$ is a quadratic function. In this setting, the exact power output is equal to

$$P = \frac{1}{t_{\text{cycle}}} k_B (T_h - T_c) \log(\frac{\sigma_b}{\sigma_a}) - \frac{1}{\gamma t_{\text{cycle}}} \int_0^{t_{\text{cycle}}} (a_t - \frac{k_B T}{\sigma_t^2})^2 \sigma_t^2 \mathrm{d}t$$

with update law for the variance given by the Lyapunov equation:

$$\frac{\mathrm{d}\sigma_t^2}{\mathrm{d}t} = -2(\frac{a_t}{\gamma} - \frac{k_B T}{\gamma \sigma_t^2}) \sigma_t^2$$

with the constraint (3.61) given by $\frac{1}{\gamma} a_t^2 \sigma_t^2 \leq M$. Then, in the limit as $t_{\text{cycle}} \to 0$, and $\sigma_b \to \sigma_a = \sigma$, the power output is equal to

$$P = k_B (T_h - T_c) \frac{\lambda}{2} - \gamma \lambda^2 \sigma^2 \tag{3.64}$$

with the constraint

$$|\gamma\lambda + \frac{k_B T_c}{\sigma^2}| \leq \frac{\sqrt{\gamma M}}{\sigma}, \tag{3.65}$$

where we introduced a new variable $\lambda = \frac{a}{\gamma} - \frac{k_B T_c}{\gamma\sigma^2}$. Next, we will show that the maximum of the expression (3.64) over all values of λ and σ that satisfy the constraint (3.65), is equal to the lower-bound.

The constraint (3.65) is expressed as:

$$0 \leq \lambda \leq \frac{\sqrt{\gamma M}}{\gamma\sigma} - \frac{k_B T_c}{\gamma\sigma^2}, \text{ for } \sigma \geq \frac{k_B T_c}{\sqrt{\gamma M}}.$$

The inequality $\lambda \geq 0$ ensures that the power is non-negative, whereas $\sigma \geq \frac{k_B T_c}{\sqrt{\gamma M}}$ ensures that the upper bound is positive. We utilize dimensionless variables

$$x := \frac{\lambda}{\lambda_0}, \quad y := \frac{\sigma_0}{\sigma}$$

for $\sigma_0 := k_B T_c/\sqrt{\gamma M}$, $\lambda_0 := M/k_B T_c$, and re-write (3.64) and the constraints,

$$P = M f(x, y), \quad 0 \leq x \leq g(y), \quad 0 < y \leq 1$$

where $f(x, y) = \frac{\Delta T}{2T_c} x - \frac{x^2}{y^2}$, $g(y) = y - y^2$. As long as $y \leq y_0$, where $y_0 = \frac{1}{1+\frac{\Delta T}{4T_c}}$, the unconstrained maximizer

$$x^*(y) = \text{argmax}_x \ f(x, y) = \frac{\Delta T}{4T_c} y^2$$

satisfies the constraint $x^*(y) \leq g(y)$. When $y_0 < y \leq 1$, the maximizer is at $x = g(y)$. Hence,

$$\max_{x \leq y - y^2} f(x, y) = \begin{cases} \frac{(\Delta T)^2}{16 T_c^2} y^2, & 0 < y \leq y_0 \\ \frac{\Delta T}{2T_c}(y - y^2) - (1 - y)^2, & y_0 \leq y \leq 1 \end{cases}.$$

Maximizing the expressions in the two cases over y gives

$$\max\left\{\left(\frac{\Delta T}{3T_c+T_h}\right)^2, \frac{(\Delta T)^2}{8T_c(T_c+T_h)}\right\} = \frac{(\Delta T)^2}{8T_c(T_c+T_h)}.$$

This is achieved for

$$\sigma = \frac{k_B T_c}{\sqrt{\gamma M}}\frac{2(T_h+T_c)}{(T_h+3T_c)},\ \lambda = \frac{M}{k_B T_c}\frac{(T_h+3T_c)(T_h-T_c)}{4(T_h+T_c)^2}. \tag{3.66}$$

The lower-bound also holds in vector setting by extending this argument and considering a d-dimensional Gaussian distributions with independent components. □

This final result is universal as it does not depend on the choice of ρ_a and ρ_b, unlike (3.59). Moreover, the bounds in this final result are especially appealing in that it the depend on the ratio T_h/T_c of the absolute temperatures of the two heat baths.

Remark 3.6.4. *It is noted that the upper bound in* (3.62) *on achievable power under the constraint* (3.61) *does not depend on* t_{cycle}*, whereas our construction for achieving the lower bound ensures that the bound is approached as* $t_{\text{cycle}} \to 0$.

Remark 3.6.5. *In the proof of Theorem 3.6.1, an operating point has been constructed to ensure that power equal the lower bound in* (3.62) *can be achieved. The parameters are given in equation* (3.66) *in the Appendix. For this operating point, which corresponds to maximal power constrained by* (3.61)*, the efficiency turns out to be*

$$\eta = \frac{T_h - T_c}{T_h + T_c}.$$

It is interesting to note that

$$\eta_{SS} \le \eta_{CA} \le \eta \le \eta_C,$$

where η_{SS} is the efficiency given in (3.34), $\eta_{CA} = 1 - \sqrt{T_c/T_h}$ is the Curzon-Ahlborn efficiency, and $\eta_C = 1 - T_c/T_h$ is the Carnot Efficiency. Furthermore, η_{CA}, η and η_C tend to 1 as $T_c \to 0$, while $\eta_{SS} \to 2/3$. Interestingly, that η may be larger than η_{SS} is due to the fact it is obtained under an added constraint on the controlling potential, that seeks to maximize power, as compared to η_{SS}; the increase in efficiency is consistent with the inherent trade-off between power and efficiency.

3.7 Concluding remarks

The present work focused on quantifying the maximal power that can be drawn by a Carnot-like heat engine operating by alternating contact with two heat reservoirs and modeled by stochastic overdamped Langevin dynamics driven by the time dependent potential. The framework that the work is based on is that of Stochastic Thermodynamics [9–13], which allows quantifying energy and heat exchange by individual particles in a thermodynamic ensemble, to be subsequently averaged, so as to quantify performance of the thermodynamic process as a whole. A physically reasonable bound is derived, which is shown to be reached within a specified factor, both depending on the ratio T_h/T_c of the absolute temperatures of the two heat baths, hot and cold, respectively. The present work is quite distinct from earlier results, within a similar framework, which is however restricted to Gaussian states. Conditions that suggest non-physical conclusions are highlighted, and a suitable constraint on the controlling potential is brought forth that underlies our analysis.

In the past few decades, there have been several attempts to quantify efficiency mainly, but also power, of thermodynamic processes operating in Carnot-like manner. It is fair to say that there has been neither a consensus on the type of assumptions that have been used by previous authors, and thereby, nor full consistency of the results. This is to be expected, since finite-period operation and finite-time thermodynamic transitions require substance/engine

dependent assumptions to capture the complexity of heat transfer in non-equilibrium states. Thus, estimated bounds may never reach the "universality" of the celebrated Carnot efficiency. They are expected to provide physical insights and guidelines for engineering design. Thus, it will be imperative that these estimates be subject to experimental testing. The notable feature of our conclusions as compared to earlier works is that the expressions we derive are given in the form of ratio of absolute temperatures–a physically suggestive feature.

The present work follows a long line of contributions within the control field to draw links between thermodynamics and control, see e.g., [70–75]. More recently, important insights have linked the Wasserstein distance of optimal mass transport, which itself is a solution to a stochastic control problem, to the dissipation mechanism in stochastic thermodynamics [11, 51–54]. Indeed, the Wasserstein metric takes the form of an action integral and arises naturally in the energy balance of thermodynamic transitions. This fact has been explored and developed for the overdamped Langevin dynamics studied herein. Whether similar conclusions can be drawn for underdamped Langevin dynamics remains an open research direction at present. Furthermore, much work remains to reconcile and compare alternative viewpoints and models for thermodynamic processes including those based on the Boltzmann equation.

Besides the potentially intrinsic value of the analysis and bounds that have been derived, it is hoped that the control-theoretic aspect of the problem to optimize Carnot-like cycling of thermodynamic process has been sufficiently highlighted, and that this work will serve to raise attention on this important and foundational topic to the control community.

Chapter 4

Harvesting energy from a sinusoidally driven thermodynamic engine

4.1 Introduction

As mentioned earlier, the natural processes in living organisms do not follow Carnot setting. It is often mediated by continuous processes and energy differentials, whereas the Carnot cycle reflects the switching mechanics of an idealized engine. Therefore, in this and also the following chapter, we put forth stochastic models for non-equilibrium thermodynamic processes in contact with a heat source having periodically and continuously varying temperature. It then explores the question of how to optimize energy harvesting, both in terms of power and efficiency. In the current chapter, we will study power extracted from sinusoidally driven heat bath, which is a stochastic control problem of a somewhat non-traditional nature; the coupling between controlling potential and heat bath renders the models nonlinear. Expressions in closed form are not possible. Thus, we resort to approximation and numerical verification for limiting cases. Conclusions are drawn as to the nature of optimal operation.

Related work treating thermodynamic systems in the linear response regime can be found in [76].

The main contributions of our paper are to be viewed within the context of stochastic control. The motivation and inspiration come from non-equilibrium thermodynamics processes. Natural processes would somehow self-organize, to match driving potentials and optimize efficiency and power, which remains speculatory at present. Specific physical processes need to be modeled, validated, and compared before any definitive statement is made.

The paper develops as follows. In Section 4.2 we present certain basic stochastic model of thermodynamic processes. Section 4.3 details our results for sinusoidal control gain and 4.4 extends the conclusion with a more general control gain, and in Section 4.5 we discuss future directions and open questions.

4.2 Mathematical model

Consider the stochastic dynamics as

$$\gamma dX_t = -\nabla U(t, X_t)dt + \sqrt{2k_B T(t)\gamma}dB_t,$$

which represents a standard model for molecular systems interacting with a thermal environment. Throughout, X_t denotes the location of a particle, and without loss of generality, we assume the mean of X_t is zero. $U(t, x)$ denotes the time-varying potential that can be externally controlled, γ is the viscosity coefficient, k_B is the Boltzmann constant, and B_t is the Brownian motion describing thermal fluctuation, and

$$T(t) = T_0 + T_1 \cos(\omega t), \;\; T_0 > T_1 \geq 0, \tag{4.1}$$

is the temperature of a fluctuating heat bath. We restrict our attention to the case where the potential is quadratic $U(t, X_t) = \frac{1}{2}q(t)X_t^2$.

The variance of X_t, $\Sigma(t) := \mathbb{E}\{X_t^2\}$, satisfies the following differential Lyapunov equation

$$\dot{\Sigma}(t) = -\frac{2}{\gamma}q(t)\Sigma(t) + \frac{2}{\gamma}k_B T(t). \tag{4.2}$$

Then the power that is extracted by the interaction of the stochastic system with the heat bath during one period is defined as

$$\text{power} := -\frac{\omega}{2\pi}\mathbb{E}\left\{\int_0^{\frac{2\pi}{\omega}} \dot{q}(t)\frac{\partial U}{\partial q}(t, X_t)dt\right\} = -\frac{\omega}{4\pi}\int_0^{\frac{2\pi}{\omega}} \dot{q}(t)\Sigma(t)dt.$$

We will consider the following optimization problem

$$\max\{-\frac{\omega}{4\pi}\int_0^{\frac{2\pi}{\omega}} \dot{q}(t)\Sigma(t)dt \mid \text{ subject to } (4.1), (5.8)\}, \tag{4.3}$$

in two different cases. First, we will consider the optimization problem (4.3) under the assumption that the potential has a sinusoidal intensity: $q(t) = q_0 + q_1\cos(\omega t + \phi)$. Next, we will explore the same problem for the general periodic intensity $q(t)$.

4.3 Sinusoidal control gain

Consider the potential with a sinusoidal intensity

$$q(t) = q_0 + q_1\cos(\omega t + \phi), \quad q_0 > 0, \tag{4.4}$$

then the power extracted during one period is given as the following

$$\text{power} := -\frac{\omega}{4\pi}\int_0^{\frac{2\pi}{\omega}} \dot{q}(t)\Sigma(t)dt = \frac{\omega^2 q_1}{4\pi}\int_0^{\frac{2\pi}{\omega}} \sin(\omega t+\phi)\Sigma(t)dt. \tag{4.5}$$

From (4.2), we know the variance $\Sigma(t)$ will evolve according to

$$\dot{\Sigma}(t) = -\frac{2}{\gamma}\left[q_0 + q_1\cos(\omega t+\phi)\right]\Sigma(t) + \frac{2}{\gamma}k_B T(t). \tag{4.6}$$

4.3.1 Fundamental limits for the power

The problem we will study in this section is as the following:

Problem 4.3.1. *Determine the phase shift ϕ in the control law $q(t)$ that maximize Eq.* (4.5) *subject to Eq.* (4.4) *and* (4.6) *and the boundary condition $\Sigma(0) = \Sigma(\frac{2\pi}{\omega})$.*

Proposition 4.3.1. *There exists a unique attractive periodic solution for Eq.* (4.6).

Proof. Solving above linear non-homogeneous ODE gives the closed form gives

$$\begin{aligned}\Sigma(t) = \Sigma_0 \exp\left\{-\frac{2q_1}{\gamma\omega}\sin(\omega t+\phi) - \frac{2}{\gamma}q_0 t + \frac{2q_1}{\gamma\omega}\sin(\phi)\right\}\\ + \exp\left\{-\frac{2q_1}{\gamma\omega}\sin(\omega t+\phi) - \frac{2}{\gamma}q_0 t\right\}\int_0^t \frac{2k_B}{\gamma}T(t)\exp\left\{\frac{2q_1}{\gamma\omega}\sin(\omega\tau+\phi) + \frac{2}{\gamma}q_0\tau\right\}d\tau.\end{aligned} \tag{4.7}$$

In order to have the $\frac{2\pi}{\omega}$-periodic solution, $\Sigma(0) = \Sigma(\frac{2\pi}{\omega})$, which gives

$$\Sigma(0) = \frac{\frac{2k_B}{\gamma}\exp\left\{-\frac{q_1}{\gamma\omega}\sin(\phi)\right\}}{\exp\{\frac{4\pi}{\gamma\omega}q_0\} - 1}\int_0^{\frac{2\pi}{\omega}} (T_0 + T_1\cos(\omega t))\exp\left\{\frac{2q_1}{\gamma\omega}\sin(\omega t+\phi) + \frac{2}{\gamma}q_0 t\right\}dt. \tag{4.8}$$

Attractiveness

In addition, $\exp\left\{\int_0^{\frac{2\pi}{\omega}} -2\left[q_0 + q_1\cos(\omega t+\phi)\right]\right\}dt < 1$ will ensure the stability of the solution

according to Floquet Theory.

Thus, Eq. (4.6) admits a stable and unique periodic solution, which is given in the form of Eq. (4.7) with the initial condition as Eq. (4.8). □

Theorem 4.3.2. *Problem* 1 *has a unique minimizer. Thus power has a unique maximizer* $\phi = \frac{\pi}{2} - \arctan(\frac{\gamma\omega}{2q_0})$.

Proof. Define:

$$q(t) = \sum_{n=-\infty}^{\infty} a_n e^{in\omega t}, \quad T(t) = \sum_{n=-\infty}^{\infty} b_n e^{in\omega t}, \quad \Sigma(t) = \sum_{n=-\infty}^{\infty} c_n e^{in\omega t}.$$

Eq (4.6) gives

$$\dot{\Sigma} = \sum_{n=-\infty}^{\infty} in\omega c_n e^{in\omega t} = -\frac{2}{\gamma} \sum_{n=-\infty}^{\infty} a_n e^{in\omega t} \sum_{n=-\infty}^{\infty} c_n e^{in\omega t} + \frac{2k_B}{\gamma} \sum_{n=-\infty}^{\infty} b_n e^{in\omega t}.$$

From $q(t) = q_0 + q_1 \cos(\omega t + \phi) q_0 + q_1 = \frac{e^{i(\omega t+\phi)} + e^{-i(\omega t+\phi)}}{2}$, we have

$$\sum_{n=-\infty}^{\infty} a_n e^{in\omega t} = q_0 + q_1 \cos(\omega t + \phi),$$

and

$$a_0 = q_0, \quad a_1 = \frac{q_1}{2} e^{i\phi}, \quad a_{-1} = \frac{q_1}{2} e^{-i\phi},$$

Doing the similar analysis for $T(t)$ gives

$$b_0 = T_0, \quad b_1 = \frac{T_1}{2}, \quad b_{-1} = \frac{T_1}{2}, \quad b_n = 0, \ n \neq 0, -1, 1.$$

Therefore, we have

$$\dot{\Sigma} = -\frac{2}{\gamma}\left(a_0 + a_1 e^{i\omega t} + a_{-1} e^{-i\omega t}\right) \sum_{n=-\infty}^{\infty} c_n e^{in\omega t} + \frac{2k_B}{\gamma}\left(b_0 + b_1 e^{i\omega t} + b_{-1} e^{-i\omega t}\right),$$

and

$$in\omega c_n = -\frac{2}{\gamma}\left(a_0 c_n + a_1 p_{n-1} + a_{-1} p_{n+1}\right) + \frac{2k_B}{\gamma} b_n$$

with

$$c_n = \frac{-\frac{2}{\gamma}a_1}{in\omega + \frac{2}{\gamma}a_0} c_{n-1} + \frac{-\frac{2}{\gamma}a_{-1}}{in\omega + \frac{2}{\gamma}a_0} c_{n+1} + \frac{\frac{2k_B}{\gamma}}{in\omega + \frac{2}{\gamma}a_0} b_n \tag{4.9}$$

matricially expressed as: $c = Ac + b$. The solution is given by the Von-Neumann series

$$c = \sum_{k=0}^{\infty} A^k b. \tag{4.10}$$

The truncated series $c^m = \sum_{k=0}^{m} A^k b$ up to the order $m = 2$ is shown in Table 4.1. In time domain, the approximated solution is:

$$\begin{aligned}
\Sigma^2(t) =& c_0^{(2)} + 2\mathrm{Re}(c_1^{(2)} e^{i\omega t}) + 2\mathrm{Re}(c_2^{(2)} e^{i2\omega t}) \\
=& \frac{k_B T_0}{q_0} - 2\mathrm{Re}\left(\frac{q_1 e^{i\phi}}{2q_0} \frac{k_b T_1 e^{i\theta}}{\sqrt{4q_0^2 + \gamma^2\omega^2}}\right) + 2\mathrm{Re}\left(\frac{\frac{-k_B T_0 q_1}{q_0} e^{i\phi} + k_B T_1}{\sqrt{4q_0^2 + \gamma^2\omega^2}} e^{i(\omega t - \theta)}\right) \\
&+ 2\mathrm{Re}\left(\frac{-k_B T_1 q_1}{2\sqrt{\gamma^2\omega^2 + q_0^2}\sqrt{\gamma^2\omega^2 + 4q_0^2}} e^{i(2\omega t + \phi - \theta - \theta_2)}\right) \\
=& \frac{k_B T_0}{q_0} - \frac{k_B T_1 q_1}{q_0\sqrt{\gamma^2\omega^2 + 4q_0^2}}\cos(\phi + \theta) - \frac{2k_B T_0 q_1}{q_0\sqrt{\gamma^2\omega^2 + 4q_0^2}}\cos(\omega t + \phi - \theta) \\
&+ \frac{2k_B T_1}{\sqrt{\gamma^2\omega^2 + 4q_0^2}}\cos(\omega t - \theta) - \frac{k_B T_1}{\sqrt{\gamma^2\omega^2 + 4q_0^2}\sqrt{\gamma^2\omega^2 + q_0^2}}\cos(2\omega t + \phi - \theta - \theta_2),
\end{aligned}$$

where $\theta = \arctan(\frac{\gamma\omega}{2q_0})$, and $\theta_2 = \arctan(\frac{\gamma\omega}{q_0})$.

n	iteration $m=0$	iteration $m=1$	iteration $m=2$
2	$c_2^{(0)}=0$	$c_2^{(1)}=0$	$c_2^{(2)}=\frac{-a_1}{\gamma i\omega+a_0}c_1^{(1)}$
1	$c_1^{(0)}=0$	$c_1^{(1)}=\frac{2k_B}{i\gamma\omega+2a_0}b_1$	$c_1^{(2)}=\frac{-2a_1}{i\gamma\omega+2a_0}c_0^{(1)}+\frac{2k_B}{i\gamma\omega+2a_0}b_1$
0	$c_0^{(0)}=0$	$c_0^{(1)}=\frac{k_Bb_0}{a_0}$	$c_0^{(2)}=-\frac{a_1}{a_0}c_{-1}^{(1)}-\frac{a_{-1}}{a_0}c_1^{(1)}+\frac{k_B}{a_0}b_0$
-1	$c_{-1}^{(0)}=0$	$c_{-1}^{(1)}=\frac{2k_B}{-i\omega\gamma+2a_0}b_{-1}$	$c_{-1}^{(2)}=\frac{-2a_{-1}}{-i\gamma\omega+2a_0}c_0^{(1)}+\frac{2k_B}{-i\gamma\omega+2a_0}b_{-1}$
-2	$c_{-2}^{(0)}=0$	$c_{-2}^{(1)}=0$	$c_{-2}^{(2)}=\frac{-a_{-1}}{\gamma i\omega+a_0}c_{-1}^{(1)}$

Table 4.1 – The Fourier coefficients for first order approximation of the variance

For the power extracted during one period is computed as

$$\begin{aligned}
\text{power} &= -\frac{\omega}{4\pi}\int_0^{\frac{2\pi}{\omega}}\dot{q}(t)\Sigma(t)dt = \frac{\omega}{4\pi}\int_0^{\frac{2\pi}{\omega}}\sum_{n=-\infty}^{\infty}in\omega c_n e^{in\omega t}\sum_{n=-\infty}^{\infty}a_n e^{in\omega t}dt \\
&= \frac{\omega}{4\pi}\int_0^{\frac{2\pi}{\omega}}(a_0+a_1e^{i\omega t}+a_{-1}e^{-i\omega t})\sum_{n=-\infty}^{\infty}in\omega c_n e^{in\omega t}dt \\
&= \frac{1}{2}(-i\omega a_1c_{-1}+i\omega a_{-1}c_1) = \frac{i\omega}{2}(-a_1c_{-1}+a_{-1}c_1) \\
&= i\omega\mathrm{Im}\{a_{-1}c_1\}i = -\omega\mathrm{Im}\{a_{-1}c_1\} \\
&= \frac{\omega q_1k_B}{2(4q_0^2+\gamma^2\omega^2)q_0}\left(-q_1T_0\gamma\omega+\cos(\phi)\gamma\omega T_1q_0+2\sin(\phi)T_1q_0^2\right) \\
&= -\frac{k_B\gamma T_0q_1^2\omega^2}{2(4q_0^2+\gamma^2\omega^2)q_0}+\frac{k_B\omega T_1q_1}{2(4q_0^2+\gamma^2\omega^2)}\left(\gamma\omega\cos(\phi)+2q_0\sin(\phi)\right) \\
&= -\frac{k_B\gamma T_0q_1^2\omega^2}{2(4q_0^2+\gamma^2\omega^2)q_0}+\frac{k_B\omega T_1q_1}{2\sqrt{4q_0^2+\gamma^2\omega^2}}\sin(\phi+\theta),
\end{aligned} \tag{4.11}$$

where $\theta=\arctan(\frac{\gamma\omega}{2q_0})$.

The maximum is achieved at $\phi=\frac{\pi}{2}-\theta$. Particularly, when $\frac{\gamma\omega}{2q_0}=1$, the maximum power will be obtained when $\phi=\frac{\pi}{4}$. The power can also be expressed as

$$\text{power} = \frac{\mathrm{k_B T_0 q_0}}{2\gamma}\left(-\frac{\omega_\mathrm{r}^2}{1+\omega_\mathrm{r}^2}\mathrm{q_r^2}+\frac{\omega_\mathrm{r}\mathrm{T_r}\mathrm{q_r}}{\sqrt{1+\omega_\mathrm{r}^2}}\sin(\phi+\theta)\right), \tag{4.12}$$

where

$$\omega_r = \frac{\omega}{\frac{2q_0}{\gamma}}, \quad q_r = \frac{q_1}{q_0}, \quad T_r = \frac{T_1}{T_0}, \quad \theta = \arctan(\omega_r)$$

are dimensionless variables. ω_r can be interpreted as the ratio of the applied frequency to the natural frequency of the system. The maximum power, over ϕ, is achieved at $\phi = \frac{\pi}{2} - \theta$. The maximum power is then equal to

$$\max_{\phi} \text{power} = \frac{\mathrm{k_B T_0 q_0}}{2\gamma} \frac{\omega_\mathrm{r}}{\sqrt{1+\omega_\mathrm{r}^2}} \left(-\frac{\omega_\mathrm{r}}{\sqrt{1+\omega_\mathrm{r}^2}} \mathrm{q_r^2} + \mathrm{T_r q_r} \right). \tag{4.13}$$

This explains the quadratic curve in Figure 4.2. For fixed ω_r and T_r, the maximum power, over q_r, is achieved at

$$q_r = \frac{T_r}{2\omega_r}\sqrt{1+\omega_r^2},$$

and the maximum power is equal to

$$\text{power} = \frac{\mathrm{k_B T_0 q_0}}{8\gamma} \mathrm{T_r^2}.$$

Define $T_h := T_0 + T_1$ and $T_c := T_0 - T_1$, then

$$\text{power} = \frac{\mathrm{k_B T_0 q_0}}{8\gamma} (\frac{\mathrm{T_h - T_c}}{\mathrm{T_h + T_c}})^2.$$

□

4.4 Generalization for the control

In this section, we will consider the temperature fluctuation about the mean value T_0 as follows:

$$T(t) = T_0 + \epsilon T_1 \cos(\omega t), \tag{4.14}$$

and study the potential for drawing the power by applying periodic control

$$q(t) = q_0 + \epsilon u(t), \tag{4.15}$$

where $\epsilon u(t)$ is the deviation of the control input from the nominal value q_0.

4.4.1 Fundamental limits for the power

In this part, the limit case when perturbation ϵ is small is studied. We seek to determine a control input that maximizes power, that is,

$$\begin{aligned} \max_{u} \quad & -\frac{\omega}{4\pi}\int_0^{\frac{2\pi}{\omega}} \epsilon \dot{u}(t)\Sigma(t)\mathrm{d}t \\ \text{subject to} \quad & \dot{\Sigma}(t) = -\frac{2}{\gamma}(q_0 + \epsilon u(t))\Sigma(t) + \frac{2}{\gamma}k_B T(t), \\ & u(0) = u(\frac{2\pi}{\omega}). \end{aligned} \tag{4.16}$$

To this end, we carry out a perturbation analysis about $\epsilon = 0$. The variance $\Sigma(t)$ is expressed as

$$\Sigma(t) = \sum_{k=0}^{\infty} \epsilon^k \Sigma^{(k)}(t), \tag{4.17}$$

where $\Sigma^{(k)}(t)$ solves the Lyapunov equation (4.6) for ϵ^k order. In particular, the leading two terms satisfy

$$\dot{\Sigma}^{(0)}(t) = -\frac{2q_0}{\gamma}\Sigma^{(0)}(t) + \frac{2k_BT_0}{\gamma} \tag{4.18}$$

$$\dot{\Sigma}^{(1)}(t) = -\frac{2q_0}{\gamma}\Sigma^{(1)}(t) - \frac{2u(t)}{\gamma}\Sigma^{(0)}(t) + \frac{2k_BT_1}{\gamma}\cos(\omega t). \tag{4.19}$$

We truncate all but the first two terms in the objective function of the optimal control problem, and consider the problem to optimize

$$\max_u \ -\frac{\omega}{4\pi}\int_0^{\frac{2\pi}{\omega}} \left(\epsilon\dot{u}(t)\Sigma^{(0)}(t) + \epsilon^2\dot{u}(t)\Sigma^{(1)}(t)\right) \mathrm{d}t. \tag{4.20}$$

The solution of the optimal control problem (4.20) can now be expressed as follows.

Theorem 4.4.1. *Consider the optimal control problem* (4.20)*. The optimal control law is*

$$u^*(t) = q_1^*\cos(\omega t - \phi^*), \tag{4.21}$$

where

$$q_1^* = \frac{q_0 g T_1}{2\gamma\omega T_0}, \quad \phi^* = \angle(\gamma\omega - i2q_0), \tag{4.22a}$$

with $g = \sqrt{\gamma^2\omega^2 + 4q_0^2}$*, giving power output*

$$\text{power} = \epsilon^2\frac{k_Bq_0T_1^2}{8\gamma T_0} + O(\epsilon^3),$$

and the expression for the variance $\Sigma(t)$ *up to first order in* ϵ *is:*

$$\Sigma(t) = \frac{k_BT_0}{q_0} + \epsilon\frac{k_BT_1}{\gamma\omega}\sin(\omega t) + O(\epsilon^2).$$

Proof. Solving (4.18), together with the periodic condition for $\Sigma(t)$, we can obtain that

$\Sigma^{(0)}(t)$ is constant and satisfies

$$\Sigma^{(0)}(t) = \frac{k_B T_0}{q_0},$$

therefore, the optimization problem (4.20) becomes

$$\begin{aligned} \max_{u} \quad & -\frac{\omega}{4\pi}\int_0^{\frac{2\pi}{\omega}} \epsilon^2 \dot{u}(t)\Sigma^{(1)}(t)\mathrm{d}t, \\ \text{subject to} \quad & \dot{\Sigma}^{(1)}(t) = -\frac{2q_0}{\gamma}\Sigma^{(1)}(t) - \frac{2u(t)}{\gamma}\Sigma^{(0)}(t) + \frac{2k_B T_1}{\gamma}\cos(\omega t) \\ & u(0) = u(\frac{2\pi}{\omega}). \end{aligned} \tag{4.23}$$

Define $\mathcal{L} := \dot{u}(t)\Sigma^{(1)}(t) + \lambda(\dot{\Sigma}^{(1)}(t) + \frac{2q_0}{\gamma}\Sigma^{(1)}(t) + \frac{2u(t)}{\gamma}\Sigma^{(0)}(t) - \frac{2k_B T_1}{\gamma}\cos(\omega t))$ with λ lagrangian multiplier, then the optimal u and $\Sigma^{(1)}(t)$ should satisfy the following Euler Lagrange equation

$$\frac{\partial \mathcal{L}}{\partial u} - \frac{\mathrm{d}(\frac{\partial \mathcal{L}}{\partial \dot{u}})}{dt} = \frac{2\lambda k_B T_0}{\gamma q_0} - \dot{\Sigma}^{(1)} = 0 \tag{4.24}$$

$$\frac{\partial \mathcal{L}}{\partial \Sigma^{(1)}} - \frac{\mathrm{d}(\frac{\partial \mathcal{L}}{\partial \dot{\Sigma}^{(1)}})}{dt} = \dot{u} + \frac{2\lambda q_0}{\gamma} - \dot{\lambda} = 0. \tag{4.25}$$

Then substituting Eq. (4.24) into Eq. (4.25), we have

$$\dot{u} + \frac{q_0^2}{k_B T_0}\dot{\Sigma}^{(1)} - \frac{\gamma q_0}{2k_B T_0}\ddot{\Sigma}^{(1)} = 0, \tag{4.26}$$

which together with the Lyapunov equation Eq. (4.18) gives

$$\dot{\Sigma}^{(1)} = -\frac{2q_0}{\gamma}\Sigma^{(1)} - \frac{2u}{\gamma}\frac{k_B T_0}{q_0} + \frac{2k_B T_1}{\gamma}\cos(\omega t). \tag{4.27}$$

and

$$\ddot{\Sigma}^{(1)} = -\frac{k_B T_1 \omega}{\gamma}\sin(\omega t), \tag{4.28}$$

from which, we can easily get (4.21), (4.22) and (4.23). □

4.4.2 Efficiency at maximal power

For the Carnot-like cycle, Q_h is defined as the heat transfer during the time interval that the system is in contact with hot thermal bath $T = T_h$. However, when the temperature profile is continuous, it is not clear to distinguish the heat transfer to the hot bath from the heat transfer to the cold bath. This issue is resolved in the linear response regime by defining Q_h according to

$$Q_h[q] = \int_0^{t_f} \frac{T(t) - T_c}{T_h - T_c} \dot{Q}(t)[q] \mathrm{d}t = \int_0^{t_f} \frac{\cos(\omega t) + 1}{2} \dot{Q}(t)[q] \mathrm{d}t, \tag{4.29}$$

where $\dot{Q}(t)[q]$ is the heat transfer rate given by the sum of the rate of change in energy and power output

$$\dot{Q}(t)[q] = \frac{\mathrm{d}}{\mathrm{d}t} \mathbb{E}\{U(t, X_t)\} + t_f \dot{\mathcal{P}}[q] = \frac{1}{2} q(t) \dot{\Sigma}(t).$$

In this chapter, we will use the definition (4.29) to define Q_h and efficiency. The efficiency is defined to be the ratio between the work output $\mathcal{W} = t_f \text{power}$ and the heat input Q_h : $\eta = \frac{t_f \text{power}}{Q_h}$. The heat is computed as the following:

$$\begin{aligned}
Q_h :&= \frac{1}{4} \int_0^{t_f} (\cos(\omega t) + 1)\, q(t) \dot{\Sigma}(t) dt \\
&= -\frac{1}{4} \int_0^{t_f} \dot{q}(t) \Sigma(t) dt + \frac{1}{4} \int_0^{t_f} \cos(\omega t) q(t) \dot{\Sigma}(t) dt \\
&= \frac{1}{2} t_f \text{power} + \frac{1}{4} \int_0^{t_f} \cos(\omega t) \left(q_0 + \epsilon q_1^* \cos(\omega t - \phi^*)\right) \left(\epsilon \frac{k_B T_1}{\gamma} \cos(\omega t) + O(\epsilon)\right) \mathrm{d}t \\
&= \frac{1}{2} t_f \text{power} + \epsilon \frac{q_0 k_B T_1 t_f}{2\gamma} + O(\epsilon^2).
\end{aligned}$$

Therefore, the efficiency at maximal power can be computed as

$$\begin{aligned}\eta &= \frac{t_f \text{power}}{Q_h} = \frac{t_f \text{power}}{\frac{1}{2} t_f \text{power} + \epsilon \frac{q_0 k_B T_1 t_f}{2\gamma} + O(\epsilon^2)} = \frac{\epsilon T_1}{T_0} + O(\epsilon^2) \\ &= \frac{T_h - T_c}{T_h + T_c} + O((T_h - Tc)^2) = \frac{\eta_C}{2 - \eta_C} + O((T_h - T_c)^2), \end{aligned} \tag{4.30}$$

where we define $T_h := T_0 + \epsilon T_1,\ T_c := T_0 - \epsilon T_1$, and η_C Carnot efficiency.

Remark 4.4.1. *We note that Carnot efficiency, for quasistatic operation between two heat baths of temperatures T_h and T_c, for hot and cold respectively, is $\eta_C = \frac{T_h - T_c}{T_h}$. Letting $T_h = T_0 + \epsilon T_1$ and $T_c = T_0 - \epsilon T_1$, and evaluating the expression for η_C, gives*

$$\frac{2\epsilon T_1}{T_0 + \epsilon T_1} = 2\epsilon \frac{T_1}{T_0} + O(\epsilon^2).$$

4.4.3 Numerical validation

In this section, we provide numerical validation and insight into the effect of higher order terms in the expansions. Specifically, we consider a sinusoidal control input and use Fourier representations to numerically solve the Lyapunov equation and obtain expressions for the power. Our interest mainly focuses on how maximal power depends on the amplitude and phase of the control, and on how efficiency at maximal power depends on the amplitude of the temperature fluctuations of the heat bath. Starting from the choice of control

$$q(t) = q_0 + q_1 \cos(\omega t - \phi) = a_{-1} e^{-i\omega t} + a_0 + a_1 e^{i\omega t},$$

where $a_0 = q_0$ and $a_1 = \bar{a}_{-1} = \frac{1}{2} q_1 e^{-i\phi}$, the power drawn can be expressed as

$$\text{power} = -\frac{i\omega}{2} \left(a_1 c_{-1} - a_{-1} c_1 \right), \tag{4.31}$$

where c_n is the nth Fourier coefficient of covariance Σ. The Fourier coefficients can be obtained by expressing the Lyapunov equation in the Fourier domain, to obtain a set of linear coupled equations for the various terms. We truncate and keep the first 100 modes, and solve the resulting finite-dimensional problem for a range of values for q_1 and ϕ.

The numerical result for the power is depicted in Figure 4.1, Figure 4.2 and Figure 4.3. It is observed that the pair of values (ϕ, q_1) that maximize the power are close to the analytical expressions obtained by ignoring second-order terms. The model parameters used to obtain the numerical result are presented in Table 4.2.

Notation vs. value	notation	value
perturbation	ϵ	1
viscosity coefficient	γ	1
frequency	ω	2
temperature	T_1	0.5
temperature	T_0	1
nominal gain	q_0	1

Table 4.2 – Parameters selected in the simulations.

The thermal efficiency is also evaluated numerically. The efficiency as a function of the temperature fluctuation T_1 is depicted in Figure 4.4. The numerical result is compared with the analytical result obtained up to first-order approximation. It is observed that the analytical expression captures the behavior of the efficiency for small values of T_1 for the model. We also plot the optimal control law and temperature profile, interestingly which shows that the optimal control law is sinusoidal function when the temperature is sinusoidal, but with different phases, which is shown in Figure 4.5.

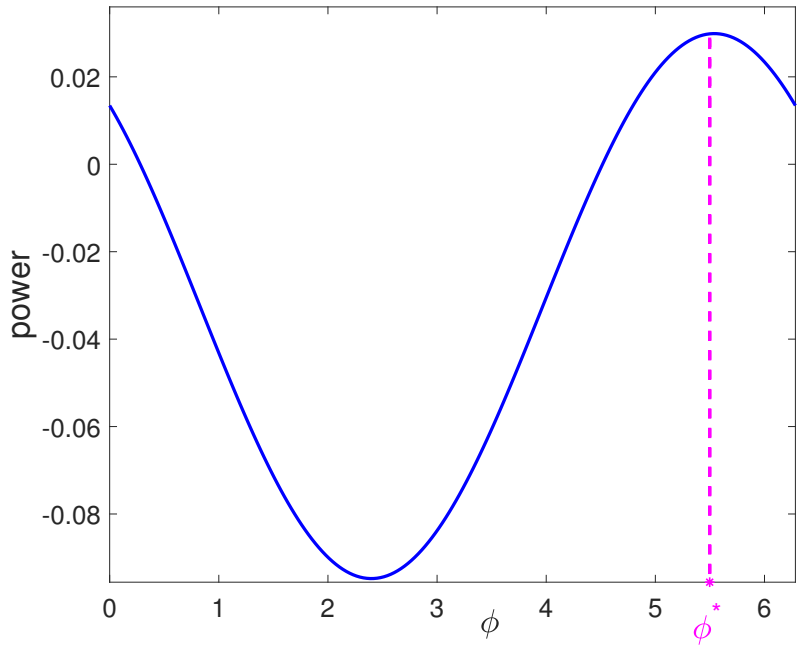

Figure 4.1 – Numerical evaluation of the power output as a function of the control phase ϕ as described in Section 4.4.3. "ϕ^*" is the optimal control parameter computed analytically using (4.22).

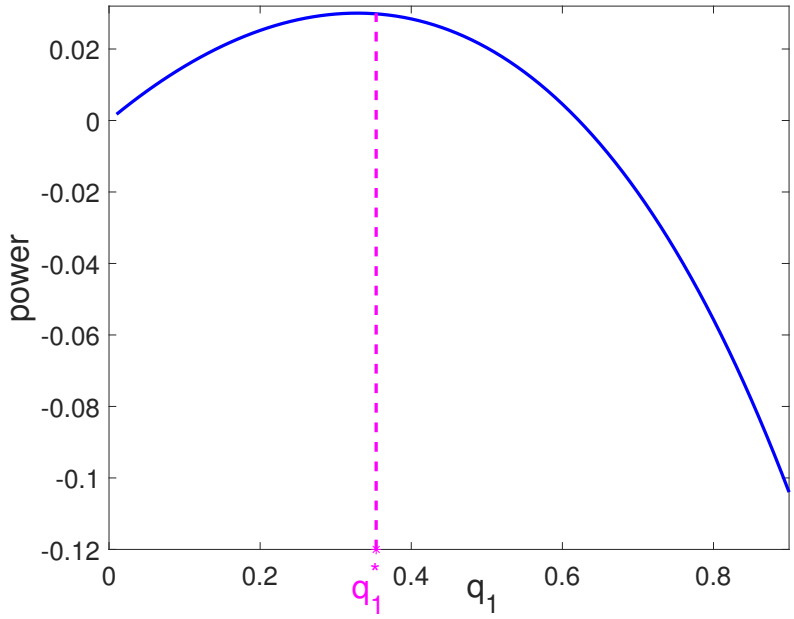

Figure 4.2 – Numerical evaluation of the power output as a function of the control phase q_1 as described in Section 4.4.3. "q_1^*" is the optimal control parameter computed analytically using (4.22).

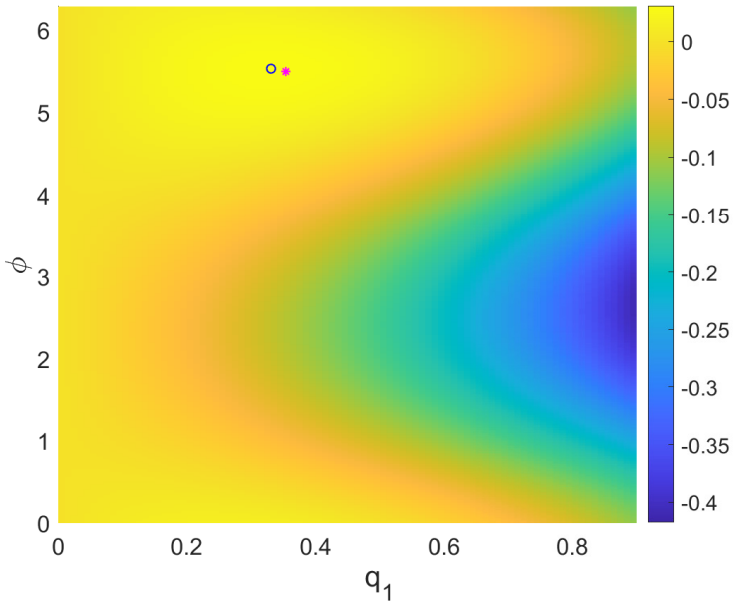

Figure 4.3 – Numerical evaluation of the power output as a function of the control phase ϕ and control amplitude q_1 as described in Section 4.4.3. The points marked by "$\circ$" and "$*$" correspond to optimal control parameters. The first ("$\circ$") was computed numerically, and the second ("$*$") analytically using (4.22).

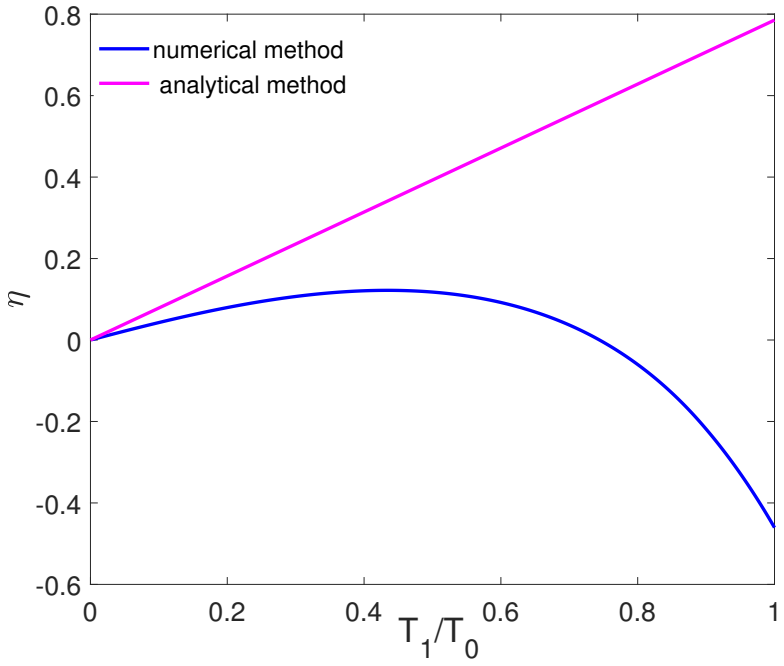

Figure 4.4 – Numerical evaluation of efficiency as a function of temperature fluctuation T_1, at maximum power. The numerical result is compared with the analytical expressions derived using first-order approximations in (4.30).

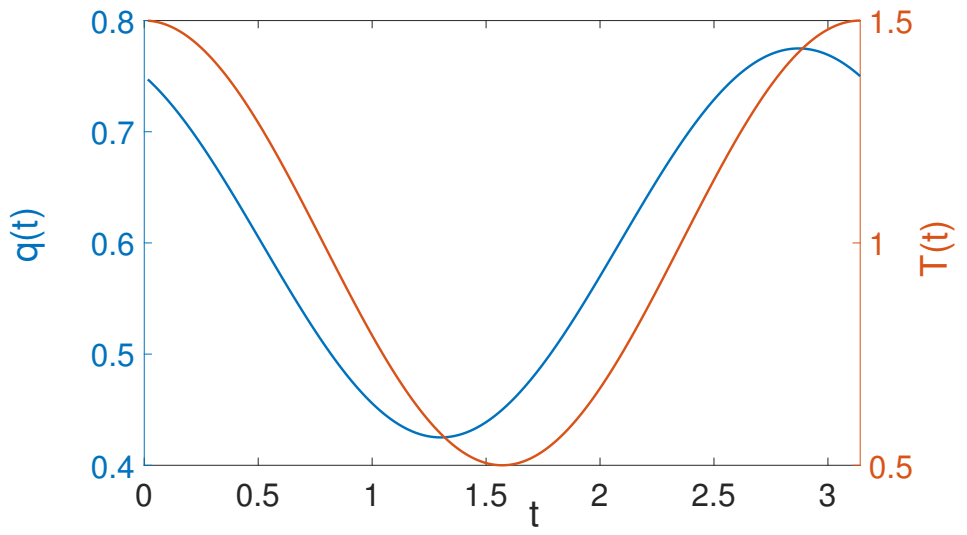

Figure 4.5 – The comparison between the optimal control law and the temperature profile.

4.5 Concluding remarks

We addressed the question of maximal power and efficiency for thermodynamic processes in contact with a heat bath having periodic continuously varying temperature. Our analysis is approximate and focuses on sinusoidal fluctuations. It is of interest to study the effect of the temperature profile, and properties of optimal controlling potential, in a more general setting.

Chapter 5

Thermodynamic bounds on power under arbitrary temperature profile

5.1 Introduction

In this chapter, we consider the problem of maximal power extracted from a thermodynamic engine in contact with a heat source having periodically and continuously varying temperature, and seek to determine the control input that giving the maximal power in finite time transitions. Under the appropriate constraints, we draw some conclusions on the bounds of maximal power and also the corresponding efficiency. The main contribution of this chapter is that our results give the upper bound of the maximal power any periodically and continuous heat engine can achieve, which is proportional to the average fluctuations in the temperature profile. Moreover, we show the surprising result that the efficiency at maximum power does not depend on the temperature profile, just the maximum and minimum values.

The chapter is developed as follows. In Section 5.2 we introduce certain basic stochastic model of thermodynamic processes, and show some relevant known results. Sections 5.3 and

5.4 detail our results about the bounds of maximal power and corresponding efficiency for our overdamped stochastic model, and in Section 5.6 we discuss future directions and open questions.

5.2 Mathematical model

The Langevin equation

$$\gamma \mathrm{d}X_t = -\nabla U(t, X_t)\mathrm{d}t + \sqrt{2k_B T(t)\gamma}\mathrm{d}B_t. \tag{5.1}$$

This model has been studied extensively when the system undergoes a Carnot cycle [77–79]: the system is in contact with the heat bath of temperature T_h for t_1 amount of time, and then it switches instantaneously to contact with a cold bath of temperature T_c for t_2 amount of time, i.e.

$$T(t) = \begin{cases} T_h, & t \in (0, t_1) \\ T_c, & t \in (t_1, t_1 + t_2). \end{cases}$$

Let $t_f = t_1 + t_2$ denotes the cycle period. The first question of interest is to understand the maximum power that can be extracted from the system by controlling the potential function. The power extracted from the system over a cycle is defined according to

$$\mathcal{P} := -\mathbb{E}\left[\frac{1}{t_f}\int_0^{t_f} \frac{1}{2}\dot{q}(t) X_t^2 \mathrm{d}t.\right] \tag{5.2}$$

Let $P_{\max}$ denote the maximum power that can be achieved by controlling the potential function. The second question we will tackle is that of the efficiency of the system operating at maximum power, denoted by η_{mp}. These problems have also been studied for Carnot

cycle, the results are as follows:

$$P_{\max} = \frac{\left(k_B(T_h - T_c)\right)^2}{16\gamma\sigma_{\min}^2}, \tag{5.3a}$$

$$\eta_{\mathrm{mp}} = \frac{2\left(T_h - T_c\right)}{3T_h + T_c}, \tag{5.3b}$$

where $\sigma_{\min}$ is the minimum value of the variance of X_t over the cycle.

The objective of this paper is to extend those results to arbitrary temperature profile. The motivations come from analysis of biological thermodynamic systems that undergo a continuously changing temperature profile, instead of abrupt change in the Carnot cycle model. In particular, we assume that the temperature profile is of the form:

$$T(t) = \frac{T_h + T_c}{2} + \frac{T_h - T_c}{2}\alpha(t), \tag{5.4}$$

where $\alpha(t)$ is a periodic function of period t_f taking values in $[-1, +1]$. Note that the special case where $\alpha(t) = 1$ for $t \in (0, t_1)$, and $\alpha(t) = -1$ for $t \in (t_1, t_f)$ corresponds to the Carnot cycle. The extension of existing results (5.3a) to an arbitrary temperature profile is:

$$P_{\max} \leq \frac{\left(k_B(T_h - T_c)\right)^2}{16\gamma\sigma_{\min}^2}\mathrm{Var}(\alpha), \tag{5.5}$$

where $\mathrm{Var}(\alpha) := \frac{1}{t_f}\int_0^{t_f}\alpha(t)^2 dt - \left(\frac{1}{t_f}\int_0^{t_f}\alpha(t)dt\right)^2$ is the variation of α in one cycle. The upper bound can be achieved when the total period t_f tends to zero, and the corresponding efficiency is given:

$$\eta_{\mathrm{mp}} = \frac{2\left(T_h - T_c\right)}{3T_h + T_c}. \tag{5.6}$$

Some striking implications of our result are as follows:

1. Assuming $\alpha(t) = \pm 1$, the result for the Carnot cycle is recovered.

2. The maximum power is proportional to the variations in the temperature.

3. Over the set of admissible temperature profiles, maximum power is achieved when the temperature profile is according to Carnot cycle.

4. The efficiency at maximum power η_{mp} does not depend on the temperature profile.

5.3 The analysis for the power and efficiency

In this section, we will present the analysis of the maximum power and the corresponding efficiency.

5.3.1 Analysis of maximal power

In this section, we will study maximizing power for the dynamics (5.1). For this model, the expression for power (5.2) simplifies to

$$\mathcal{P} := -\mathbb{E}\left[\frac{1}{t_f}\int_0^{t_f}\frac{1}{2}\dot{q}(t)X_t^2 \mathrm{d}t.\right] = -\frac{1}{2t_f}\int_0^{t_f}\dot{q}(t)\Sigma(t)dt \tag{5.7}$$

where $\Sigma(t) := \mathbb{E}[X_t^2]$ is governed by the Lyapunov equation:

$$\dot{\Sigma}(t) = -\frac{2q(t)}{\gamma}\Sigma(t) + \frac{2k_B T(t)}{\gamma}. \tag{5.8}$$

The objective is to analyze the maximal power that can be extracted by controlling the coefficient $q(t)$:

$$\mathcal{P}^* = \max_q \ \left\{\mathcal{P}[q], \text{ s.t. } \Sigma(t) \text{ solves (5.8)}, q(0) = q(t_f),\ \Sigma(0) = \Sigma(t_f), \quad \Sigma(t) \geq \sigma_{\min}^2\right\}, \tag{5.9}$$

because without any constraint, the power can become unbounded. The key step is to simplify the expression for the power by a change of variable.

Lemma 5.3.1. *The expression for the power is equal to*

$$\mathcal{P}[q] = J[x] := \frac{k_B(T_h - T_c)}{2t_f} \int_0^{t_f} \left[-\lambda t_f \dot{x}(t)^2 e^{2x(t)} + \alpha \dot{x}(t)\right] \mathrm{d}t, \tag{5.10}$$

where $x(t) := \frac{1}{2}\log(\frac{\Sigma(t)}{\sigma_{min}^2})$, $\lambda := \frac{2\gamma\sigma_{min}^2}{t_f k_B(T_h - T_c)}$ *is a dimensionless constant. Moreover, the problem for maximizing the power is equivalent to*

$$P_{\max} = \max_x \ \left\{J[x], \quad s.t. \quad x(0) = x(t_f), \quad x(t) \geq 0\right\}. \tag{5.11}$$

Proof. Define the function $F(t)$ according to

$$F(t) := \frac{1}{2} q(t)\Sigma(t) - \frac{k_B T(t)}{2} \log(\Sigma(t)).$$

Taking the time derivative yields

$$\begin{aligned}
\frac{\mathrm{d}F(t)}{\mathrm{d}t} &= \frac{1}{2}\dot{q}(t)\Sigma(t) + \frac{1}{2}q(t)\dot{\Sigma}(t) - \frac{1}{2}k_B\dot{T}(t)\log(\Sigma(t)) - \frac{1}{2\Sigma(t)}k_B T(t)\dot{\Sigma}(t) \\
&= \frac{1}{2}\dot{q}(t)\Sigma(t) - \frac{q(t)^2}{\gamma}\Sigma(t) + 2q(t)\frac{k_B T(t)}{\gamma} + \frac{k_B^2 T(t)^2}{\gamma\Sigma(t)} - \frac{1}{2}k_B\dot{T}(t)\log(\Sigma(t)) \\
&= \frac{1}{2}\dot{q}(t)\Sigma(t) - \gamma\left(\frac{q(t)\sqrt{\Sigma(t)}}{\gamma} - \frac{k_B T(t)}{\gamma\sqrt{\Sigma(t)}}\right)^2 - \frac{1}{2}k_B\dot{T}(t)\log(\Sigma(t)) \\
&= \frac{1}{2}\dot{q}(t)\Sigma(t) - \gamma\dot{\sigma}(t)^2 - k_B\dot{T}(t)\log(\sigma(t)),
\end{aligned}$$

where $\sigma(t) := \sqrt{\Sigma(t)}$ and we used $\dot{\sigma}(t) = \frac{q(t)\sqrt{\Sigma(t)}}{\gamma} - \frac{k_B T(t)}{\gamma\sqrt{\Sigma(t)}}$ from (5.8). Integrating over a cycle yields

$$F(t_f) - F(t_0) = -t_f\mathcal{P}[q] - \int_0^{t_f}\left[\gamma\dot{\sigma}(t)^2 + k_B\dot{T}(t)\log(\sigma(t))\right]\mathrm{d}t,$$

where we used the definition of power (5.7). Using $F(0) = F(t_f)$ because of the boundary conditions $\Sigma(0) = \Sigma(t_f)$ and $q(0) = q(t_f)$, yields the following expression for power

$$\mathcal{P}[q] = -\frac{1}{t_f}\int_0^{t_f}\left[\gamma\dot{\sigma}(t)^2 + k_B\dot{T}(t)\log(\sigma(t))\right]\mathrm{d}t.$$

Using the definition $x(t) := \log(\frac{\sigma(t)}{\sigma_{\min}})$ and dimensionless number $\lambda := \frac{2\gamma\sigma_{\min}^2}{t_f k_B(T_h - T_c)}$ concludes

$$\begin{aligned}\mathcal{P}[q] &= -\frac{1}{t_f}\int_0^{t_f}\left[\gamma\sigma_{\min}^2\dot{x}(t)^2e^{2x(t)} + k_B\frac{T_h - T_c}{2}\dot{\alpha}x(t)\right]\mathrm{d}t \\ &= -\frac{k_B(T_h - T_c)}{2t_f}\int_0^{t_f}\left[\lambda t_f\dot{x}(t)^2e^{2x(t)} + \dot{\alpha}(t)x(t)\right]\mathrm{d}t \\ &= -\frac{k_B(T_h - T_c)}{2t_f}\int_0^{t_f}\left[\lambda t_f\dot{x}(t)^2e^{2x(t)} - \alpha(t)\dot{x}(t)\right]\mathrm{d}t,\end{aligned}$$

where the last step follows from integration by parts and periodic boundary conditions.

For maximizing the power, we are allowed to choose any $x(t)$ under two constraints: (i) $x(t)$ is periodic, i.e. $x(0) = x(t_f)$, (ii) $x(t)$ is positive, i.e. $x(t) \geq 0$. The two constraints are due to the definition $x(t) = \frac{1}{2}\log(\frac{\Sigma(t)}{\sigma_{\min}^2})$ and the constraints $\Sigma(0) = \Sigma(t_f)$ and $\Sigma(t) \geq \sigma_{\min}^2$.

□

The following theorem provides bounds for the maximal power.

Theorem 5.3.1. *Consider the over-damped Langevin dynamics* (5.1) *under the temperature profile* (5.4)*. Then, the maximal power* (5.11) *satisfies the upper-bound:*

$$\mathcal{P}^* \leq \frac{k_B^2T_c^2}{16\gamma\sigma_{min}^2}(\frac{T_h}{T_c} - 1)^2\,Var(\alpha). \tag{5.12}$$

And there exists an explicit protocol $q_0(t)$ that achieves the upper-bound within a gap that

converges to zero as $t_f \to 0$, and the optimal protocol and the resulting variance are given by

$$q_0(t) = \frac{k_B}{\sigma^2_{min}} T(t) \exp(-\frac{1}{\lambda t_f} \int_{t_0}^{t} (\alpha(s) - \bar{\alpha})\, \mathrm{d}s) - \frac{k_B(T_h - T_c)}{4\sigma^2_{min}} (\alpha(t) - \bar{\alpha}) \tag{5.13}$$

and

$$\Sigma_0(t) = \sigma^2_{min} \exp(\frac{1}{\lambda t_f} \int_{t_0}^{t} (\alpha(s) - \bar{\alpha})\, \mathrm{d}s), \tag{5.14}$$

where t_0 is a time so that $\int_{t_0}^{t} (\alpha(s) - \bar{\alpha})\, \mathrm{d}s \geq 0$ for all t.

Proof. The proof is based on the equivalent formulation of the power according to Lemma 5.3.1. The upper-bound follows by proving an upper-bound for $J[x]$ for all x.

$$\begin{aligned}
J[x] &= \frac{k_B(T_h - T_c)}{2t_f} \int_0^{t_f} -\lambda t_f \dot{x}(t)^2 e^{2x(t)} + \alpha(t)\dot{x}(t)\mathrm{d}t \\
&\leq \frac{k_B(T_h - T_c)}{2t_f} \int_0^{t_f} -\lambda t_f \dot{x}(t)^2 + \alpha(t)\dot{x}(t)\mathrm{d}t \\
&= -\frac{k_B(T_h - T_c)}{2t_f} \int_0^{t_f} \lambda t_f \left[\dot{x}(t) - \frac{1}{2\lambda t_f}(\alpha(t) - \bar{\alpha})\right]^2 \mathrm{d}t + \frac{k_B(T_h - T_c)}{8\lambda t_f^2} \int_0^{t_f} (\alpha(t) - \bar{\alpha})^2 \mathrm{d}t \\
&\leq \frac{k_B(T_h - T_c)}{8\lambda t_f^2} \int_0^{t_f} (\alpha(t) - \bar{\alpha})^2 \mathrm{d}t,
\end{aligned}$$

where $\bar{\alpha} = \frac{1}{t_f} \int_0^{t_f} \alpha(t)\mathrm{d}t$. The second line follows because $e^{2x(t)} \geq 1$ when $x(t) \geq 0$. The third line follows by completion of squares and $\int \dot{x}(t)\mathrm{d}t = x(t_f) - x(0) = 0$. This proves the upper-bound:

$$P_{\max} = \max_x \ \{J[x], \quad \text{s.t.} \quad x(0) = x(t_f), \quad x(t) \geq 0\} \leq \frac{k_B(T_h - T_c)}{8\lambda t_f^2} \int_0^{t_f} (\alpha(t) - \bar{\alpha})^2 \mathrm{d}t.$$

The proof for the lower-bound follows by using a particular choice for $x(t)$ and computing

the power for that particular choice. Let

$$x^*(t) = \frac{1}{2\lambda t_f}\int_{t_0}^{t}(\alpha(s)-\bar{\alpha})\,\mathrm{d}s,$$

this is a valid choice for $x(t)$: the periodic boundary condition is true because of definition of $\bar{\alpha}$; the positivity condition is true by suitable choice of t_0. Inserting x^* back to $J[x]$ gives

$$\begin{aligned}
J[x^*] &= \frac{k_B(T_h-T_c)}{2t_f}\int_0^{t_f} -\lambda t_f \dot{x}^*(t)^2 e^{2x^*(t)} + \alpha(t)\dot{x}^*(t)\mathrm{d}t \\
&\geq \frac{k_B(T_h-T_c)}{2t_f}\int_0^{t_f} -\lambda t_f \dot{x}^*(t)^2 e^{2x^*_{\max}} + \alpha(t)\dot{x}^*(t)\mathrm{d}t \\
&= (2-e^{2x^*_{\max}})\frac{k_B(T_h-T_c)}{8\lambda t_f^2}\int_0^{t_f}(\alpha(t)-\bar{\alpha})^2\mathrm{d}t \\
&= \frac{k_B(T_h-T_c)}{8\lambda t_f}\mathrm{Var}(\alpha) - (e^{2x^*_{\max}}-1)\frac{k_B(T_h-T_c)}{8\lambda t_f}\mathrm{Var}(\alpha),
\end{aligned}$$

where

$$x^*_{\max} := \max_t x^*_t \leq \frac{1}{2\lambda t_f}\int_{t_0}^{t}|\alpha(s)-\bar{\alpha}|\mathrm{d}s \leq \frac{1}{\lambda} = \frac{k_B(T_h-T_c)}{2\gamma\sigma^2_{\min}}t_f.$$

Therefore, in the limit as $t_f \to 0$, we have $e^{2x^*_{\max}} \to 1$, hence $J[x^*] \to \frac{k_B(T_h-T_c)}{8\lambda t_f^2}\mathrm{Var}(\alpha)$ concluding

$$P_{\max} \geq J[x^*] \geq \frac{k_B(T_h-T_c)}{8\lambda t_f}\mathrm{Var}(\alpha) - \epsilon,$$

where ϵ is positive, has the same order as $O(t_f)$. □

5.3.2 Analysis of efficiency

The efficiency is defined to be the ratio between the work output $W = t_f \mathcal{P}$ and heat input Q_h.

$$\eta = \frac{t_f \mathcal{P}}{Q_h}. \tag{5.15}$$

For the Carnot-like cycle, Q_h is defined as the heat transfer during the time interval that the system is in contact with hot thermal bath $T = T_h$. However, when the temperature profile is arbitrary, it is not clear to distinguish the heat transfer to the hot bath from the heat transfer to the cold bath. This issue is resolved in the linear response regime by defining Q_h according to

$$Q_h[q] = \int_0^{t_f} \frac{T(t) - T_c}{T_h - T_c} \dot{Q}(t)[q] \mathrm{d}t = \int_0^{t_f} \frac{\alpha(t) + 1}{2} \dot{Q}(t)[q] \mathrm{d}t, \tag{5.16}$$

where $\dot{Q}(t)[q]$ is the heat transfer rate given by the sum of the rate of change in energy and power output

$$\dot{Q}(t)[q] = \frac{\mathrm{d}}{\mathrm{d}t} \mathbb{E}\{U(t, X_t)\} + t_f \dot{\mathcal{P}}[q] = \frac{1}{2} q(t) \dot{\Sigma}(t).$$

In this paper, we use the definition (5.16) to define Q_h and efficiency, although the setting we consider is far from linear response regime. The next theorem characterizes the efficiency operating with the protocol (5.13) that achieves the upper-bound for maximum power as $t_f \to 0$:

$$\eta[q_0] := \frac{t_f \mathcal{P}[q_0]}{Q_h[q_0]}$$

under the following assumption on the temperature profile.

Assumption A1: There exists $\tau \in [0, t_f]$ such that $\alpha(t-\tau)$ is an odd function.

Theorem 5.3.2. *Consider the over-damped Langevin dynamics* (5.1) *under the protocol* (5.13) *for* $q(t)$ *and Assumption A1. Then, the efficiency satisfies*

$$\eta[q_0] = \frac{2\left(T_h - T_c\right)}{3T_h + T_c} - \frac{4(T_h + T_c)}{3T_h + T_c}\frac{1}{3T_h + T_c}t_f\epsilon - \frac{4(T_h + T_c)}{3T_h + T_c}\left(\frac{1}{(3T_h + T_c)^2}t_f^2\epsilon^2 + ...\right), \tag{5.17}$$

where ϵ *is positive, and have the same order as* $O(t_f)$*.*

Assumption A1: There exists $\tau \in [0, t_f]$ such that $\alpha(t-\tau)$ is an odd function.

Proof. Without loss generality, we can assume $\alpha(t)$ is odd, because time shifting doesn't affect the power and heat. For the heat absorbed from the hot reservoir,

$$\begin{aligned}
Q_h :&= \frac{1}{4}\int_0^{t_f} \left(\alpha(t)+1\right) q(t)x^2\dot{\rho}(x,t)dxdt \\
&= -\frac{1}{4}\int_0^{t_f} \dot{q}(t)x^2\rho(x,t)dxdt - \frac{1}{4}\int_0^{t_f} \dot{\alpha}(t)q(t)\Sigma(t)dt - \frac{1}{4}\int_0^{t_f} \alpha(t)\dot{q}(t)\Sigma(t)dt \\
&= -\frac{1}{4}\int_0^{t_f} \dot{q}(t)\Sigma(t)dt + \frac{1}{4}\int_0^{t_f} \alpha(t)q(t)\dot{\Sigma}(t)dt \\
&= \frac{1}{2}t_f\text{power} + \frac{1}{8}\int_0^{t_f} 2k_B\alpha(t)\left(\frac{T_h+T_c}{2} + \frac{T_h-T_c}{2}\alpha(t)\right)\frac{\dot{\Sigma}(t)}{\Sigma(t)}dt - \frac{1}{8}\int_0^{t_f}\frac{\gamma\alpha(t)\dot{\Sigma}(t)^2}{\Sigma(t)}dt \\
&= \frac{t_f}{2}\text{power} + \frac{k_B}{2}\int_0^{t_f} \alpha(t)\left(\frac{T_h+T_c}{2} + \frac{T_h-T_c}{2}\alpha(t)\right)\frac{\dot{\sigma}(t)}{\sigma(t)}dt - \frac{\gamma}{2}\int_0^{t_f}\alpha(t)\dot{\sigma}(t)^2dt \\
&= \frac{t_f}{2}\text{power} + \frac{k_B}{8\lambda t_f}\int_0^{t_f} \left(T_h+T_c\right)\alpha(t)^2 + (T_h-T_c)\alpha(t)^3dt \\
&\quad - \frac{\gamma\sigma_{\min}^2}{2}\int_0^{t_f}\frac{1}{4\lambda^2t_f^2}\alpha(t)^3e^{\frac{1}{\lambda t_f}\int_{t_0}^t\alpha(s)ds}dt \\
&= \frac{t_f}{2}\text{power} + \frac{k_B^2(T_h^2 - T_c^2)}{16\gamma\sigma_{\min}^2}\int_0^{t_f}\alpha(t)^2dt.
\end{aligned}$$

To show the last equality is true, we will show both $\int_0^{t_f}\alpha(t)^3$ and $\int_0^{t_f}\alpha(t)^3e^{\frac{1}{\lambda t_f}\int_{t_0}^t\alpha(s)ds}dt$ are

zero when $\alpha(t)$ is odd. For $\int_0^{t_f} \alpha(t)^3 dt$,

$$\begin{aligned}\int_0^{t_f} \alpha(t)^3 dt &= \int_0^{\frac{t_f}{2}} \alpha(t)^3 dt + \int_{\frac{t_f}{2}}^{t_f} -\alpha(-t)^3 dt = \int_0^{\frac{t_f}{2}} \alpha(t)^3 dt + \int_{-\frac{t_f}{2}}^{-t_f} \alpha(t)^3 dt \\ &= \int_0^{\frac{t_f}{2}} \alpha(t)^3 dt + \int_{\frac{t_f}{2}}^{0} \alpha(t)^3 dt = 0\end{aligned}$$

and for the third term $\int_0^{t_f} \alpha(t)^3 e^{\frac{1}{\lambda t_f}\int_{t_0}^t \alpha(s)ds} dt$, define

$$f(t) := \int_{t_0}^{t} \alpha(s) ds,$$

then $f(t)$ is an even function given $\alpha(t)$ is odd, which can be proved as the following: Define $g(t) := f(t) - f(-t)$, then $g'(t) = f'(t) + f'(-t) = \alpha(t) - \alpha(t) = 0$, then $g(t)$ is constant and $g(t) = g(0) = 0$, which gives $f(t) = f(-t) :$ $f(t)$ is even, from which $g_2(t) := e^{\frac{1}{\lambda t_f} f(t)}$ is even, therefore $\alpha(t)^3 g_2(t)$ is an odd function and

$$\int_0^{t_f} \alpha(t)^3 e^{\frac{1}{\lambda t_f}\int_0^t \alpha(s)ds} = 0,$$

which gives

$$Q_h = \frac{1}{2}\text{power} + \frac{k_B^2(T_h^2 - T_c^2)}{16\gamma\sigma_{\min}^2} \int_0^t \alpha(t)^2 dt,$$

from which the efficiency under the protocol (5.13) can be computed as

$$\begin{aligned}\eta[q_0] &= \frac{t_f \mathcal{P}[q_0]}{\frac{1}{2} t_f \mathcal{P}[q_0] + \frac{k_B^2(T_h^2 - T_c^2)}{16\gamma\sigma_{\min}^2} \int_0^t \alpha(t)^2 dt} = \frac{1}{\frac{1}{2} + \frac{k_B^2(T_h^2 - T_c^2)}{16\gamma\sigma_{\min}^2 t_f \mathcal{P}[q_0]} \int_0^t \alpha(t)^2 dt} \\ &= \frac{1}{\frac{1}{2} + \frac{k_B^2(T_h^2 - T_c^2)}{16\gamma\sigma_{\min}^2 t_f} \frac{\int_0^t \alpha(t)^2 dt}{\frac{(k_B(T_h - T_c))^2}{16\gamma\sigma_{\min}^2 t_f} \int_0^t \alpha(t)^2 dt \left(1 - \frac{t_f \epsilon}{T_h - T_c}\right)}} \\ &= \frac{1}{\frac{1}{2} + \frac{1}{\frac{T_h - T_c}{T_h + T_c} - \frac{t_f}{T_h + T_c}\epsilon}} = \frac{2\left(T_h - T_c - t_f \epsilon\right)}{3T_h + T_c - t_f \epsilon}\end{aligned}$$

$$= \frac{2\left(T_h - T_c\right)}{3T_h + T_c} - \frac{4(T_h + T_c)}{3T_h + T_c}\frac{1}{3T_h + T_c}t_f\epsilon - \frac{4(T_h + T_c)}{3T_h + T_c}\left(\frac{1}{(3T_h + T_c)^2}t_f^2\epsilon^2 + ...\right),$$

where we use $\mathcal{P}[q_0] = \frac{(k_B(T_h - T_c))^2}{16\gamma\sigma_{\min}^2 t_f}\mathrm{Var}(\alpha) - \frac{(T_h - T_c)k_B^2\int_0^t \alpha^2(t)dt}{16\gamma\sigma_{\min}^2}\epsilon$ for the computation, where ϵ is positive, and same order as $O(t_f)$. □

5.3.3 Consistent verification for optimal performance in the Carnot-limit

As we mentioned above, the upper bound of the power in (5.12) will be achieved when we apply the control protocol (5.13) as the period of the cycle tends to zero. In this part, we will compare our result in the Carnot limit where the temperature is piecewise constant with the known results for Carnot-like heat engine in [35] when the total period $t_f \to 0$.

$$T(t) = \begin{cases} T_h, & t \in (0, t_1) \\ T_c, & t \in (t_1, t_1 + t_2). \end{cases}$$

- The optimal protocol

  From Eq.(13) in [35], the optimal protocol for the variance defined as $\sigma_1(t)^2$ in the Carnot setting should satisfy

  $$\sigma_1(t) = \sigma_a + 2\frac{\sigma_b - \sigma_a}{t_f}t, \;\; 0 \le t < \frac{t_f}{2}, \tag{5.18}$$

  where σ_a^2 and σ_b^2 are the variance at initial point and half period point, respectively. In our current chapter, we know, from (5.14), the optimal protocol for the variance defined as $\sigma_2(t)^2$ during the interval $[0, \frac{t_f}{2})$ satisfy,

  $$\sigma_2(t) = \sigma_{\min}\exp\left(\frac{1}{2\lambda t_f}\int_{t_0}^{t}\left(\alpha(s) - \bar{\alpha}\right)\mathrm{d}s\right) = \sigma_{\min}\exp\left(\frac{1}{2\lambda t_f}\int_{t_0}^{t}\alpha(s)\mathrm{d}s\right), \tag{5.19}$$

where $\lambda := \frac{2\gamma\sigma_{\min}^2}{t_f k_B(T_h - T_c)}$ is a dimensionless constant, and t_0 is a time so that $\int_{t_0}^{t} (\alpha(s) - \bar{\alpha})\, \mathrm{d}s \geq 0$ for all t. In the Carnot limit, it can be written as

$$\begin{aligned}\sigma_2(t) =& \sigma_{\min} \exp\left(\frac{1}{2\lambda t_f}\int_{t_0}^{t} \alpha(s)\mathrm{d}s\right) = \sigma_{\min}\exp\left(\frac{1}{2\lambda t_f}\int_{t_0}^{t}\mathrm{d}s\right)\\ =& \sigma_{\min}\exp\left(\frac{t}{2\lambda t_f}\right) = \sigma_{\min}\exp\left(\frac{k_B(T_h - T_c)}{4\gamma\sigma_{\min}^2}t\right).\end{aligned}$$

When $t_f \to 0$, we can do the Taylor expansion

$$\sigma_2(t) = \sigma_{\min}\exp\left(\frac{k_B(T_h - T_c)}{4\gamma\sigma_{\min}^2}t\right) = \sigma_{\min}\left(1 + \frac{k_B(T_h - T_c)}{4\gamma\sigma_{\min}^2}t\right) + O(t^2).$$

From the Eq (5.14), in the limit case, we can write

$$\begin{aligned}&\sigma_{\min} = \sigma_a,\\ &\sigma_{\min}\left(1 + \frac{k_B(T_h - T_c)}{4\gamma\sigma_{\min}^2}\frac{t_f}{2}\right) = \sigma_a\left(1 + \frac{k_B(T_h - T_c)}{4\gamma\sigma_a^2}\frac{t_f}{2}\right) = \sigma_b\end{aligned} \tag{5.20}$$

and

$$\sigma_2(t) = \sigma_a\left(1 + \frac{k_B(T_h - T_c)}{4\gamma\sigma_a^2}t\right) = \sigma_a + 2\frac{\sigma_b - \sigma_a}{t_f}t,$$

which is consistent with Eq. (5.18). For the optimal variance in the other interval $[\frac{t_f}{2}, t_f)$, the same results hold based on the similar analysis.

For the optimal control in these two cases, because both satisfy the same Lyapunov equation, so they will coincide automatically.

- The optimal power extracted

  From Eq.(20) and Eq.(21) in [35], the maximal power extracted power_1 of the Carnot-

like heat engine should satisfy

$$\text{power}_1 = \frac{(T_h - T_c)^2 \ln \frac{\sigma_b}{\sigma_a}}{16\gamma(\sigma_b - \sigma_a)^2}.$$

For the result in the current chapter, the maximal power power_2 extracted in the Carnot limit and also $t_f \to 0$ can be written as

$$\text{power}_2 = \frac{k_B^2 T_c^2}{16\gamma\sigma_{\min}^2}(\frac{T_h}{T_c} - 1)^2 = \frac{k_B^2(T_h - T_c)^2}{16\gamma\sigma_{\min}^2}. \tag{5.21}$$

From Eq. (5.20), we can write power_1 as

$$\text{power}_1 = \frac{(T_h - T_c)^2 \ln \frac{\sigma_b}{\sigma_a}}{16\gamma(\sigma_b - \sigma_a)^2} = \frac{(T_h - T_c)^2}{16\gamma} \frac{\frac{k_B(T_h - T_c)}{8\gamma\sigma_a^2} t_f}{\frac{k_B(T_h - T_c)}{8\gamma\sigma_a} t_f} = \frac{k_B^2(T_h - T_c)^2}{16\gamma\sigma_a^2},$$

which coincides with Eq. (5.21).

- The efficiency at maximal power

  From Eq. (23) in [35], we know the efficiency at maximal power η_1 is given as

$$\eta_1 = \frac{2(T_h - T_c)}{3T_h + T_c}, \tag{5.22}$$

  which just depends on the temperatures.

  From the result in current chapter Eq. (5.17), the efficiency at maximal power η_2 in the Carnot limit and also $t_f \to 0$ can be written as

$$\eta_2 = \frac{2(T_h - T_c)}{3T_h + T_c},$$

  which coincides with Eq. (5.22).

5.4 Optimization under different constraints

In the previous section, we have investigated the optimization problem (5.11), however, the constraint $\Sigma(t) \geq \sigma^2_{\min}$ may not be easily implemented in some sense. In this section, we will focus on maximizing power under different constraints, which are technically easier to implement.

Problem 1: Bounded on the intensity of the potential $q(t)$

$$\mathcal{P}^*_{q_{\max}} = \max_q \ \{\mathcal{P}[q], \text{ s.t. } \Sigma(t) \text{ solves (5.8)}, \ q(0) = q(t_f), \Sigma(0) = \Sigma(t_f), \ q(t) \leq q_{\max}\}, \tag{5.23}$$

Corollary 5.4.1. *Consider the over-damped Langevin dynamics* (5.1) *under the temperature profile* (5.4)*. Then, the optimization problem* (5.23) *satisfies the upper-bound:*

$$\mathcal{P}^*_M \leq \frac{k_B T_c q_{max}}{16\gamma}(\frac{T_h}{T_c} - 1)^2 \, Var(\alpha). \tag{5.24}$$

Proof. There exists time t_0 such that $\Sigma(t_0) = \Sigma_{\min}$, where $\Sigma_{\min}$ is the smallest value of $\Sigma(t), \ t \in [0, t_f]$, then

$$\dot{\Sigma}(t_0) = 0,$$

which gives

$$q_{\max}\Sigma_{\min} \geq q(t_0)\Sigma_{\min} \geq k_B T_c,$$

then

$$\Sigma(t) \geq \Sigma_{\min} \geq \frac{k_B T_c}{q_{\max}},$$

substitute into (5.12) gives

$$\mathcal{P}_M^* \leq \frac{k_B T_c q_{\max}}{16\gamma}(\frac{T_h}{T_c} - 1)^2 \mathrm{Var}(\alpha).$$

□

Next, we will consider optimization under the constraint of control input as follows.

Problem 2: Bounded on the gradient of the potential

$$\mathcal{P}_M^* = \max_q \left\{ \mathcal{P}[q], \text{s.t. } \Sigma(t) \text{ solves (5.8)}, q(0) = q(t_f), \Sigma(0) = \Sigma(t_f), \frac{1}{\gamma}\mathbb{E}\{\|\nabla U\|^2\} \leq M \right\}. \tag{5.25}$$

Corollary 5.4.2. *Consider the over-damped Langevin dynamics* (5.1) *under the temperature profile* (5.4)*. Then, the optimization problem* (5.25) *satisfies the upper-bound:*

$$\mathcal{P}_M^* \leq \frac{M}{16}(\frac{T_h}{T_c} - 1)^2 \, Var(\alpha). \tag{5.26}$$

Proof. The constraint on the gradient potential given in Eq. (5.25) can be written as follows with quadratic potential $U(t,x) = \frac{1}{2}q(t)x^2$

$$\frac{1}{\gamma}q(t)^2\Sigma(t) \leq M. \tag{5.27}$$

There exists time t_0 such that $\Sigma(t_0) = \Sigma_{\min}$, where $\Sigma_{\min}$ is the smallest value for $\Sigma(t),\ t \in$

$[0, t_f]$, then

$$\dot{\Sigma}(t_0) = 0,$$

which gives $q(t_0)\Sigma_{\min} = k_B T(t_0) \geq k_B T_c$, then

$$q(t_0)^2 \Sigma_{\min}^2 \geq k_B^2 T_c^2,$$

together with the constraint $q(t_0)^2 \Sigma_{\min} < \gamma M$ giving

$$k_B^2 T_c^2 \leq q(t_0)^2 \Sigma_{\min}^2 \leq \gamma M \Sigma_{\min},$$

then

$$\Sigma(t) \geq \Sigma_{\min} \geq \frac{k_B^2 T_c^2}{\gamma M},$$

substitute into (5.12) gives

$$\mathcal{P}_M^* \leq \frac{M}{16}(\frac{T_h}{T_c} - 1)^2 \mathrm{Var}(\alpha).$$

□

5.5 Numerical verification

In this chapter, we will verify our results for:

$$\alpha(t) = \cos(\omega t).$$

Through the illustrating example of a sinusoidal temperature profile, we compare the results obtained in the linear response regime in Chapter 4 with the ones obtained in the present chapter. In particular, we consider a temperature profile of period $t_f = \frac{2\pi}{\omega}$ given by

$$T(t) = \frac{T_h + T_c}{2} + \frac{T_h - T_c}{2}\cos(\omega t).$$

First, we plot the optimal protocols obtained in each of the frameworks, along with the temperature profile in Figure 5.1, which shows that the optimal control law obtained in Eq. (5.13) coincides with the one obtained in the linear response regime in Chapter 4 when the difference between T_h and T_c is small. Next, we numerically solve the full Langevin equation with the optimal protocol in each of frameworks to get power output, which are compared with the upper bound Eq. (5.12) of the maximal power that can be drawn from continuous and periodic heat bath for different temperatures together. The results of this analysis are portrayed in Figure 5.2, and the model parameters used to obtain the numerical result are presented in Table 5.1.

Notation vs. value	notation	value
Boltzmann constant	k_B	1
viscosity coefficient	γ	1
period	t_f	0.6
minimal variance	σ^2_{min}	0.6

Table 5.1 – Parameters selected in the simulations.

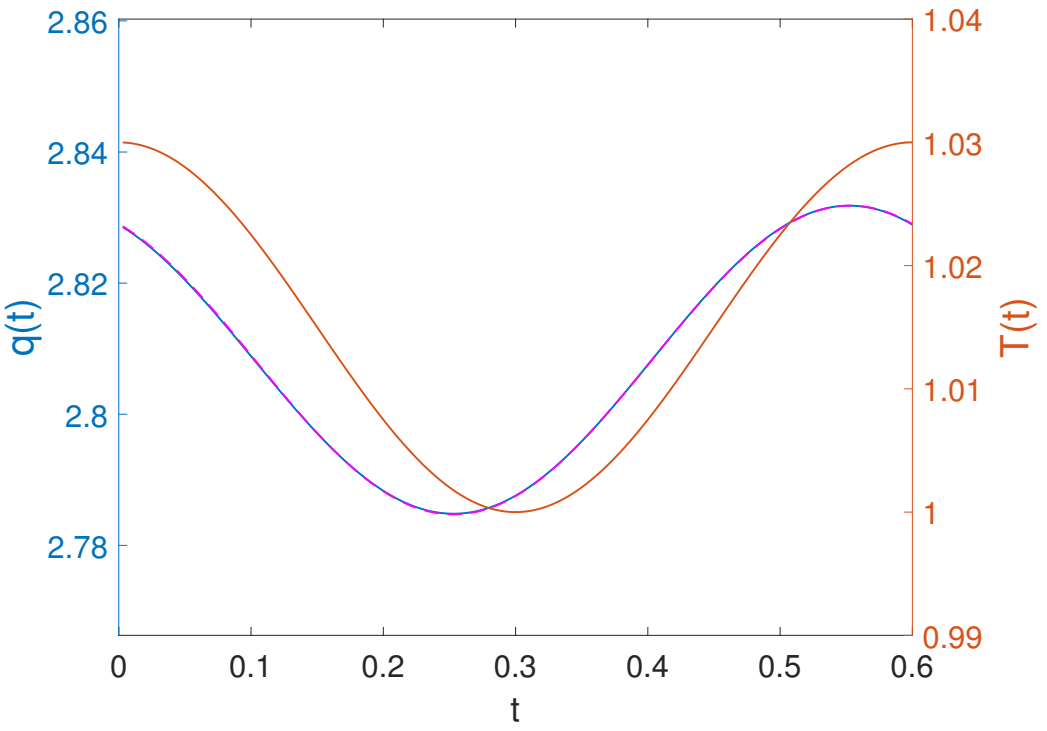

Figure 5.1 – Protocol for the sinusoidal temperature $T(t) = \frac{T_h+T_c}{2} + \frac{T_h-T_c}{2}\cos(\omega t)$: The solid line in blue is the protocol Eq. (5.13), while the dotted line in magenta is the protocol Eq. (4.21) in Chapter 4 when $T_h - T_c$ is small.

5.6 Conclusion and Remark

The present work focused on quantifying the maximal power that can be drawn from thermodynamic processes driven by the heat bath with periodic and continuous temperature profile. The motivation comes from the fact that biological processes in our nature rarely conforms to the setting of Carnot's cyclic contact with alternating heat baths, or the physics of the thermocouple with a stationary thermal gradient. Instead, it is the periodic fluctuations in chemical concentrations in conjunction with the variability of electrochemical potentials that provide the universal source of cellular energy. The framework is based on stochastic thermodynamics, which allows quantifying energy and heat exchanged by individual particles in a thermodynamic ensemble, to be subsequently averaged, so as to quantify performance of the thermodynamic as a whole. A universal and meaningful bound is derived, which is

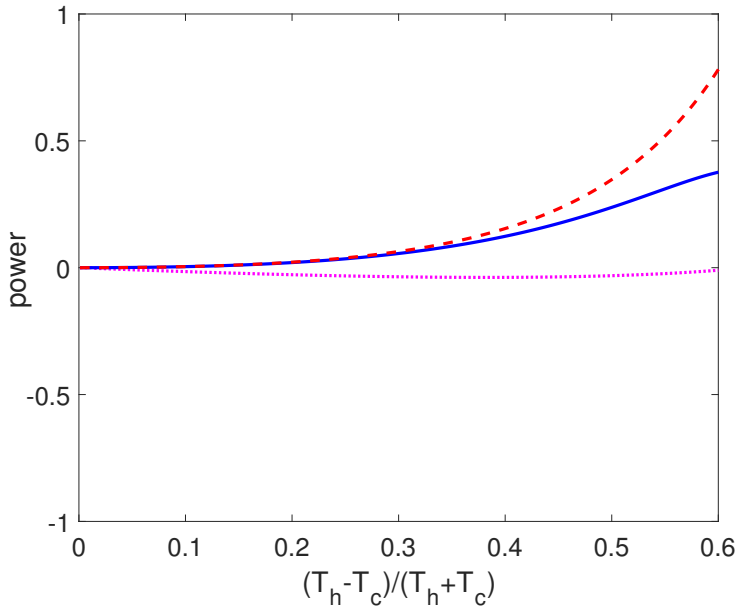

Figure 5.2 – The power output compared in different settings: The solid line in blue is the power obtained by using the protocol Eq. (5.13) with $T(t) = \frac{T_h+T_c}{2} + \frac{T_h-T_c}{2}\cos(\omega t)$, and the dotted line in magenta is the one obtained by using the protocol Eq. (4.21) in Chapter 4, while the dashed line in red is the upper bound in Eq. (5.12).

proportional to the average fluctuations in the temperature profile.

Chapter 6

Conclusion and Future work

6.1 Conclusion

In this dissertation, we focused on quantifying the maximal power that can be drawn from a thermodynamic heat engine that is periodically in contact with a heat reservoir and is modeled by stochastic overdamped Langevin dynamics driven by a time dependent potential that can be externally controlled. The framework that the work is based on is that of Stochastic Thermodynamics [9–13], from which fresh approaches have pointed at general laws applicable to non-equilibrium systems. The new advances appear to bridge the gap between theory and practice in that they lead to physically meaningful expressions for the dissipation cost in operating a thermodynamic engine over a finite time window. They allow quantifying energy and heat exchange by individual particles in a thermodynamic ensemble, to be subsequently averaged, so as to quantify performance of the thermodynamic process as a whole.

In Chapter 3, we considered the maximal power drawn from a Carnot-like heat engine operating by periodically alternating contact with two heat reservoirs. A physically reasonable

bound is derived, which is shown to be reached within a specified factor, both depending on the ratio T_h/T_c of the absolute temperatures of the two heat baths, hot and cold, respectively. The present work is quite distinct from earlier that are restricted to Gaussian states and quadratic potential. Conditions that suggest non-physical conclusions are highlighted, and a suitable constraint on the controlling potential is brought forth that underlies our analysis.

Motivated by the biological processes in nature that undergo a continuously changing temperature profile, instead of abrupt change in the Carnot cycle mode, Chapter 4 considers quantifying the maximal power that can be drawn from the thermodynamic heat engine periodically driven by a heat reservoir with sinusoidal temperature profile, and seek to determine the control law that maximizes the power. The analysis of the maximal power, and of efficiency at maximal power, is given in the linear response regime and shows that the optimal control law is sinusoidal function just as the temperature profile, but with different phase. Interestingly, the efficiency at maximal power just depends on the ratio of upper and lower bound of the sinusoidal temperature profile. The results are numerically verified via the Fourier analysis.

In Chapter 5, we studied the problem of quantifying the maximal power that can be drawn from a thermodynamic heat engine that is periodically driven by a heat reservoir with continuous temperature profile in general. We give an upper bound of the power that any heat engine can achieve under these conditions. The results are numerically verified by the special case of sinusoidal temperature profile.

In the past few decades, there have been several attempts to quantify efficiency mainly, but also power, of thermodynamic processes. It is fair to say that there has been neither a consensus on the type of assumptions that have been used by previous authors, and nor full consistency of the results. This is to be expected, since finite-period operation and finite-time thermodynamic transitions require substance/engine-dependent assumptions to

capture the complexity of heat transfer in non-equilibrium states. Thus, estimated bounds may never reach the "universality" of the celebrated Carnot efficiency. They are expected to provide physical insights and guidelines for engineering design. Thus, it is imperative that these estimates be subject to experimental testing. The notable feature of our conclusions as compared to earlier works is that: (1) For the Carnot heat engine, the expressions we derive are given in the form of the ratio of absolute temperatures–a physically suggestive feature. (2) For periodically and continuously driven thermodynamic engine, we give the universal upper bound of the maximal power that the heat engine can achieve.

6.2 Future work

The present work follows a long line of contributions within the control field to draw links between thermodynamics and control, see e.g. [70–75]. There is a wide range of possible applications as well as extensions of the theory that lay ahead.

- Underdamped description for the Carnot-like thermodynamic engine

  The Wasserstein distance of optimal mass transport, which itself is a solution to a stochastic control problem, has been linked to the dissipation mechanism in stochastic thermodynamics [11,51–54]. Indeed, the Wasserstein metric takes the form of an action integral and arises naturally in the energy balance of thermodynamic transitions. This fact has been explored and developed for the overdamped Langevin dynamics studied herein. Whether similar conclusions can be drawn for underdamped Langevin dynamics remains an open research direction at present. Furthermore, much work remains to reconcile and compare alternative viewpoints and models for thermodynamic processes including those based on the Boltzmann equation.

- Optimal performance in Stirling engines, thermocouples, and biological processes.

It is an open question whether methods similar to those in Chapter 3 can be used to study the optimal performance of Stirling engines, thermocouples, and biological processes.

- Optimal performance in the nonlinear regime

  In Chapter 4, we studied the optimal performance of the sinusoidally driven heat engine and determined the control law in the linear response regime. While in Chapter 5 we obtained a universal upper bound for the maximal power that can be extracted in general and a respective optimal control law for a special case, derivation of optimal control laws in the nonlinear response regime still remains open.

Bibliography

[1] Zahra Askarzadeh, Rui Fu, Abhishek Halder, Yongxin Chen, and Tryphon T Georgiou. Stability theory of stochastic models in opinion dynamics. *IEEE Transactions on Automatic Control*, 65(2):522–533, 2019.

[2] Zahra Askarzadeh, Rui Fu, Abhishek Halder, Yongxin Chen, and Tryphon T Georgiou. Opinion dynamics over influence networks. In *2019 American Control Conference (ACC)*, pages 1873–1878. IEEE, 2019.

[3] Rui Fu, Amirhossein Taghvaei, Yongxin Chen, and Tryphon T. Georgiou. Maximal power output of a stochastic thermodynamic engine. *Automatica*, 123:109366, 2021.

[4] Sadi Carnot. *Reflexions on the motive power of fire: a critical edition with the surviving scientific manuscripts*. Manchester University Press, 1986.

[5] Herbert B Callen. Thermodynamics and an introduction to thermostatistics, 1998.

[6] J Casas-Vázquez and D Jou. Temperature in non-equilibrium states: a review of open problems and current proposals. *Reports on Progress in Physics*, 66(11):1937, 2003.

[7] Georgy Lebon, David Jou, and José Casas-Vázquez. *Understanding non-equilibrium thermodynamics*, volume 295. Springer, 2008.

[8] Sybren R. De Groot and Peter Mazur. *Non-equilibrium thermodynamics*. Courier Corporation, 2013.

[9] Udo Seifert. Stochastic thermodynamics: principles and perspectives. *The European Physical Journal B*, 64(3-4):423–431, 2008.

[10] Ken Sekimoto. *Stochastic energetics*, volume 799. Springer, 2010.

[11] Udo Seifert. Stochastic thermodynamics, fluctuation theorems and molecular machines. *Reports on progress in physics*, 75(12):126001, 2012.

[12] Juan MR Parrondo, Jordan M Horowitz, and Takahiro Sagawa. Thermodynamics of information. *Nature physics*, 11(2):131–139, 2015.

[13] Andreas Dechant, Nikolai Kiesel, and Eric Lutz. Underdamped stochastic heat engine at maximum efficiency. *Europhysics Letters*, 119(5):50003, 2017.

[14] Roger W Brockett. Thermodynamics with time: Exergy and passivity. *Systems & Control Letters*, 101:44–49, 2017.

[15] Udo Seifert. Entropy production along a stochastic trajectory and an integral fluctuation theorem. *Physical review letters*, 95(4):040602, 2005.

[16] Christopher Jarzynski. Equalities and inequalities: Irreversibility and the second law of thermodynamics at the nanoscale. *Annu. Rev. Condens. Matter Phys.*, 2(1):329–351, 2011.

[17] Gavin E Crooks. Entropy production fluctuation theorem and the nonequilibrium work relation for free energy differences. *Physical Review E*, 60(3):2721, 1999.

[18] Ken Sekimoto. Langevin equation and thermodynamics. *Progress of Theoretical Physics Supplement*, 130:17–27, 1998.

[19] Christian Van den Broeck and Massimiliano Esposito. Ensemble and trajectory thermodynamics: A brief introduction. *Physica A: Statistical Mechanics and its Applications*, 418:6–16, 2015.

[20] Rosemary J Harris and Gunther M Schütz. Fluctuation theorems for stochastic dynamics. *Journal of Statistical Mechanics: Theory and Experiment*, 2007(07):P07020, 2007.

[21] B Gaveau, M Moreau, and LS Schulman. Stochastic thermodynamics and sustainable efficiency in work production. *Physical review letters*, 105(6):060601, 2010.

[22] Philipp Strasberg, Gernot Schaller, Neill Lambert, and Tobias Brandes. Nonequilibrium thermodynamics in the strong coupling and non-markovian regime based on a reaction coordinate mapping. *New Journal of Physics*, 18(7):073007, 2016.

[23] ML Rosinberg, T Munakata, and G Tarjus. Stochastic thermodynamics of langevin systems under time-delayed feedback control: Second-law-like inequalities. *Physical Review E*, 91(4):042114, 2015.

[24] Saurav Talukdar, Shreyas Bhaban, and Murti V Salapaka. Memory erasure using time-multiplexed potentials. *Physical Review E*, 95(6):062121, 2017.

[25] James Melbourne, Saurav Talukdar, and Murti V Salapaka. Realizing information erasure in finite time. In *2018 IEEE Conference on Decision and Control (CDC)*, pages 4135–4140. IEEE, 2018.

[26] Patrick R Zulkowski and Michael R DeWeese. Optimal finite-time erasure of a classical bit. *Physical Review E*, 89(5):052140, 2014.

[27] Paul Chambadal. Les centrales nucléaires. *Manuales Ingenieria*, 4:1–58, 1957.

[28] I. I. Novikov. The efficiency of atomic power stations (a review). *Journal of Nuclear Energy (1954)*, 7(1-2):125–128, 1958.

[29] F. L. Curzon and B. Ahlborn. Efficiency of a Carnot engine at maximum power output. *American Journal of Physics*, 43(1):22–24, 1975.

[30] Massimiliano Esposito, Katja Lindenberg, and Christian Van den Broeck. Universality of efficiency at maximum power. *Physical review letters*, 102(13):130602, 2009.

[31] Massimiliano Esposito, Ryoichi Kawai, Katja Lindenberg, and Christian Van den Broeck. Efficiency at maximum power of low-dissipation carnot engines. *Physical review letters*, 105(15):150603, 2010.

[32] Lixuan Chen and Zijun Yan. The effect of heat-transfer law on performance of a two-heat-source endoreversible cycle. *The Journal of Chemical Physics*, 90(7):3740–3743, 1989.

[33] Frank Jülicher, Armand Ajdari, and Jacques Prost. Modeling molecular motors. *Reviews of Modern Physics*, 69(4):1269, 1997.

[34] R Dean Astumian and Peter Hänggi. Brownian motors. *Physics today*, 55(11):33–39, 2002.

[35] Tim Schmiedl and Udo Seifert. Efficiency at maximum power: An analytically solvable model for stochastic heat engines. *EPL (Europhysics Letters)*, 81(2):20003, 2007.

[36] Ignacio A Martínez, Édgar Roldán, Luis Dinis, Dmitri Petrov, Juan MR Parrondo, and Raúl A Rica. Brownian carnot engine. *Nature physics*, 12(1):67–70, 2016.

[37] Andreas Dechant, Nikolai Kiesel, and Eric Lutz. Underdamped stochastic heat engine at maximum efficiency. *EPL (Europhysics Letters)*, 119(5):50003, 2017.

[38] Massimiliano Esposito, Ryoichi Kawai, Katja Lindenberg, and Christian Van den Broeck. Quantum-dot Carnot engine at maximum power. *Physical review E*, 81(4):041106, 2010.

[39] Carnot cycle. `https://en.wikipedia.org/wiki/Carnot_cycle`.

[40] Carnot cycle. `https://en.wikipedia.org/wiki/Fluctuation_theorem`.

[41] Christopher Jarzynski. Nonequilibrium equality for free energy differences. *Physical Review Letters*, 78(14):2690, 1997.

[42] Christopher Jarzynski. Equilibrium free-energy differences from nonequilibrium measurements: A master-equation approach. *Physical Review E*, 56(5):5018, 1997.

[43] Aykut Argun, Jalpa Soni, Lennart Dabelow, Stefano Bo, Giuseppe Pesce, Ralf Eichhorn, and Giovanni Volpe. Experimental realization of a minimal microscopic heat engine. *Physical Review E*, 96(5):052106, 2017.

[44] Jong-Min Park, Hyun-Myung Chun, and Jae Dong Noh. Efficiency at maximum power and efficiency fluctuations in a linear brownian heat-engine model. *Physical Review E*, 94(1):012127, 2016.

[45] Alex Gomez-Marin, Tim Schmiedl, and Udo Seifert. Optimal protocols for minimal work processes in underdamped stochastic thermodynamics. *The Journal of chemical physics*, 129(2):024114, 2008.

[46] J Horowitz and C Jarzynski. Comment on "failure of the work-hamiltonian connection for free-energy calculations". *Physical review letters*, 101(9):098901, 2008.

[47] Luca Peliti. Comment on "failure of the work-hamiltonian connection for free-energy calculations". *Physical review letters*, 101(9):098903, 2008.

[48] Luca Peliti. On the work–hamiltonian connection in manipulated systems. *Journal of Statistical Mechanics: Theory and Experiment*, 2008(05):P05002, 2008.

[49] Jose MG Vilar and J Miguel Rubi. Failure of the work-hamiltonian connection for free-energy calculations. *Physical review letters*, 100(2):020601, 2008.

[50] Richard Jordan, David Kinderlehrer, and Felix Otto. The variational formulation of the Fokker–Planck equation. *SIAM journal on mathematical analysis*, 29(1):1–17, 1998.

[51] Erik Aurell, Carlos Mejía-Monasterio, and Paolo Muratore-Ginanneschi. Optimal protocols and optimal transport in stochastic thermodynamics. *Physical review letters*, 106(25):250601, 2011.

[52] Erik Aurell, Krzysztof Gawędzki, Carlos Mejía-Monasterio, Roya Mohayaee, and Paolo Muratore-Ginanneschi. Refined second law of thermodynamics for fast random processes. *Journal of statistical physics*, 147(3):487–505, 2012.

[53] Yongxin Chen, Tryphon Georgiou, and Allen Tannenbaum. Stochastic control and nonequilibrium thermodynamics: fundamental limits. *IEEE Transactions on Automatic Control*, 65(1):252–262, 2020.

[54] Andreas Dechant and Yohei Sakurai. Thermodynamic interpretation of Wasserstein distance. *arXiv preprint arXiv:1912.08405*, 2019.

[55] Cédric Villani. *Topics in optimal transportation*. Number 58 in Graduate Studies in Mathematics. American Mathematical Soc., 2003.

[56] Luigi Ambrosio, Nicola Gigli, and Giuseppe Savaré. *Gradient flows: in metric spaces and in the space of probability measures*. Springer Science & Business Media, 2008.

[57] Felix Otto and Cédric Villani. Generalization of an inequality by talagrand and links with the logarithmic sobolev inequality. *Journal of Functional Analysis*, 173(2):361–400, 2000.

[58] Ivan Gentil, Christian Léonard, Luigia Ripani, and Luca Tamanini. An entropic interpolation proof of the HWI inequality. *Stochastic Processes and their Applications*, 2019.

[59] David R Owen. *A first course in the mathematical foundations of thermodynamics*. Springer Science & Business Media, 2012.

[60] Christopher Jarzynski. Equilibrium free-energy differences from nonequilibrium measurements: A master-equation approach. *Physical Review E*, 56(5):5018, 1997.

[61] Jorge Kurchan. Fluctuation theorem for stochastic dynamics. *Journal of Physics A: Mathematical and General*, 31(16):3719, 1998.

[62] R Graham. Path-integral methods in nonequilibrium thermodynamics and statistics. In *Stochastic processes in nonequilibrium systems*, pages 82–138. Springer, 1978.

[63] Michele Pavon and Francesco Ticozzi. On entropy production for controlled markovian evolution. *Journal of mathematical physics*, 47(6):063301, 2006.

[64] Yongxin Chen, Tryphon T Georgiou, and Michele Pavon. On the relation between optimal transport and schrödinger bridges: A stochastic control viewpoint. *Journal of Optimization Theory and Applications*, 169(2):671–691, 2016.

[65] Cédric Villani. *Optimal transport: old and new*, volume 338. Springer Science & Business Media, 2008.

[66] Jean-David Benamou and Yann Brenier. A computational fluid mechanics solution to the Monge-Kantorovich mass transfer problem. *Numerische Mathematik*, 84(3):375–393, 2000.

[67] Yury Polyanskiy and Yihong Wu. Wasserstein continuity of entropy and outer bounds for interference channels. *IEEE Transactions on Information Theory*, 62(7):3992–4002, 2016.

[68] Marco F Huber, Tim Bailey, Hugh Durrant-Whyte, and Uwe D Hanebeck. On entropy approximation for Gaussian mixture random vectors. In *2008 IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems*, pages 181–188. IEEE, 2008.

[69] Eric A. Carlen and Wilfrid Gangbo. Constrained deepest descent in the 2-Wasserstein metric. *Annals of Mathematics*, pages 807–846, 2003.

[70] Roger W. Brockett and Jan C. Willems. Stochastic control and the second law of thermodynamics. In *1978 IEEE Conference on Decision and Control including the 17th Symposium on Adaptive Processes*, pages 1007–1011. IEEE, 1979.

[71] Michele Pavon. Stochastic control and nonequilibrium thermodynamical systems. *Applied Mathematics and Optimization*, 19(1):187–202, 1989.

[72] Sanjoy K Mitter and Nigel J Newton. Information and entropy flow in the Kalman–Bucy filter. *Journal of Statistical Physics*, 118(1-2):145–176, 2005.

[73] Henrik Sandberg, Jean-Charles Delvenne, Nigel J Newton, and Sanjoy K Mitter. Maximum work extraction and implementation costs for nonequilibrium Maxwell's demons. *Physical Review E*, 90(4):042119, 2014.

[74] Tanmay Rajpurohit and Wassim Haddad. Stochastic thermodynamics: A dynamical systems approach. *Entropy*, 19(12):693, 2017.

[75] David Wallace. Thermodynamics as control theory. *Entropy*, 16(2):699–725, 2014.

[76] Michael Bauer, Kay Brandner, and Udo Seifert. Optimal performance of periodically driven, stochastic heat engines under limited control. *Physical Review E*, 93(4):042112, 2016.

[77] T. Schmiedl and U. Seifert. Efficiency at maximum power: An analytically solvable model for stochastic heat engines. *Europhysics Letters*, 81(2), 2007.

[78] Andreas Dechant, Nikolai Kiesel, and Eric Lutz. Underdamped stochastic heat engine at maximum efficiency. *Europhysics Letters*, 119(5), 2017.

[79] Nikolas Zöller. *Optimization of Stochastic Heat Engines in the Underdamped Limit*. Springer, 2017.