UNIVERSITY OF CALIFORNIA
SANTA CRUZ

MORI-ZWANZIG EQUATION: THEORY AND APPLICATION

A dissertation submitted in partial satisfaction of the
requirements for the degree of

DOCTOR OF PHILOSOPHY

in

APPLIED MATHEMATICS AND STATISTICS

by

Yuanran Zhu

December 2019

The Dissertation of Yuanran Zhu
is approved:

Daniele Venturi, Chair

Hongyun Wang

Qi Gong

Quentin Williams
Vice Provost and Dean of Graduate Studies

Contents

List of Figures

List of Tables

Abstract

Mori-Zwanzig equation: Theory and Application

by

Yuanran Zhu

The Mori-Zwanzig (MZ) formulation is a technique from irreversible statistical mechanics that allows the development of formally exact evolution equation for the quantities of interest such as macroscopic observables in high-dimensional dynamical systems. Although being widely used in physics and applied mathematics as a common tool of dimension reduction, the analytical properties of the equation are still unknown, which makes the quantification and approximation of the MZ equation arduous tasks for the whole community. In this dissertation, we address this problem from both theoretically and computational points of view. For the first time, we study the MZ equation, especially the memory integral term, using the theory of strongly continuous semigroups, thereby establish an estimation theory for the memory effects which is based on solid mathematical foundations. In particular, some recent results from the Hörmader analysis of hypoelliptic equations are applied to get exponentially decaying estimates of the MZ memory kernel. We also develop a series expansion technique to approximate the MZ equation, and provide associated combinatorial algorithms to calculate the expansion coefficients from the first principle. The new approximation methods are tested on various linear and nonlinear dynamical systems, with convergence results obtained both theoretically and numerically. Further developments of the Mori-Zwanzig formulation based on the mathematical framework provided in this work can be expected, which can be used in more general dimension reduction problems from physics and mathematics.

To my parents: Shenglong Zhu and Yuefen Wang.

Acknowledgments

First of all, I would like to express my sincere gratitude to my advisor Professor Daniele Venturi for guiding me throughout my whole graduate study. It is you who brought me into this fascinating research direction and inspired me with the spirit of excellence. It is your endless encouragement that made this dissertation possible. I would also like to thank Professor Qi Gong, Hongyun Wang, Dongwook Lee, Pascale Garaud and Nic Brummell for teaching me and providing me academic and emotional support duing my five-year study in UC Santa Cruz.

I want thank my fellow graduate students and friends in the department of applied mathematics and statistics. In particular, I want thank Sky, George, Sara, Sisi, Bram, Catherine, Tenavi, Rene, Alex, Jason and many Daniels. You guys made my life in this little town so colorful. I also want to thank my roommates and friends, Wendy, Gao Lao, Tianrui, Xiao Mai, Jude and Zoe for their company in this foreign country so I would not be homesick during holidays. I will remember the wonderful time we spent together in the valley, in Napa, in Santa Barbara, in Yosemite and Yellowstone, forever.

Last but certainly not least, I would like the thank my parents for their unconditional support. I am so lucky to be born in this family full of love and caring. I am sorry I cannot spend more time with you during my undergraduate and graduate study. But I am and will love you in my whole life, just like the way you love me.

Chapter 1

Introduction

Modern applied mathmatics often needs to address high dimensional (stochastic) systems with complex dynamics. Despite the fact that the current computational power enables us to perform large-scale simulation of complex systems such as turbulence, running simulations that fully account for the variability of these physical models can be infeasible due to the curse of dimensionality. In practice, however, it is often the case that only some local observables, such as the low dimensional phase functions of a high-dimensional system are of interest. This necessitates the development of theoretical and computational strategies to construct effective models that describe the dynamics of these reduced-order quantities. To this end, the Mori-Zwanzig (MZ) formuation, which was first established in irreversible statistical mechanics [67, 108, 107], provides us a general framework to derive the exact evolution equation of such quantities from the underlying high dimensional systems. Once the exact evolution equation is built, then one can avoid integrating the full (possibly high-dimensional) dynamical system and instead solve directly for the quantities of interest, the whole computational cost would be reduced significantly.

Technically speaking, the Mori-Zwanzig equation is a general framework which is applicable to dimenison reduction problems in any spatial scales. In practice, the dynamics of a general physical model can be written as the evolution equation of the form

$$\frac{\partial}{\partial t}e^{t\mathcal{G}}u(0) = \mathcal{G}u(0), \tag{1.1}$$

where $e^{t\mathcal{G}}$ is the evolution operator and $\mathcal{G}$ is the corresponding infinitesimal generator. In the microscopic scale, the generator $\mathcal{G}$ can be the Liouville operator $\mathcal{L}$ and the resulting Liouville equation describes the dynamics of a system of a large number of interacting particles. Through some systematic coarse-graining procesure such as the BBGKY hierarchy, we may get effective generators $\mathcal{G}$ which is the Fokker-Planck operator corresponding to the dynamics of Brownian particles in mesoscopic scales or the Boltzmann operator for the phase space distribution function in macroscopic scales. Now matter which scale we are interested in, the evolution equation of low-dimensional observables can be formally derived as an operator equation, which is now known as the Mori-Zwanzig equation. The basic idea is to use suitable projection operators $\mathcal{P}$ and $\mathcal{Q} = \mathcal{I} - \mathcal{P}$ to split the dynamics of the original, high-dimensional system into the part for the relevant variables (streaming term), irrelevant rest degree of freedom (fluctuation term) and the interaction between these two (memory term). For kinetic equation (1.1) (Here we assumed that the generator $\mathcal{G}$ is time independent), the MZ equation would be like

$$\frac{\partial}{\partial t}e^{t\mathcal{G}}u(0) = \underbrace{e^{t\mathcal{G}}\mathcal{P}\mathcal{G}u(0)}_{\text{Streaming}} + \underbrace{e^{t\mathcal{Q}\mathcal{G}\mathcal{Q}}\mathcal{Q}\mathcal{G}u(0)}_{\text{Fluctuation}} + \underbrace{\int_0^t e^{s\mathcal{G}}\mathcal{P}\mathcal{G}e^{(t-s)\mathcal{Q}\mathcal{G}\mathcal{Q}}\mathcal{Q}\mathcal{G}u(0)ds}_{\text{Memory}} \tag{1.2}$$

where $\mathcal{Q}\mathcal{G}\mathcal{Q}$ is the infinitesimal generator of the orthogonal propagator $e^{t\mathcal{Q}\mathcal{G}\mathcal{Q}}$. As an example, consider a large system of interacting particles, and suppose we are interested in studying the motion of one specific particle. By applying the MZ formulation to the Liouville equation of the full particle system, it is possible extract a formally exact MZ equation governing the position and the momentum of the particle of interest. This is at the basis of microscopic physical theories of Brownian motion [46].

Although exact, the Mori-Zwanzig equation (1.2) is a complicated operator equation which involves the orthogonal dynamics $e^{t\mathcal{Q}\mathcal{G}\mathcal{Q}}$ and a convoluted memory integral which encoded the interaction between the orthogonal dynamics and the streaming term. These two terms, especially the memory integral, are extremely complicated. The approximation and quantification of them constitutes the main theme of most research on the MZ equation. Over the years, many techniques have been proposed to address this problem numerically. These techniques can be grouped in two categories: i) data-driven methods; ii) methods based on first-principles. Data-driven methods aim at recovering the MZ memory integral/fluctuation term based on data, usually in the form of sample trajectories of the full system. Typical examples are the NARMAX technique developed by Lu *et al.* [61], the rational function approximation recently proposed by Lei *et al.* [57] (see also [19]), and the conditional expectation technique developed by Brennan and Venturi [11]. On the other hand, methods based on first principles aim at approximating the MZ memory integral and fluctuation term based on the structure of the nonlinear system (microscopic equations of motion), without using any simulation data. The first effective method developed within this class is the continued fraction expansion of Mori [67], which can be conveniently formulated in terms of recurrence relations [56, 33]. Other methods based on first-principles include perturbation methods [100, 96], mode coupling techniques, [81, 36], optimal prediction methods [15, 18, 85], and various series expansion [87, 74, 73, 103].

First-principle calculation methods can effectively capture non-Markovian memory effects, e.g., in coarse-grained particle simulations [102, 41]. However, they are often quite involved and they do not generalize well to systems with no scale separation [35]. At the same time, data-driven methods can yield accurate results, but they often require a large number of sample trajectories to faithfully capture memory effects [11, 19, 57, 58, 59]. On the other hand, the theoretical study of the MZ theory developed slowly since the establishment of the framework in 60s. Although the MZ equation has been used in physics and applied math as practical computational tools for many years, rigorous theoretical studies of the MZ equation are still rather lacking and only some prelimilary results were obtained in this direction. An representative example is the well-posedness study done by Kupferman, Givon and Hald. In [34], they extended the Friedrichs' existence proof for symmetric hyperbolic systems and got the existence and uniqueness of the orthogonal dynamics for classical dynamical system under Mori's projection $\mathcal{P}$. The lack of theoretical studies partially attributes to the complexity of the orthogonal dynamics. Since the analytical properties of the propagator $e^{t\mathcal{Q}\mathcal{G}\mathcal{Q}}$ are barely known, any quantifications of the memory integral and fluctuation term would be arduous tasks.

In this dissertation, we address this problem from both theoretical and numerical points of view. For the theoretic part, we will use the analytical properties of the kinetic equation (1.1) to derive estimates of strongly continuous semigroup $e^{t\mathcal{L}}$ and $e^{t\mathcal{Q}\mathcal{G}\mathcal{Q}}$. The obtained semigroup estimations yield immediately the *prior* estimate of the MZ memory integral, and the convergence analysis of frequently used approximation schemes. The results will be presented for both classical dynamical systems and stochastic systems driven by white noise. For the numerical parts, we introduce a series expansion method to approximate the MZ equation. Its convergence for linear system can be obtained analytically with an accurate estimate of the convergence rate. For general nonlinear systems, we propose a combinatorial algorithm to calculate the expansion coefficients from the first principle.

The whole dissertation is organized as follows. In Chapter 1, we formally derive the MZ equation for classical and stochastic dynamical systems. The derivation of the MZ equation for the first case is explained rather clearly in many previous studies such as [68, 15, 18, 85], Hence we will focus more on the analysis of the projection operators that are frequently used in the literature. For stochastic dynamical system, we provide a self-consistent way to derive the MZ equation for stochastic differential equations driven by white noise. In Section 2.3, an new MZ equation, which is called as the effective MZ equation (EMZE) is proposed based on the ensemble average of the original MZ equation for SDEs. Chapter 3 and 4 constitute the main theoretical studies we have done for the MZ equation. In Chapter 3, we will focus on the analysis of deterministic systems. In particular, the properties of Liouville operator $\mathcal{L}$ are used to derive estimates for MZ memory integral and its various approximation schemes. In Chapter 4, we will show that the generator $\mathcal{G}$ that appears in the MZ equation for SDEs is Kolmogorov operator $\mathcal{K}$. In addition, if $\mathcal{K}$ is proven to be hypoelliptic (see the definitions in Chapter 4). Then the exponentially decaying estimate can be obtained for semigroup $e^{-t\mathcal{K}}$, $e^{-t\mathcal{QKQ}}$ and the MZ memory kernel. Chapter 5 and 6 are devoted to the approximation of the MZ equation which constitutes the main body of numerical studies of this dissertation. In Chapter 5, we develop new series expansions of the MZ memory integral based on operator series of the orthogonal dynamics propagator. We also develop exact MZ equations for the mean and the auto-correlation function of an observable of interest, and determine their analytical solution through Laplace transforms. A thorough convergence analysis of the memory approximation methods is also given for linear systems. The new approximation scheme is tested in various random wave propagation problems such as the one in an annulus, Bethe lattice and random graphs with arbitrary topology. In Chapter 6, we present a new method to compute the MZ memory integral and the flucuation term from first principles for nonlinear systems with local polynomial interactions. The method essentially is combinatorial which is tested in nonlinear wave propagation problems. We also propose a stochastic modelling diagram which is based on the MZ framework and bi-orthogonal series expansion theory. The main results of this dissertation are summarized in Chapter 7.

Chapter 2

Mori-Zwanzig formulation

In this section, we represent systematic ways to derive the Mori-Zwanzig equation of the reduced-order quantities for both classical and stochastic dynamical systems. For the deterministic, classical system, the derivation of the MZ equation is rather standard hence we will focus more on the properties of the projection operator $\mathcal{P}$ and its relationship with the MZ equation. For the stochastic system, we will concentrate on the analysis of system driven by white noise. In this case, stochastic analysis has to be introduced in order to derive the MZ equation and the effective MZ equation.

2.1 The Mori-Zwanzig equation for classical system

Consider the nonlinear dynamical system

$$\frac{dx}{dt} = F(x), \qquad x(0) = x_0 \tag{2.1}$$

evolving on a smooth manifold $\mathcal{S}$. For simplicity, let us assume that $\mathcal{S} = \mathbb{R}^n$. We will consider the dynamics of scalar-valued observables $g : \mathcal{S} \to \mathbb{C}$, and for concreteness, it will be desirable to identify structured spaces of such observable functions. In [22], it was argued that C^*-algebras of observables such as $L^\infty(\mathcal{S}, \mathbb{C})$ (the space of all measureable, essentially bounded functions on $\mathcal{S}$) and $C_0(\mathcal{S}, \mathbb{C})$ (the space of all continuous functions on $\mathcal{S}$, vanishing at infinity) make natural choices. In what follows, we do not require the observables to comprise a C^*-algebra, but we will want them to comprise a Banach space as the estimation theorems of section 3.1 make extensive use of the norm of this space. Having the structure of a Banach space of observables also gives greater context to the meaning of the linear operators $\mathcal{L}$, $\mathcal{K}$, $\mathcal{P}$, and $\mathcal{Q}$ to be defined hereafter.

The dynamics of any scalar-valued observable $g(x)$ (quantity of interest) can be expressed in terms of a semi-group $\mathcal{K}(t, s)$ of operators acting on the Banach space of observables. This is the Koopman operator [53] which acts on the function g as

$$g(x(t)) = [\mathcal{K}(t, s)g]\,(x(s)), \tag{2.2}$$

where

$$\mathcal{K}(t, s) = e^{(t-s)\mathcal{L}}, \qquad \mathcal{L}g(x) = F(x) \cdot \nabla g(x). \tag{2.3}$$

Rather than compute the Koopman operator applicable to all observables, it is often more tractable to compute the evolution only of a (closed) subspace of quantities of interest. This subspace can be described conveniently by means of a projection operator $\mathcal{P}$ with the subspace as its image. Both $\mathcal{P}$ and the complementary projection $\mathcal{Q} = \mathcal{I} - \mathcal{P}$ act on the space of observables. The nature, mathematical properties and connections between $\mathcal{P}$ and the observable g are discussed in detail in [22], and summarized in section 2.1.1. For now it suffices to assume that $\mathcal{P}$ is a bounded linear operator, and that $\mathcal{P}^2 = \mathcal{P}$. The

MZ formalism describes the evolution of observables initially in the image of $\mathcal{P}$. Because the evolution of observables is governed by the semi-group $\mathcal{K}(t,s)$, we seek an evolution equation for $\mathcal{K}(t,s)\mathcal{P}$. By using the definition of the Koopman operator (2.3), and the well-known Dyson identity

$$e^{t\mathcal{L}} = e^{t\mathcal{Q}\mathcal{L}} + \int_0^t e^{s\mathcal{L}}\mathcal{P}\mathcal{L}e^{(t-s)\mathcal{Q}\mathcal{L}}ds$$

we obtain the operator equation

$$\frac{d}{dt}e^{t\mathcal{L}} = e^{t\mathcal{L}}\mathcal{P}\mathcal{L} + e^{t\mathcal{Q}\mathcal{L}}\mathcal{Q}\mathcal{L} + \int_0^t e^{s\mathcal{L}}\mathcal{P}\mathcal{L}e^{(t-s)\mathcal{Q}\mathcal{L}}\mathcal{Q}\mathcal{L}ds. \tag{2.4}$$

By applying this equation to an observable function u_0, we obtain the well-known MZ equation in phase space

$$\frac{\partial}{\partial t}e^{t\mathcal{L}}u_0 = e^{t\mathcal{L}}\mathcal{P}\mathcal{L}u_0 + e^{t\mathcal{Q}\mathcal{L}}\mathcal{Q}\mathcal{L}u_0 + \int_0^t e^{s\mathcal{L}}\mathcal{P}\mathcal{L}e^{(t-s)\mathcal{Q}\mathcal{L}}\mathcal{Q}\mathcal{L}u_0ds. \tag{2.5}$$

The three terms at the right hand side are called, respectively, streaming term, fluctuation (or noise) term, and memory term. It is often more convenient (and tractable) to compute the evolution of the observable $\boldsymbol{u}(t)$ within a closed linear space, e.g., the image of the projection operator $\mathcal{P}$. To this end, we apply such projection to both sides of equation (2.5). This yields the evolution equation for projected dynamics[1]

$$\frac{\partial}{\partial t}\mathcal{P}e^{t\mathcal{L}}u_0 = \mathcal{P}e^{t\mathcal{L}}\mathcal{P}\mathcal{L}u_0 + \int_0^t \mathcal{P}e^{s\mathcal{L}}\mathcal{P}\mathcal{L}e^{(t-s)\mathcal{Q}\mathcal{L}}\mathcal{Q}\mathcal{L}u_0ds. \tag{2.6}$$

Depending on the choice of the projection operator, the MZ equation (2.6) can yield evolution equations for different quantities. For example, if we use Chorin's

[1]Note that the second term in (2.5), i.e., $\mathcal{P}e^{t\mathcal{Q}\mathcal{L}}\mathcal{Q}\mathcal{L}x_0 = 0$ vanishes since $\mathcal{P}\mathcal{Q} = 0$.

projection [15, 16, 103, 95], then (2.6) is an evolution equation for the conditional mean of $\boldsymbol{u}(t)$. Similarly, if we use Mori's projection [104, 83], then (2.6) is an evolution equation for the temporal auto-correlation function of $\boldsymbol{u}(t)$.

2.1.1 Projection Operators

In this section, we make a summary on the commonly used projection operators $\mathcal{P}$ in the Mori-Zwanzig framework. To make our definition mathematically sound, we begin by assuming that the Liouville operator (2.3) acts on observable functions in a C^*-algebra $\mathfrak{A}$, for instance $L^\infty(\mathcal{M}, \Sigma, \mu)$, where $\mathcal{M}$ is a space such as $\mathbb{R}^N$, Σ is a σ-algebra on $\mathcal{M}$, and μ is a measure on Σ. Let $\sigma \in \mathfrak{A}_*$ be a positive linear functional on $\mathfrak{A}$. We define the weighted pre-inner product

$$\langle f, g \rangle_\sigma = \sigma(f^* g).$$

This can be used to define a Hilbert space $\mathcal{H} = L^2(\mathcal{M}, \sigma)$, which is the completion of the quotient space

$$\mathcal{H}' = \{f \in \mathfrak{A} : \sigma(f^* f) < \infty\} / \{f \in \mathfrak{A} : \sigma(f^* f) = 0\}$$

endowed with the inner product $\langle \cdot, \cdot \rangle_\sigma$. The L^2 norm induced by the inner product is denoted as $\| \cdot \|_\sigma$. The positive linear functional σ is set to be induced by a probability distribution $\tilde{\sigma}$ through

$$\sigma(u) = \int_{\mathcal{M}} \tilde{\sigma}(\omega) u(\omega) d\omega,$$

where $\tilde{\sigma}$ is typically chosen to be the probability density of the initial condition ρ_0, or the equilibrium distribution ρ_{eq} in statistical physics. To conform to the literature, we also use notation $\langle\cdot,\cdot\rangle_{\rho_0}$, $\langle\cdot,\cdot\rangle_{\rho_{eq}}$, $\langle\cdot,\cdot\rangle_{eq}$ to represent the weighted inner product corresponding to different probability measures $\tilde{\sigma}(\omega)d\omega$. With the Hilbert space determined, we now focus on the following two broad class of orthogonal projections on $\mathcal{H}$.

Infinite-Rank Projections

The first class of projection operators to consider in this setting are the conditional expectations $\mathcal{P}$ such that $\mathcal{P}_*\sigma = \sigma$. In this case, the properties of conditional expectations (in particular that $\mathcal{P}[\mathcal{P}(f)g\mathcal{P}(h)] = \mathcal{P}(f)\mathcal{P}(g)\mathcal{P}(h)$ [91]) and the fact that $\mathcal{P}_*\sigma = \sigma$ imply that

$$\langle \mathcal{P}f, g\rangle_\sigma = \sigma[(\mathcal{P}f)^*g] = \mathcal{P}_*(\sigma)[(\mathcal{P}f)^*g] = \sigma[\mathcal{P}((\mathcal{P}f)^*g)] = \sigma[(\mathcal{P}f)^*(\mathcal{P}g)]$$
$$\langle f, \mathcal{P}g\rangle_\sigma = \sigma[f^*\mathcal{P}g] = \mathcal{P}_*(\sigma)[f^*\mathcal{P}g] = \sigma[\mathcal{P}(f^*\mathcal{P}g)] = \sigma[(\mathcal{P}f^*)(\mathcal{P}g)] = \sigma[(\mathcal{P}f)^*(\mathcal{P}g)]$$

so that

$$\langle \mathcal{P}f, g\rangle_\sigma = \langle f, \mathcal{P}g\rangle_\sigma$$

for all $f, g \in \mathcal{H}$. It follows that

$$\langle \mathcal{Q}f, g\rangle_\sigma = \langle f, g\rangle_\sigma - \langle \mathcal{P}f, g\rangle_\sigma = \langle f, g\rangle_\sigma - \langle f, \mathcal{P}g\rangle_\sigma = \langle f, \mathcal{Q}g\rangle_\sigma.$$

Therefore both $\mathcal{P}$ and $\mathcal{Q}$ are self-adjoint (i.e. orthogonal) projections onto closed subspaces of $\mathcal{H}$, hence contractions $\|\mathcal{P}\|_\sigma \leq 1$, $\|\mathcal{Q}\|_\sigma \leq 1$. Chorin's projection

[18, 15] is one of this class, and is defined as

$$\Big(\mathcal{P}g\Big)(\hat{x}_0) = \frac{\displaystyle\int_{-\infty}^{+\infty} g(\hat{x}(t;\hat{x}_0,\tilde{x}_0),\tilde{x}(t;\hat{x}_0,\tilde{x}_0))\rho_0(\hat{x}_0,\tilde{x}_0)d\tilde{x}_0}{\displaystyle\int_{-\infty}^{+\infty} \rho_0(\hat{x}_0,\tilde{x}_0)d\tilde{x}_0} = \mathbb{E}_{\rho_0}[g|\hat{x}_0]. \tag{2.7}$$

Here $x(t;x_0) = (\hat{x}(t;\hat{x}_0,\tilde{x}_0),\tilde{x}(t;\hat{x}_0,\tilde{x}_0))$ is the flow map generated by (2.1) split into resolved ($\hat{x}$) and unresoved ($\tilde{x}$) variables, and $g(x) = g(\hat{x},\tilde{x})$ is the quantity of interest. For Chorin's projection, the positive functional σ defining the Hilbert space $\mathcal{H}$ may be taken to be integration with respect to the probability measure $\rho_0(\hat{x}_0,\tilde{x}_0)$.

Clearly, if x_0 is deterministic then $\rho_0(\hat{x}_0,\tilde{x}_0)$ is a product of Dirac delta functions. On the other hand, if $\hat{x}(0)$ and $\tilde{x}(0)$ are statistically independent, i.e. $\rho_0(\hat{x}_0,\tilde{x}_0) = \hat{\rho}_0(\hat{x}_0)\tilde{\rho}_0(\tilde{x}_0)$, then the conditional expectation $\mathcal{P}$ simplifies to

$$\Big(\mathcal{P}u\Big)(\hat{x}_0) = \int_{-\infty}^{+\infty} u(\hat{x}(t;\hat{x}_0,\tilde{x}_0),\tilde{x}(t;\hat{x}_0,\tilde{x}_0))\tilde{\rho}_0(\tilde{x}_0)d\tilde{x}_0. \tag{2.8}$$

In the special case where $u(\hat{x},\tilde{x}) = \hat{x}(t;\hat{x}_0,\tilde{x}_0)$ we have

$$\Big(\mathcal{P}\hat{x}\Big)(\hat{x}_0) = \int_{-\infty}^{+\infty} \hat{x}(t;\hat{x}_0,\tilde{x}_0)\tilde{\rho}_0(\tilde{x}_0)d\tilde{x}_0, \tag{2.9}$$

i.e., the conditional expectation of the resolved variables $\hat{x}(t)$ given the initial condition $\hat{x}_0$. This means that an integration of (2.9) with respect to $\hat{\rho}_0(\hat{x}_0)$ yields the mean of the resolved variables, i.e.,

$$\mathbb{E}_{\rho_0}[\hat{x}(t)] = \int_{-\infty}^{\infty} \Big(\mathcal{P}\hat{x}\Big)(\hat{x}_0)\hat{\rho}_0(\hat{x}_0)d\hat{x}_0 = \int_{-\infty}^{\infty} \hat{x}(t,x_0)\rho_0(x_0)dx_0. \tag{2.10}$$

Obviously, if the resolved variables $\hat{x}(t)$ evolve from a deterministic initial state $\hat{x}_0$ then the conditional expectation (2.9) represents the the average of the reduced-order flow map $\hat{x}(t;\hat{x}_0,\tilde{x}_0)$ with respect to the PDF of $\tilde{x}_0$, i.e., the flow map

$$\mathcal{P}e^{t\mathcal{L}}\hat{x}(0) = X_0(t;\hat{x}_0) = \int_{-\infty}^{+\infty} \hat{x}(t;\hat{x}_0,\tilde{x}_0)\tilde{\rho}_0(\tilde{x}_0)d\tilde{x}_0. \tag{2.11}$$

In this case, the MZ equation (2.6) is an exact (unclosed) evolution equation (PDE) for the multivariate field $X_0(t,\hat{x}_0)$. In order to close such an equation, a mean field approximation of the type $\mathcal{P}f(\hat{x}) = f(\mathcal{P}\hat{x})$ was introduced by Chorin *et al.* in [15, 18, 17], together with the assumption that the probability distribution of x_0 is invariant under the flow generated by (2.1).

Finite-Rank Projections

Another class of projections is defined by choosing a closed (typically finite-dimensional) linear subspace $V \subset \mathcal{H} = L^2(\mathcal{M},\sigma)$ and letting $\mathcal{P}$ be the orthogonal projection onto V in the σ inner product. An example of such projection is Mori's projection [68], widely used in statistical physics. For finite-dimensional V, given a linearly independent set $\{u_1,...,u_M\} \subset V$ that spans V, $\mathcal{P}$ can be defined by first constructing the positive definite Gram matrix

$$G_{ij} = \int_{\mathcal{M}} u_i(\boldsymbol{x})u_j(\boldsymbol{x})\rho(\boldsymbol{x})d\boldsymbol{x}. \tag{2.12}$$

With G_{ij} available, we define

$$\mathcal{P}f = \sum_{i,j=1}^{M} G_{ij}^{-1}\langle u_i(0), f\rangle_\rho u_j(0), \qquad f \in H. \tag{2.13}$$

In classical statistical dynamics of Hamiltonian systems, a common choice for the density ρ is the Boltzmann-Gibbs distribution

$$\rho_{eq}(\boldsymbol{x}) = \frac{1}{Z} e^{-\beta\mathcal{H}(\boldsymbol{x})}, \tag{2.14}$$

where $\mathcal{H}(\boldsymbol{x}) = \mathcal{H}(\boldsymbol{q}, \boldsymbol{p})$ is the Hamiltonian of the system, $\boldsymbol{x} = (\boldsymbol{q}, \boldsymbol{p})$ are generalized coordinates/momenta, and Z is the partition function. For other systems, ρ can be, e.g., the probability density function of the random initial state (see Eq. (2.1)). Next, suppose that each observable $u_i(\boldsymbol{x})$ $(i = 1, \dots, M)$ belongs to the linear space $\mathcal{P}H \cap \mathcal{D}(\mathcal{L})$, where $\mathcal{P}H = V$ and $\mathcal{D}(\mathcal{L})$ denotes the domain of the Liouville operator $\mathcal{L}$ defined in (2.3). The MZ equation (2.5), with $\mathcal{P}$ defined in (2.13), reduces to

$$\frac{d\boldsymbol{u}(t)}{dt} = \boldsymbol{\Omega}\boldsymbol{u}(t) + \int_0^t \boldsymbol{K}(t-s)\boldsymbol{u}(s)ds + \boldsymbol{f}(t), \tag{2.15}$$

where[2]

$$G_{ij} = \langle u_i(0), u_j(0)\rangle_\rho \quad \text{(Gram matrix)}, \tag{2.16a}$$

$$\Omega_{ij} = \sum_{k=1}^{M} G_{jk}^{-1} \langle u_k(0), \mathcal{L}u_i(0)\rangle_\rho \quad \text{(streaming matrix)}, \tag{2.16b}$$

$$K_{ij}(t) = \sum_{k=1}^{M} G_{jk}^{-1} \langle u_k(0), \mathcal{L}e^{t\mathcal{QL}}\mathcal{QL}u_i(0)\rangle_\rho \quad \text{(memory kernel)}, \tag{2.16c}$$

$$\boldsymbol{f}(t) = e^{t\mathcal{QL}}\mathcal{QL}\boldsymbol{u}(0) \quad \text{(fluctuation term)}. \tag{2.16d}$$

[2]Note that the ith component of the system (2.15) can be explicitly written as

$$\frac{du_i(t)}{dt} = \sum_{j=1}^{M} \Omega_{ij}u_j(t) + \sum_{j=1}^{M}\int_0^t K_{ij}(t-s)u_j(s)ds + f_i(t).$$

Equation (2.15) is often referred to as generalized Langevin equation (GLE) in classical statistical physics and other disciplines [83]. By applying Mori's projection to (2.15) we obtain the following linear (and closed) evolution equation for the projected phase space function

$$\frac{d}{dt}\mathcal{P}\boldsymbol{u}(t) = \boldsymbol{\Omega}\mathcal{P}\boldsymbol{u}(t) + \int_0^t \boldsymbol{K}(t-s)\mathcal{P}\boldsymbol{u}(s)\, ds. \tag{2.17}$$

Acting with the inner product $\langle u_j(0), \cdot\rangle_\rho$ on both sides of equation (2.17), yields the following exact equation for the temporal auto-correlation matrix $C_{ij}(t) = \langle u_j(0), u_i(t)\rangle_\rho$

$$\frac{d}{dt}C_{ij}(t) = \sum_{k=1}^{M} \Omega_{ik}C_{kj}(t) + \sum_{k=1}^{M}\int_0^t K_{ik}(t-s)C_{kj}(s)ds. \tag{2.18}$$

Suppose that the system (2.1) is Hamiltonian, and that the random initial state $\boldsymbol{x}_0$ is distributed according to the Boltzmann-Gibbs distribution (2.14), i.e., $\rho_0 = \rho_{eq}$. In these assumptions, the Liouville operator $\mathcal{L}$ is skew-adjoint relative to the inner product $\langle\rangle_{eq} = \int \rho_{eq} dp dq$, i.e., we have

$$\langle f, \mathcal{L}g\rangle_{eq} = -\langle \mathcal{L}f, g\rangle_{eq} \qquad f, g \in L^2(\mathcal{M}, \rho_{eq}) \cap \mathcal{D}(\mathcal{L}). \tag{2.19}$$

This allows us to simplify the expression of the memory kernel (2.16c) as

$$\begin{aligned} K_{ij}(t) &= -\sum_{k=1}^{M} G^{-1}_{jk}\langle \mathcal{Q}\mathcal{L}u_k(0), e^{t\mathcal{Q}\mathcal{L}}\mathcal{Q}\mathcal{L}u_i(0)\rangle_{eq}, \\ &= -\sum_{k=1}^{M} G^{-1}_{jk}\langle f_k(0), f_i(t)\rangle_{eq}, \end{aligned} \tag{2.20}$$

where $f_k(t)$ is the k-th component of the fluctuation term (2.16d). The identity (2.20) is known as Kubo's second fluctuation-dissipation theorem [54]. There are several advantages in using Mori's projection (2.13) over other projection operators, e.g., Chorin's projection [17]. For example, the MZ equation corresponding to Mori's projection is linear and closed, which allows us perform rigorous convergence analysis [104, 103]. Secondly, the streaming matrix (2.16b) and the memory kernel (2.16c) are exactly the same for both the projected and the unprojected equations ,i.e., (2.17) and (2.15)). Thirdly, we have that the second-fluctuation dissipation theorem (2.20) holds true, which allows us to express the MZ memory kernel in a relatively simple form, i.e., in terms of averages of random forces.

2.2 Mori-Zwanzig equation for stochastic system

In this section, we derive the MZ equation for stochastic systems. In particular, we discuss the stochastic differential equations (SDEs) deriven by white noise. In [22], Dominy and Venturi already developed a rather abstract Mori-Zwanzig framework for N-dimensional stochastic dynamical systems using the language of C^* algebra. In this section, we formulated a MZ framework in a self-consistent way without the usage of C^* algebra. This part can be viewed as a generalization of some previous MZ formulation for SDEs [29, 44]. To this end, let us consider a d-dimensional stochastic differential equation evolving in a smooth manifold $\mathcal{M}$

$$\frac{d}{dt}x(t) = F(x(t)) + \sigma(x(t))\xi(t), \qquad x(0) \sim \rho_0(x) \tag{2.21}$$

where $F(x(t)) : [0, T] \times \Omega \to \mathbb{R}^d$ is a d-dimensional vector function, $\sigma(x(t)) : [0, T] \times \Omega \to \mathbb{R}^{d\times m}$ is a $d \times m$ matrix function, the m-dimensional vector-valued Gaussian white noise $\xi(t)$ satisfies $\langle \xi_i(t), \xi_j(s)\rangle = \delta_{ts}\delta_{ij}$ for $1 \leq i, j \leq d$. $x(0)$ is the initial state of the system, which is characterized by a probability distribution

$\rho_0(x)$. For system with deterministic initial condition, $\rho_0(x) = \prod_{i=1}^{N} \delta(x_i - x_i(0))$. To derive the Mori-Zwanzig equation in a self-consistent and convenient way, here we follow physicists' notation [82] to write the SDE in terms of white noise $\xi(t)$, instead of the differential of the Wiener process $d\mathcal{W}(t)$. The strong solutions of (2.21) for different initial positions may be pieced together to give a random process $\{\phi_{t,s}, t \geq s \geq 0\}$ with values in Diff$(\mathcal{M})$, the space of smooth diffeomorphisms. The process $\{\phi_{t,s}, t \geq s \geq 0\}$ is called the *stochastic flow of diffeomophisms* associated with (2.21) in sense of Kunita [55, 101]. Similar to the deterministic case, we define a (stochastic) Koopman operator $\mathcal{E}(t,s)$ associated with the stochastic flow $\{\phi_{t,s}, t \geq s \geq 0\}$ such that any phase space function $u = u(x(t))$ can be expressed in terms of the operator acting on the space of observable, i.e.,

$$u(x(t)) = \mathcal{E}(t,s)u(x(s)) = u \bullet \phi_{t,s}(x(s)), \tag{2.22}$$

where $\bullet$ is the composition operator in the common sense. In physics, $\mathcal{E}(t,s)$ is normally understood as the evolution operator for some dynamical process, which can be formally written as

$$\mathcal{E}(t,s) = \overrightarrow{\mathcal{T}} e^{\int_s^t \mathcal{L}_\omega(\tau)d\tau},$$

where $\overrightarrow{\mathcal{T}}$ is the time ordering operator placing later operator to the right, $\mathcal{L}_\omega(\tau)$ is the (random) infinitesimal generator of the Koopman operator. The problem has a dual construction in the space of probability density function. The probability density $\rho(t) = \rho(u(x(t)))$ gives a full description of the statistical properties of observable $u(x(t))$. The evolution operator for $\rho(t)$ is given by the transfer

(Perron-Frobenius) operator associated with the flow map, which is

$$\rho(t) = \overleftarrow{\mathcal{T}} e^{\int_s^t \mathcal{L}_\omega^*(\tau)d\tau} \rho(s) = \mathcal{E}^*(t,s)u(x(s)). \tag{2.23}$$

The Perron-Frobenius operator $\mathcal{E}^*(t,s)$ and its infinitesimal generator $\mathcal{L}_\omega^*(\tau)$ are the dual of the Koopman operator $\mathcal{E}(t,s)$ and its generator $\mathcal{L}_\omega(\tau)$ in some suitable function space. The duality between these two construction and their equivalence in describing the dynamics of the system (2.21) are well summarized in [22]. From a physical point of view, the formulation in the phase space can be interpreted as the Schrödinger picture, while the dual construction in the probability function space can be interpreted as the Heisenberg picture. In this paper, we shall concentrate on the Schrödinger picture, i.e. the phase space formulation (2.22).

When interpreted as the evolution operator in physical sense, the infinitesimal generator $\mathcal{L}_\omega(\tau)$ of the Koopman operator $\mathcal{E}(t,s)$ can be determined by Dyson's series expansion [77] and following the listed procedure. We first write the evolution operator $\mathcal{E}(t,0)$ in (2.22) as a Dyson series

$$\mathcal{E}(t,s) = \sum_{n=1}^{+\infty} \mathcal{D}_n(t,s), \tag{2.24}$$

where

$$\mathcal{D}_n(t,s) = \overrightarrow{\mathcal{T}} \frac{1}{n!} \int_s^t \int_s^t \cdots \int_s^t dt_1 dt_2 \cdots dt_n \mathcal{L}_\omega(t_1)\mathcal{L}_\omega(t_2) \cdots \mathcal{L}_\omega(t_n).$$

Then the series expansion of the observable $u(x(t))$

$$u(x(t)) = \sum_{n=1}^{+\infty} \mathcal{D}_n(t,s) u(x(s)) \tag{2.25}$$

should match the stochastic Taylor expansion (See [52], Chapter 5) of $u(x(t))$. By comparing term by term with the Taylor series, one can get the explicit expression of $\mathcal{L}_\omega(\tau)$. Specifically, if the SDE (2.21) is interpreted in the Itô sense, using the Itô-Taylor expansion, one can get $\mathcal{L}_\omega(t) = \mathcal{L}^I_\omega(t)$ which is explicitly given by

$$\mathcal{L}^I_\omega(t) = \sum_{k=1}^{d} F_k \frac{\partial}{\partial x_{k0}} + \frac{1}{2}\sum_{j=1}^{m}\sum_{i,k=1}^{d} \sigma_{i,j}\sigma_{k,j}\frac{\partial^2}{\partial x_{i0}\partial x_{k0}} + \sum_{j=1}^{m}\sum_{i=1}^{d}\sigma_{i,j}\frac{\partial}{\partial x_{i0}}\xi_j(t),$$

where $x_{k0} = x_k(0)$, $F_k = F_k(x(0))$ and $\sigma_{ij} = \sigma_{ij}(x(0))$. The superscript I indicates the Itô's interpretation. Similarly, using the Stratonovich type Taylor expansion, one can get that the infinitesimal generator under Stratonovich's interpretation is given by

$$\mathcal{L}^S_\omega(t) = \sum_{k=1}^{d} F_k \frac{\partial}{\partial x_{k0}} + \sum_{j=1}^{m}\sum_{i=1}^{d}\sigma_{i,j}\frac{\partial}{\partial x_{i0}} \circ \xi_j(t),$$

where $\circ\xi_j(t)$ stands for the stochastic increment under Stratonovich interpretation (Again, we used physicists' notation $\circ\xi_j(t)$ instead of $\circ d\mathcal{W}_j(t)$). With the specific form of the stochastic generator $\mathcal{L}_\omega$ determined, we now follow the procedure in [22] to derive the phase space Mori-Zwanzig equation. First, we introduce a projection operator $\mathcal{P}$. For now we assume that $\mathcal{P}$ is defined on some rather general Banach space satisfying $\mathcal{P}^2 = \mathcal{P}$. We also denote the complementary operator $\mathcal{Q} = \mathcal{I} - \mathcal{P}$, where $\mathcal{I}$ is the identity operator. By differentiating the time-dependent Dyson's identity

$$\mathcal{E}(t,0) = \mathcal{Y}(t,0) + \int_0^t \mathcal{E}(s,0)\mathcal{P}\mathcal{L}_\omega(s)\mathcal{Y}(t-s,0)ds,$$

where $\mathcal{Y}(t,s)$ denotes the orthogonal evolution operator $\overrightarrow{\mathcal{T}} e^{\int_s^t \mathcal{Q}\mathcal{L}_\omega(\tau)d\tau}$, we shall get

$$\frac{d}{dt}\mathcal{E}(t,0) = \mathcal{E}(t,0)\mathcal{L}_\omega(t) + \mathcal{Y}(t,0)\mathcal{Q}\mathcal{L}_\omega(t) + \int_0^t \mathcal{E}(s,0)\mathcal{P}\mathcal{L}_\omega(s)\mathcal{Y}(t,s)\mathcal{Q}\mathcal{L}_\omega(t)ds. \tag{2.26}$$

Applying the operator equation (2.26) into a general observable $u(x(0))$ in the phase space yields the time-dependent MZ equation for the full dynamics

$$\begin{aligned}\frac{d}{dt}u(x(t,\omega)) = \mathcal{E}(t,0)\mathcal{P}\mathcal{L}_\omega(t)u(x(0,\omega)) + \mathcal{Y}(t,0)\mathcal{Q}\mathcal{L}_\omega(t)u(x(0,\omega))\\ + \int_0^t \mathcal{E}(s,0)\mathcal{P}\mathcal{L}_\omega(s)\mathcal{Y}(t,s)\mathcal{Q}\mathcal{L}_\omega(t)u(x(0,\omega))ds\end{aligned} \tag{2.27}$$

In practice, a projected version of the MZ equation is also frequently used. To get this, we apply the projection operator $\mathcal{P}$ in two sides of the equation (2.27) and get

$$\begin{aligned}\frac{d}{dt}\mathcal{P}u(x(t,\omega)) =& \mathcal{P}\mathcal{E}(t,0)\mathcal{P}\mathcal{L}_\omega(t)u(x(0,\omega))\\ &+ \int_0^t \mathcal{P}\mathcal{E}(s,0)\mathcal{P}\mathcal{L}_\omega(s)\mathcal{Y}(t,s)\mathcal{Q}\mathcal{L}_\omega(t)u(x(0,\omega))ds,\end{aligned} \tag{2.28}$$

where the projected fluctuating force $\mathcal{P}\mathcal{Y}(t,0)\mathcal{Q}\mathcal{L}_\omega(t)u(x(0,\omega))$ vanished since $\mathcal{P}\mathcal{Q} = 0$. The time dependent MZ equation in the dual space is given in [22], which is not explicitly stated here. In the context of classical dynamical systems, the MZ equation (2.27) and (2.28) describe, respectively, the evolution of a phase space function (observable) $u(x(t,\omega))$ and projected quantity $\mathcal{P}u(x(t,\omega))$, where the physical meaning of $\mathcal{P}u(x(t,\omega))$ is associated with the projection operator $\mathcal{P}$ that is used.

If the projection operator $\mathcal{P}$ is specified to be a Mori-type, finite rank projection operator in some Hilbert space, one can get a matrix form, linear generalized

Langevin equation (GLE) from the Mori-Zwanzig equation (2.27) and (2.28) [22].

2.3 Effective Mori-Zwanzig equation

The Mori-Zwanzig equation we derived in the previous subsection provided a general framework to handle dimension reduction problems for stochastic systems. However, it is hard to use directly because of the existence of the time dependent, white noise term $\xi_i(t)$ in the generator $\mathcal{L}_\omega(t)$ and $\mathcal{Q}\mathcal{L}_\omega(t)$. Such stochastic term exists essentially because the quantity of interest $u(x(t))$, as a function of $x(t) = x(t, \omega)$, is a stochatic process determined by the SDE (2.21). As we mentioned before, the MZ equation (2.27) fully described the dynamics of the stochastic process $u(t)$. In practice, however, such full description is often not needed since only some *determnistic* quantities of such stochastic process, such as the moments and correlation functions, are of theoretical importance. For instance, if $u(x(t))$ is proven to be a Gaussian process, knowing the first two moments is already enough to fully characterize the process itself. Even for the non-Gaussian processes, we will show in Chapter 6 that with the moment information, it is enough to create a biorthogonal series that well-approximated the original process. All these observations inspire us to derive an *effective Mori-Zwanzig equation* (EMZE) that describes the evolution of such deterministic quantities themselves. To this end, we first notice that for (2.21) the randomness introduced by the white noise $\xi(t)$ is often important for the first moment, this allow us to consider the Markovian semigroup $\mathcal{M}(t, 0)$ (see details in [80]) defined as

$$\mathcal{M}(t, 0)u(x(0)) := \mathbb{E}_{\xi(t)}[u(x(t))|x(0)], \tag{2.29}$$

where the $\mathbb{E}_{\xi(t)}[\cdot|x_0]$ is the conditional average with respect to $\xi(t)$ given the initial condition $x(0)$. To be noticed that the Markovian semigroup $\mathcal{M}(t,0)$ only encoded *one-point information*, which is the mean of the whole stochastic process under the noise $\xi(t)$. The standard stochastic analysis indicates that the infinitesimal generator of the Markovian semigroup $\mathcal{M}(t,0)$ is given by the following Komogorov (backward) operator [52]:

$$\begin{aligned}
\mathcal{K}^I &= \sum_{k=1}^{d} F_k(\boldsymbol{x})\frac{\partial}{\partial x_k} + \frac{1}{2}\sum_{j=1}^{m}\sum_{i,k=1}^{d}\sigma_{i,j}\sigma_{k,j}\frac{\partial}{\partial x_i \partial x_k} \\
\mathcal{K}^S &= \sum_{k=1}^{d}\left[F_k(\boldsymbol{x}) - \frac{1}{2}\sum_{j=1}^{m}\sigma_{k,j}\sum_{i=1}^{d}\sigma_{i,j}\frac{\partial}{\partial x_i}\right]\frac{\partial}{\partial x_k},
\end{aligned}$$

where the superscript I and S indicates whether it is Itô or Stratonovich interpretation. To be noticed that $\mathcal{K}^I$ and $\mathcal{K}^S$ are no longer time-dependent nor stochastic, hence the Markovian semigroup can be written as $\mathcal{M}(t,0) = e^{t\mathcal{K}}$ where $\mathcal{K}$ can be $\mathcal{K}^I$ or $\mathcal{K}^S$. For physical quantities that can be represented as $\mathcal{M}(t,0)x(0)$ or the function of it, by introducing some suitable projection operator $\mathcal{P}$ and then differentiating the Dyson's idenity

$$e^{t\mathcal{K}}u(0) = e^{t\mathcal{QK}}u(0) + \int_0^t e^{s\mathcal{K}}\mathcal{PK}e^{(t-s)\mathcal{QK}}u(0)ds$$

we shall get the following effective Mori-Zwanzig equation (EMZE) and projected EMZE:

$$\frac{\partial}{\partial t}e^{t\mathcal{K}}u(0) = e^{t\mathcal{K}}\mathcal{PK}u(0) + e^{t\mathcal{QKQ}}\mathcal{QK}u(0) + \int_0^t e^{s\mathcal{K}}\mathcal{PK}e^{(t-s)\mathcal{QKQ}}\mathcal{QK}u(0)ds. \tag{2.30}$$

$$\frac{\partial}{\partial t}\mathcal{P}e^{t\mathcal{K}}u(0) = \mathcal{P}e^{t\mathcal{K}}\mathcal{PK}u(0) + \int_0^t \mathcal{P}e^{s\mathcal{K}}\mathcal{PK}e^{(t-s)\mathcal{QKQ}}\mathcal{QK}u(0)ds. \tag{2.31}$$

The (projected) EMZE for the stochastic system are of the same form as the one for classical dynamical systems, i.e. eqn (2.5) and (2.6). The only difference is that the Liouville operator $\mathcal{L}$ for classical systems is now replaced by a Kolmogorov operator $\mathcal{K}$. Please note that we rewrite the orthogonal evolution operator $e^{t\mathcal{QK}}$ as $e^{t\mathcal{QKQ}}$ when it appears in the memory integral and the fluctuation term. This always can be done since $e^{t\mathcal{QK}}$ acts on $\mathcal{QK}u(0)$ and $\mathcal{Q}$ is a projection operator satisfying $\mathcal{Q}^2 = \mathcal{Q}$. As we will see in Section 4.1, evolution operator $e^{t\mathcal{QKQ}}$ is a better starting point to do mathematical analysis. We can similarly derive a linear generalized Langevin equation (GLE) based on EMZE. To this end, we consider a Hilbert space $H = L^2(\mathcal{M}, \rho)$, where ρ is some suitable Lebesgue measure. To construct a Mori-type projection $\mathcal{P}$, we first compute the positive-definite Gram matrix $G_{ij} = \langle u_i(0), u_j(0)\rangle_\rho$, where phase space functions $u_i \in \mathcal{H}$, with initial conditions in $u_i(0) \in D(\mathcal{P})$. With G_{ij} available, we define the Mori's projection as

$$\mathcal{P}f = \sum_{i,j=1}^{M} G_{ij}^{-1}\langle u_i(0), f\rangle_\rho u_j(0) \qquad \forall f \in H. \tag{2.32}$$

The projection operator $\mathcal{P}$ defined an orthogonal projection onto a closed linear subspace $V = \{u_i(0)\}_{i=1}^M \subset H$. Using the definition (2.32) we can explicitly write the system of EMZE for the phase space functions $\{u_1(t), \ldots, u_M(t)\}$ as

$$\frac{d}{dt}u(t) = \Omega u(t) + \int_0^t K(t-s)u(s)ds + f(t), \tag{2.33}$$

$$\frac{d}{dt}\mathcal{P}u(t) = \Omega\mathcal{P}u(t) + \int_0^t K(t-s)\mathcal{P}u(s)\, ds. \tag{2.34}$$

where $u(t) = [u_1(t), \dots, u_M(t)]^T$ and

$$G_{ij} = \langle u_i(0), u_j(0) \rangle_\rho, \tag{2.35a}$$

$$\Omega_{ij} = \sum_{k=1}^{M} G_{jk}^{-1} \langle u_k(0), \mathcal{K} u_i(0) \rangle_\rho \quad \text{(streaming matrix)}, \tag{2.35b}$$

$$K_{ij}(t) = \sum_{k=1}^{M} G_{jk}^{-1} \langle u_k(0), \mathcal{K} e^{t\mathcal{QKQ}} \mathcal{QK} u_i(0) \rangle_\rho \quad \text{(memory kernel)}, \tag{2.35c}$$

$$f(t) = e^{t\mathcal{QKQ}} \mathcal{QK} u(0) \quad \text{(fluctuation term)}. \tag{2.35d}$$

EMZE (2.33) and its projected form (2.34) are also referred as the generalized Langevin equation (GLE) in statistical physics. A slight different part is that $\mathcal{K}$ is no longer skew-adjoint, hence the memory kernel cannot be represented as the inner product $\langle f(t), f(0) \rangle_\rho$. Now we give examples of deterministic quantities that can be dealt with the EMZE/GLE. We first remark that very different from classical system which is lack of ergodicity, the stochastic system (2.21) often can be proved to be ergodic when vector field $F(x(t))$ and diffusion matrix $\sigma(x(t))$ satisfy some suitable regularity conditions. The ergodic state is well-described by a probability distribution ρ_{stat} which is the unique, stationary solution to the Fokker-Planck (or Kolmogorov forward) equation corresponding to the Markovian semigroup $\mathcal{M}(t, 0)$. Hence for the ergodic case, the semigroup action $\mathcal{M}(t, 0)u(x(0))$ describes the dynamics of the observable $u(x)$ evolving from an initial state ρ_0 to the stationary state ρ_{stat} as $t \to \infty$. In the following examples, we consider different initial conditions and discuss the physical meanings of the quantities $\mathcal{M}(t, 0)u(x(0))$ for each case. Please note that only the scalar observables are considered while without the loss of generality.

Non-stationary mean When the initial condition $u(x(0))$ is deterministic and $u(x(0)) \neq \mathbb{E}_{stat}[u(x(0))]$, or $u(x(0))$ is random and $\rho_0(u(x(0))) \neq \rho_{stat}(u(x_0))$, one can derive an EMZE/GLE for the *non-stationary mean* $\mathcal{M}(t,0)u(x(0))$ since it describes the non-equilibrium dynamics of $u(x)$ evolving toward the stationary state. We use the following notation to represent the mean function when the initial is deterministic

$$\langle u(x(t))\rangle = \mathcal{M}(t,0)u(x(0)) = \mathbb{E}_{\xi(t)}[u(x(t))|x(0)],$$

where the conditional expectation is with respect to the probability density of the white noise $\xi(t)$ at time t given the initial condition $x(0)$. The non-stationary mean is an important quantity that is frequently used in physics. For example, in [29], Español used the linear EMZE/GLE for the description of the near-equilibrium dynamics of DPD collective observables. More recently, Hudson and Li [44] used Zwanzig-type projection operator to derive the EMZE/GLE for the reaction coordinate of a overdamped Langevin dynamical system. All these previous results are restated and well-explained within the EMZ framework.

Stationary correlation function When the initial condition $u(x(0))$ is deterministic and $u(x(0)) = \mathbb{E}_{stat}[u(x(0))]$, or $u(x(0))$ is random and $\rho_0(u(x(0))) = \rho_{stat}(u(x(0)))$, the EMZE/GLE for $\mathcal{M}(t,0)u(x(0))$ describes the dynamics of $u(x(t))$ in the stationary state. The dynamics of the first case is trivial since $\mathcal{M}(t,0)u(x(0))$ for $u(x(0)) = \mathbb{E}_{stat}[u(x(0))]$ equals constant. For the second case, a physically meaningful quantity within the stationary process would be the *correlation function*. It is very known that [37] various transport coefficients of a dynamical process can be deduced from the time correlation functions through Green-Kubo formula. To derive the EMZE/GLE for it, we introduce the notation

$$\langle\langle(\cdot)\rangle\rangle := \mathbb{E}_{x(0)}[\mathbb{E}_{\xi(t)}[(\cdot)|x(0)]],$$

where the first expectation is with respect to the initial condition $x(0) \sim \rho_0$. Following [76, 82], we define the mean and time autocorrelation function for any scalar observable $u(t) = u(x(t))$ of (2.21) as:

$$\langle\langle u(t)\rangle\rangle := \mathbb{E}_{x(0)}[\mathbb{E}_{\xi(t)}[u(t)|x(0)]] \qquad \langle\langle u(t), u(s)\rangle\rangle := \mathbb{E}_{x(0)}[\mathbb{E}_{\xi(t)}[u(t)u(s)|x(0)]]$$

When the initial state of the system is the stationary state, i.e. $\rho_0 = \rho_{stat}$, equivalently we have

$$\langle\langle u(t)\rangle\rangle = \langle\langle u(0)\rangle\rangle = \mathbb{E}_{x(0)}[u(0)] \tag{2.36}$$

$$\langle\langle u(t+s), u(s)\rangle\rangle = \langle\langle u(t), u(0)\rangle\rangle = \mathbb{E}_{x(0)}[\mathbb{E}_{\xi(t)}[u(t)u(0)|x(0)]] \tag{2.37}$$

where (2.36) (2.37) give the stationary mean and correlation function respectively. By taking the time derivative of eqn (2.37), we obtain

$$\begin{aligned}
\frac{d}{dt}\langle\langle u(t), u(0)\rangle\rangle &= \frac{d}{dt}\mathbb{E}_{x(0)}[\mathbb{E}_{\xi(t)}[u(t)u(0)|x(0)]] \\
&= \frac{d}{dt}\mathbb{E}_{x(0)}[\mathbb{E}_{\xi(t)}[u(t)|x(0)]u(0)] \\
&= \left\langle \frac{d}{dt}\mathcal{M}(t,0)u(0), u(0)\right\rangle_{stat}
\end{aligned}$$

where the second inequality holds since the white noise is not imposed when $t = 0$. The inner product $\langle\cdot,\cdot\rangle_{stat}$ is taken with respect to the stationary distribution ρ_{stat}. If we introduce Mori-type projection operator $\mathcal{P}$ defined as

$$\mathcal{P}(\cdot) = \frac{\langle(\cdot), u(0)\rangle_{stat}}{\langle u(0), u(0)\rangle_{stat}}u(0), \tag{2.38}$$

we obtain that

$$\mathcal{P}\mathcal{M}(t,0)u(0) = \frac{\mathbb{E}_{x(0)}\left[\mathbb{E}_{\xi(t)}[u(t)|x(0)]u(0)\right]}{\langle u(0), u(0)\rangle_{stat}} u(0)$$

By taking the inner product $\langle(\cdot), u(0)\rangle_{stat}$ in two hand side of eqn (2.34), we get the following GLE for the correlation function $C(t) = \langle\langle u(t), u(0)\rangle\rangle$:

$$\frac{d}{dt}C(t) = \Omega C(t) - \int_0^t K(t-s)C(s)ds. \tag{2.39}$$

where Ω and $K(t)$ is the same as (2.35b) with Gram matrix G equals to 1. In principle, EMZE/GLE can be used to calculate the stationary autocorrelation function for any stochastic system of the form (2.21) assuming the existence of the unique, stationary state. Examples of such state are the Gibbs-type statistical equilibrium $\rho_{eq} \propto e^{-\beta\mathcal{H}}$ which arises from the Langevin dynamics, overdamped Langevin dynamics [12] and dissipative particle systems [29, 30], and the non-equilibrium steady state (NESS) of some heat conduction model [24, 26]. Although for the later case, the memory kernel (2.35c) cannot be determined using the first principle method since the the explicit form of the NESS is unkown, the EMZ equation (2.39) can still be used in a data-driven setting and method such as the bi-orthogonal decomposition can be used to get the low-dimensional stochastic models (see Section 6).

2.3.1 Application to Ornstein-Ulenbeck process

We give a simple example for which the Mori-type EMZE/GLE for the non-stationary mean and stationary correlation function can be written explicitly. To this end, we consider the Ornstein-Ulenbeck process which is defined by the Itô-type stochastic differential equation

$$\frac{d}{dt}x = \theta(\mu - x) + \sigma\xi(t) \tag{2.40}$$

where σ, μ, θ are all positive constant and ξ is the Gaussian white nosie satisfying $\langle \xi(t), \xi(s) \rangle = \delta(t-s)$. Ornstein-Ulenbeck process is proven to be ergodic and admits an stationary state given by the Gaussian distribution $\mathcal{N}(\mu, \sigma^2/2\theta)$. The formal solution of eqn (2.40) is given by

$$x(t) = x(0)e^{-\theta t} + \mu(1 - e^{-\theta t}) + \sigma \int_0^t e^{-\theta(t-s)}\xi(s)ds. \tag{2.41}$$

We choose observable $u(t) = x(t) - \mu$ as quantity of interest, the formal solution (2.41) enables us to get the analytic expression of the non-stationary mean and time autocorrelation function:

$$\langle x(t) - \mu \rangle = \mathbb{E}_{\xi(t)}[x(t) - \mu | x(0)] = (x(0) - \mu)e^{-\theta t} \tag{2.42}$$

$$\langle x(t) - \mu, x(s) - \mu \rangle = \mathbb{E}_{\xi(t)}[(x(t) - \mu)(x(s) - \mu)|x(0)] = \frac{\sigma^2}{2\theta}\left(e^{-\theta|t-s|} - e^{-\theta(t+s)}\right) \tag{2.43}$$

and the stationary mean and time autocorrelation function:

$$\langle\langle x(t) - \mu \rangle\rangle = \mathbb{E}_{x(0)}[\mathbb{E}_{\xi(t)}[x(t) - \mu | x(0)]] = \mathbb{E}_{x(0)}[x(0)e^{-\theta t} - \mu e^{-\theta t}] = 0 \tag{2.44}$$

$$\langle\langle (x(t) - \mu)(x(s) - \mu) \rangle\rangle = \mathbb{E}_{x(0)}[\mathbb{E}_{\xi(t)}[(x(t) - \mu)(x(s) - \mu)|x(0)]] = \frac{\sigma^2}{2\theta}e^{-\theta|t-s|} \tag{2.45}$$

One can verify that $\langle\langle (x(t+s) - \mu)(x(s) - \mu) \rangle\rangle = \langle\langle (x(t) - \mu)(x(0) - \mu) \rangle\rangle$, hence $C(t) = \frac{\sigma^2}{2\theta}e^{-\theta t}$.

Non-stationary mean For observable $u(t) = x(t) - \mu$, we define $\mathcal{P}$ as (2.38), by some simple calculation, we get $\Omega = -\theta$, $K(t) = 0$ and $f(t) = 0$ in eqn (2.33). Hence the EMZE/GLE for the non-stationary mean $M(t) = \langle x(t) - \mu \rangle$ is explicitly given by

$$\frac{d}{dt}M(t) = -\theta M(t), \tag{2.46}$$

with initial condition $M(0) = x(0) - \mu$. Obviously, The solution of eqn (2.46) gives analytic result (2.42).

Stationary correlation function For observable $u(t) = x(t) - \mu$, we define $\mathcal{P}$ as (2.38), by some simple calculation, we get $\Omega = -\theta$ and $K(t) = 0$ in eqn (2.39). Hence the EMZE/GLE for the stationary correlation function $C(t) = \langle\langle x(t) - \mu, x(0) - \mu \rangle\rangle$ is explicitly given by

$$\frac{d}{dt}C(t) = -\theta C(t) \tag{2.47}$$

with initial condition $C(0) = \sigma^2/2\theta$. Obviously, eqn (2.47) gives analytic result which agrees with (2.45).

2.4 Summary

In the first chapter, we introduced the Mori-Zwanzig equation and have shown how it can be derived for classical and stochastic dynamical systems. We wish to convey the following message to the reader : The Mori-Zwanzig theory is a general framework which is applicable for any dynamical process as long as one finds suitable descriptions to the flow and meanings to the projection operator.

Chapter 3

Analysis of the MZE for classical system

In this chapter, we will focus on the analysis of the MZE for classical dynamical systems. As we mentioned in Chapter 1, the main difficulty to apply the MZ framework in computation lies on the quantification of the memory integral. We will start our discussion with the estimate of the semigroup $e^{t\mathcal{L}}$ and $e^{t\mathcal{L}\mathcal{Q}}$. From which, we build rigorous estimates of the memory integral and provide convergence analysis of different approximation models of the Mori-Zwanzig equation, such as the short-memory approximation [84], the t-model [18], and hierarchical methods [86]. In particular, we study the MZ equation corresponding to infinite-rank projections (Section 2.1.1) and finite-rank projections (Section 2.1.1). Some numerical experiments are performed to verify the theoretical findings we get in this section. We will also show that similar analysis can be done for probability density function space formulations of the MZ equation and the analysis is almost the same. These two descriptions are connected by the same duality principle that pairs the Koopman and Frobenious-Perron operators. The estimate of PDF space MZ equation is listed as an appendix of this chapter.

3.1 Analysis of the Memory Integral

In this section, we develop a thorough mathematical analysis of the MZ memory integral

$$\int_0^t \mathcal{P}e^{s\mathcal{L}}\mathcal{P}\mathcal{L}e^{(t-s)\mathcal{Q}\mathcal{L}}\mathcal{Q}\mathcal{L}u_0 ds = \int_0^t \mathcal{P}e^{s\mathcal{L}}\mathcal{P}e^{(t-s)\mathcal{L}\mathcal{Q}}\mathcal{L}\mathcal{Q}\mathcal{L}u_0 ds. \tag{3.1}$$

and its approximation. We begin by describing the behavior of the semigroup norms $\|e^{t\mathcal{L}}\|$, $\|e^{t\mathcal{Q}\mathcal{L}\mathcal{Q}}\|$, and $\|e^{t\mathcal{L}\mathcal{Q}}\|$ as functions of time, for different choices of projection $\mathcal{P}$ and different norms. As we will see, the analysis will give clear computable bounds only in some circumstances, illustrating the difficulty of this problem and the need for further development and insight.

3.1.1 Semigroup Estimates

For any Liouville operator $\mathcal{L}$ of the form (2.3) acting on $\mathfrak{A} = L^\infty(\mathcal{M}, \Sigma, \mu)$ and for any σ identified with an element of $L^1(\mathcal{M}, \Sigma, \mu)$, the functional $\mathcal{L}_*\sigma$ (assuming σ lies in the domain of $\mathcal{L}_*$) is absolutely continuous with respect to σ (essentially because $\mathcal{L}$ acts locally) and the Radon-Nikodym derivative may be identified with the negative of the divergence of the vector field F with respect to the measure induced by σ, i.e.,

$$\frac{d\mathcal{L}_*\sigma}{d\sigma} = -\operatorname{div}_\sigma F, \qquad \text{i.e.,} \qquad (\mathcal{L}_*\sigma)(u) = -\sigma(u \operatorname{div}_\sigma F), \tag{3.2}$$

where

$$\operatorname{div}_\sigma F = \frac{\nabla \cdot (\tilde{\sigma}(x)F(x))}{\tilde{\sigma}(x)}. \tag{3.3}$$

When $\mathcal{M} = \mathbb{R}^N$ and σ has the form $\sigma(u) = \int_{\mathbb{R}^N} \tilde{\sigma}(x)u(x)dx$, this can be shown more directly using integration by parts. By assuming that $\tilde{\sigma}(x)$ or F_i decays to

0 at ∞, we have

$$\mathcal{L}_*\sigma(u) = \sigma(\mathcal{L}u) = \int_{\mathbb{R}^N} \tilde{\sigma} \sum_{i=1}^{N} F_i \frac{\partial u}{\partial x_i} dx \tag{3.4a}$$

$$= - \int_{\mathbb{R}^N} u \sum_{i=1}^{N} \frac{\partial}{\partial x_i} (\tilde{\sigma} F_i) dx = - \int_{\mathbb{R}^N} u \left[\frac{1}{\tilde{\sigma}} \sum_{i=1}^{N} \frac{\partial}{\partial x_i} (\tilde{\sigma} F_i) \right] \tilde{\sigma} dx \tag{3.4b}$$

$$= \sigma \left(u \left[- \frac{1}{\tilde{\sigma}} \sum_{i=1}^{N} \frac{\partial}{\partial x_i} (\tilde{\sigma} F_i) \right] \right) = \sigma \left(u \left[- \operatorname{div}_\sigma F \right] \right). \tag{3.4c}$$

This leads to

$$\operatorname{div}_\sigma F = \frac{\nabla \cdot (\tilde{\sigma} F)}{\tilde{\sigma}} = \frac{1}{\tilde{\sigma}} \sum_{i=1}^{N} \frac{\partial}{\partial x_i} (\tilde{\sigma} F_i) = \nabla \cdot F + F \cdot \nabla \left(\ln \tilde{\sigma} \right).$$

Therefore,

$$\begin{aligned} \langle v, -u \operatorname{div}_\sigma F \rangle_\sigma &= \sigma(-v^* u \operatorname{div}_\sigma F) = (\mathcal{L}_*\sigma)(v^* u) \\ &= \sigma(\mathcal{L}(v^* u)) = \sigma(\mathcal{L}(v)^* u + v^* \mathcal{L}(u)) = \langle \mathcal{L}v, u \rangle_\sigma + \langle v, \mathcal{L}u \rangle_\sigma, \end{aligned}$$

and the following are equivalent: i) $\mathcal{L}_*\sigma = 0$ (i.e., σ is invariant); ii) $\operatorname{div}_\sigma F = 0$; iii) $\mathcal{L}$ is skew-adjoint with respect to the σ inner product. More generally, on $\mathcal{H} = L^2(\mathcal{M}, \sigma)$, we find that

$$\mathcal{L} + \mathcal{L}^\dagger = - \operatorname{div}_\sigma F$$

so that the numerical abscissa ω [89, 90] of $\mathcal{L}$ is given by

$$\omega = \sup_{0 \neq u \in \mathcal{H}} \frac{\operatorname{Re}\langle u, \mathcal{L}u \rangle_\sigma}{\langle u, u \rangle_\sigma} = \sup_{0 \neq u \in \mathcal{H}} \frac{\langle u, (\mathcal{L} + \mathcal{L}^\dagger) u \rangle_\sigma}{2 \langle u, u \rangle_\sigma} \tag{3.5}$$

$$= \sup_{0 \neq u \in \mathcal{H}} \frac{\langle u, -u \operatorname{div}_\sigma F \rangle_\sigma}{2 \langle u, u \rangle_\sigma} = -\frac{1}{2} \inf_x \operatorname{div}_\sigma F(x). \tag{3.6}$$

Using this numerical abscissa ω, we obtain the following L^2_σ estimation of the Koopman semigroup

$$\|e^{t\mathcal{L}}\|_{L^2_\sigma} \leq e^{\omega t} \tag{3.7}$$

and moreover, $e^{\omega t}$ is the smallest exponential function that bounds $\|e^{t\mathcal{L}}\|_\sigma$ [20]. When $\mathcal{P}$ and $\mathcal{Q} = \mathcal{I} - \mathcal{P}$ are orthogonal projections on $L^2(\mathcal{M}, \sigma)$, we can observe that the numerical abscissa of $\mathcal{QLQ}$ is bounded by that of $\mathcal{L}$. In fact,

$$\begin{aligned}
\sup_{0 \neq u \in \mathcal{H}} \frac{\mathrm{Re}\langle u, \mathcal{QLQ}u\rangle_\sigma}{\langle u, u\rangle_\sigma} &= \sup_{0 \neq u \in \mathcal{H}} \frac{\mathrm{Re}\langle \mathcal{Q}u, \mathcal{LQ}u\rangle_\sigma}{\langle u, u\rangle_\sigma} \\
&= \sup_{0 \neq u \in \mathrm{Im}\,\mathcal{Q}} \frac{\mathrm{Re}\langle u, \mathcal{L}u\rangle_\sigma}{\langle u, u\rangle_\sigma} \leq \sup_{0 \neq u \in \mathcal{H}} \frac{\mathrm{Re}\langle u, \mathcal{L}u\rangle_\sigma}{\langle u, u\rangle_\sigma} = \omega,
\end{aligned}$$

(see equation (3.5)) so that

$$\|e^{t\mathcal{QLQ}}\|_{L^2_\sigma} \leq e^{\omega t}. \tag{3.8}$$

It should be noticed that this bound for the orthogonal semigroup is not necessarily tight. The tightness of the bound depends on the choice of projection $\mathcal{P}$ and comes down to whether functions in the image of $\mathcal{Q}$ can be chosen localized to regions where $\mathrm{div}_\sigma F$ is close to its infimal value.

When estimating the MZ memory integral, we need to deal with the semigroup $e^{t\mathcal{LQ}}$. It turns out to be extremely difficult to prove strong continuity of such semigroup in general, due to the unboundedness of $\mathcal{LQ}$. It is shown in appendix 3.A that, when $\mathcal{PLQ}$ is an unbounded operator, as is typical when $\mathcal{P}$ is a conditional expectation, the semigroup $e^{t\mathcal{LQ}}$ can only be bounded as

$$\|e^{t\mathcal{LQ}}\|_\sigma \leq M_\mathcal{Q} e^{t\omega_\mathcal{Q}} \tag{3.9}$$

for some $M_{\mathcal{Q}} > 1$, due to the fact that $\|e^{t\mathcal{L}\mathcal{Q}}\|_\sigma$ has infinite slope at $t = 0$. More work is needed to obtain satisfactory, computable values for $M_{\mathcal{Q}}$ and $\omega_{\mathcal{Q}}$, in the case where $\mathcal{P}$ and $\mathcal{Q}$ are infinite-rank projections. It is also shown that, when either $\mathcal{P}\mathcal{L}\mathcal{Q}$ or $\mathcal{L}\mathcal{P}$ is bounded, for example when $\mathcal{P}$ is a finite-rank projection, we can get computable semigroup bounds of the form

$$\|e^{t\mathcal{L}\mathcal{Q}}\|_\sigma \leq e^{\omega_{\mathcal{Q}} t} \leq e^{\frac{1}{2}\left(\sqrt{\omega^2 + \|\mathcal{P}\mathcal{L}\mathcal{Q}\|_\sigma^2} + \omega\right)t}, \tag{3.10a}$$

$$\|e^{t\mathcal{L}\mathcal{Q}}\|_\sigma \leq e^{\omega_{\mathcal{Q}} t} \leq e^{(\omega + \|\mathcal{L}\mathcal{P}\|_\sigma)t}, \tag{3.10b}$$

where $\omega = -\inf \operatorname{div}_\sigma F$ if we use the L^2_σ estimation of $e^{t\mathcal{L}}$.

3.1.2 Memory Growth

We begin by seeking to bound the MZ memory integral (3.1) as a whole and build our analysis from there. A key assumption of our analysis is that the semigroup $e^{t\mathcal{L}\mathcal{Q}}$ is strongly continuous[1], i.e., the map $t \mapsto e^{t\mathcal{L}\mathcal{Q}} g$ is continuous in the norm topology on the space of observables for each fixed g [28]. However, as we pointed out in section 3.1.1, it is a difficult task both to prove strong continuity of $e^{t\mathcal{L}\mathcal{Q}}$ and to obtain a computable upper bound for unbounded generators of the form $\mathcal{L}\mathcal{Q}$, we leave this as an open problem and assume that there exist constants $M_{\mathcal{Q}}$ and $\omega_{\mathcal{Q}}$ such that $\|e^{t\mathcal{L}\mathcal{Q}}\| \leq M_{\mathcal{Q}} e^{t\omega_{\mathcal{Q}}}$. Throughout this section, $\|\cdot\|$ denotes a general Banach norm. We begin with the following simple estimate:

Theorem 3.1.1. (Memory growth) *Let $e^{t\mathcal{L}}$ and $e^{t\mathcal{L}\mathcal{Q}}$ be strongly continuous semigroups with upper bounds $\|e^{t\mathcal{L}}\| \leq M e^{t\omega}$ and $\|e^{t\mathcal{L}\mathcal{Q}}\| \leq M_{\mathcal{Q}} e^{t\omega_{\mathcal{Q}}}$. Then*

$$\left\| \int_0^t \mathcal{P} e^{s\mathcal{L}} \mathcal{P}\mathcal{L} e^{(t-s)\mathcal{Q}\mathcal{L}} \mathcal{Q}\mathcal{L} u_0 ds \right\| \leq M_0(t), \tag{3.11}$$

[1]As is well known, $e^{t\mathcal{L}}$ (Koopman operator) is typically strongly continuous [28]. However, no such result exists for $e^{t\mathcal{L}\mathcal{Q}}$.

where

$$M_0(t) = \begin{cases} C_1 t e^{t\omega_{\mathcal{Q}}}, & \omega = \omega_{\mathcal{Q}} \\ \dfrac{C_1}{\omega - \omega_{\mathcal{Q}}}[e^{t\omega} - e^{t\omega_{\mathcal{Q}}}], & \omega \neq \omega_{\mathcal{Q}} \end{cases} \tag{3.12}$$

and $C_1 = MM_{\mathcal{Q}}\|\mathcal{LQL}u_0\|$ *is a constant. Clearly,* $\lim_{t\to 0} M_0(t) = 0$*.*

Proof. We first rewrite the memory integral in the equivalent form

$$\int_0^t \mathcal{P}e^{s\mathcal{L}}\mathcal{PL}e^{(t-s)\mathcal{QL}}\mathcal{QL}u_0 ds = \int_0^t \mathcal{P}e^{s\mathcal{L}}\mathcal{P}e^{(t-s)\mathcal{LQ}}\mathcal{LQL}u_0 ds.$$

Since $e^{t\mathcal{L}}$ and $e^{t\mathcal{LQ}}$ are assumed to be strongly continuous semigroups, we have the upper bounds $\|e^{t\mathcal{L}}\| \leq Me^{t\omega}$, $\|e^{t\mathcal{LQ}}\| \leq M_{\mathcal{Q}}e^{t\omega_{\mathcal{Q}}}$. Therefore

$$\begin{aligned} \left\| \int_0^t \mathcal{P}e^{s\mathcal{L}}\mathcal{P}e^{(t-s)\mathcal{LQ}}\mathcal{LQL}u_0 ds \right\| &\leq \int_0^t \|e^{s\mathcal{L}}\mathcal{P}e^{(t-s)\mathcal{LQ}}\mathcal{LQL}u_0\| ds \\ &\leq MM_{\mathcal{Q}}\|\mathcal{LQL}u_0\| \int_0^t e^{s(\omega-\omega_{\mathcal{Q}})} ds \\ &= \begin{cases} C_1 t e^{t\omega_{\mathcal{Q}}}, & \omega = \omega_{\mathcal{Q}} \\ \dfrac{C_1}{\omega - \omega_{\mathcal{Q}}}[e^{t\omega} - e^{t\omega_{\mathcal{Q}}}], & \omega \neq \omega_{\mathcal{Q}} \end{cases} \end{aligned}$$

where $C_1 = MM_{\mathcal{Q}}\|\mathcal{P}\|^2\|\mathcal{LQL}u_0\|$. □

Theorem 3.1.1 provides an upper bound for the growth of the memory integral based on the assumption that $e^{t\mathcal{L}}$ and $e^{t\mathcal{LQ}}$ are strongly continuous semigroups. We emphasize that only for simple cases can such upper bounds can be computed analytically (we will compute one of the cases later in section 3.2), because of the fundamental difficulties in computing the upper bound of $e^{t\mathcal{LQ}}$. However, it will be shown later that, although the specific expression for $M_0(t)$ is unknown, the *form* of it is already useful as it enables us to derive some verifiable theoretical predictions for general nonlinear systems.

3.1.3 Short Memory Approximation and the t-model

Theorem 3.1.1 can be employed to obtain upper bounds for well-known approximations of the memory integral. Let us begin with the t-model proposed in [18]. This model relies on the approximation

$$\int_0^t \mathcal{P}e^{s\mathcal{L}}\mathcal{P}\mathcal{L}e^{(t-s)\mathcal{Q}\mathcal{L}}\mathcal{Q}\mathcal{L}u_0 ds \simeq te^{t\mathcal{L}}\mathcal{P}\mathcal{L}\mathcal{Q}\mathcal{L}u_0 \qquad (t\text{-model}). \tag{3.13}$$

Theorem 3.1.2. (Memory approximation via the t-model [18]) *Let $e^{t\mathcal{L}}$ and $e^{t\mathcal{L}\mathcal{Q}}$ be strongly continuous semigroups with upper bounds $\|e^{t\mathcal{L}}\| \leq Me^{t\omega}$ and $\|e^{t\mathcal{L}\mathcal{Q}}\| \leq M_{\mathcal{Q}}e^{t\omega_{\mathcal{Q}}}$. Then*

$$\left\| \int_0^t \mathcal{P}e^{s\mathcal{L}}\mathcal{P}\mathcal{L}e^{(t-s)\mathcal{L}\mathcal{Q}}\mathcal{L}\mathcal{Q}\mathcal{L}u_0 ds - t\mathcal{P}e^{t\mathcal{L}}\mathcal{L}\mathcal{Q}\mathcal{L}u_0 \right\| \leq M_1(t),$$

where

$$M_1(t) = \begin{cases} C_1\left(\dfrac{e^{t\omega_{\mathcal{Q}}} - e^{t\omega}}{\omega_{\mathcal{Q}} - \omega} + \dfrac{te^{t\omega}}{M_{\mathcal{Q}}}\right) & \omega \neq \omega_{\mathcal{Q}} \\ C_1\dfrac{M_{\mathcal{Q}} + 1}{M_{\mathcal{Q}}}te^{t\omega} & \omega = \omega_{\mathcal{Q}} \end{cases},$$

and $C_1 = MM_{\mathcal{Q}}\|\mathcal{P}\|^2\|\mathcal{L}\mathcal{Q}\mathcal{L}u_0\|$.

Proof. By applying the triangle inequality, we obtain that

$$
\begin{aligned}
&\left\| \int_0^t \mathcal{P} e^{s\mathcal{L}} \mathcal{P} e^{(t-s)\mathcal{L}\mathcal{Q}} \mathcal{L}\mathcal{Q}\mathcal{L} u_0 ds - t\mathcal{P} e^{t\mathcal{L}} \mathcal{P}\mathcal{L}\mathcal{Q}\mathcal{L} u_0 \right\| \\
&\leq \left(\int_0^t \|\mathcal{P}\| \left\| e^{s\mathcal{L}} \right\| \|\mathcal{P}\| \left\| e^{(t-s)\mathcal{L}\mathcal{Q}} \right\| ds + t\|\mathcal{P}\| \left\| e^{t\mathcal{L}} \right\| \|\mathcal{P}\| \right) \|\mathcal{L}\mathcal{Q}\mathcal{L} u_0\| \\
&\leq \|\mathcal{P}\|^2 \|\mathcal{L}\mathcal{Q}\mathcal{L} u_0\| \left(M M_{\mathcal{Q}} \int_0^t e^{s\omega} e^{(t-s)\omega_{\mathcal{Q}}} ds + tMe^{t\omega} \right) \\
&= C_1 e^{t\omega} \left(\int_0^t e^{s(\omega_{\mathcal{Q}} - \omega)} ds + \frac{t}{M_{\mathcal{Q}}} \right) \\
&= \begin{cases} C_1 \left(\dfrac{e^{t\omega_{\mathcal{Q}}} - e^{t\omega}}{\omega_{\mathcal{Q}} - \omega} + \dfrac{te^{t\omega}}{M_{\mathcal{Q}}} \right) & \omega \neq \omega_{\mathcal{Q}} \\ C_1 \dfrac{M_{\mathcal{Q}} + 1}{M_{\mathcal{Q}}} te^{t\omega} & \omega = \omega_{\mathcal{Q}} \end{cases}
\end{aligned}
$$

where $C_1 = M M_{\mathcal{Q}} \|\mathcal{P}\|^2 \|\mathcal{L}\mathcal{Q}\mathcal{L} u_0\|$. □

Theorem 3.1.2 provides an upper bound for the error associated with the t-model. The limit

$$\lim_{t \to 0} M_1(t) = 0, \tag{3.14}$$

guarantees the convergence of the t-model for short integration times. On the other hand, depending on the semigroup constants M, ω, $M_{\mathcal{Q}}$ and $\omega_{\mathcal{Q}}$ (which may be estimated numerically), the error of the t-model may remain small for longer integration times (see the numerical results in section 3.2.2) Next, we study the short-memory approximation proposed in [84]. The main idea is to replace the integration interval $[0, t]$ in (3.1) by a shorter time interval $[t - \Delta t, t]$, i.e.

$$\int_0^t \mathcal{P} e^{s\mathcal{L}} \mathcal{P}\mathcal{L} e^{(t-s)\mathcal{Q}\mathcal{L}} \mathcal{Q}\mathcal{L} u_0 ds \simeq \int_{t-\Delta t}^t \mathcal{P} e^{s\mathcal{L}} \mathcal{P}\mathcal{L} e^{(t-s)\mathcal{Q}\mathcal{L}} \mathcal{Q}\mathcal{L} u_0 ds$$

(short-memory approximation),

where $\Delta t \in [0, t]$ identifies the effective *memory length*. The following result provides an upper bound to the error associated with the short-memory approximation.

Theorem 3.1.3. (Short memory approximation [84]) *Let $e^{t\mathcal{L}}$ and $e^{t\mathcal{L}\mathcal{Q}}$ be strongly continuous semigroups with upper bounds $\|e^{t\mathcal{L}}\| \leq Me^{t\omega}$ and $\|e^{t\mathcal{L}\mathcal{Q}}\| \leq M_{\mathcal{Q}}e^{t\omega_{\mathcal{Q}}}$. Then the following error estimate holds true*

$$\left\| \int_0^t \mathcal{P}e^{s\mathcal{L}}\mathcal{P}\mathcal{L}e^{(t-s)\mathcal{Q}\mathcal{L}}\mathcal{Q}\mathcal{L}u_0 ds - \int_{t-\Delta t}^t \mathcal{P}e^{s\mathcal{L}}\mathcal{P}\mathcal{L}e^{(t-s)\mathcal{Q}\mathcal{L}}\mathcal{Q}\mathcal{L}u_0 ds \right\| \leq M_2(t-\Delta t, t),$$

where

$$M_2(\Delta t, t) = \begin{cases} C_1(t-\Delta t)e^{t\omega_{\mathcal{Q}}} & \omega = \omega_{\mathcal{Q}} \\ C_1 e^{\Delta t\omega_{\mathcal{Q}}} \dfrac{e^{(t-\Delta t)\omega} - e^{(t-\Delta t)\omega_{\mathcal{Q}}}}{\omega - \omega_{\mathcal{Q}}} & \omega \neq \omega_{\mathcal{Q}} \end{cases}$$

and $C_1 = MM_{\mathcal{Q}}\|\mathcal{P}\|^2\|\mathcal{L}\mathcal{Q}\mathcal{L}u_0\|$.

We omit the proof due to its similarity to that of Theorem 3.1.1. Note that $\lim_{\Delta t\to t} M_2(\Delta t, t) = 0$ for all finite $t > 0$.

3.1.4 Hierarchical Memory Approximation

An alternative way to approximate the memory integral (3.1) was proposed by Stinis in [86]. The key idea is to repeatedly differentiate (3.1) with respect to time, and establish a hierarchy of PDEs which can eventually be truncated or approximated at some level to provide an approximation of the memory. In this section, we derive this hierarchy of memory equations and perform a thorough theoretical analysis to establish accuracy and convergence of the method. To this end, let us first define

$$w_0(t) = \int_0^t \mathcal{P}e^{s\mathcal{L}}\mathcal{P}\mathcal{L}e^{(t-s)\mathcal{Q}\mathcal{L}}\mathcal{Q}\mathcal{L}u_0 ds \tag{3.15}$$

to be the memory integral (3.1). We now assume that $w_0(t)$ is differentiable with respect to time. For the hierarchical approach to the finite memory approximation to be applicable, we must assume that $w_0(t)$ is differentiable with respect to time as many times as needed. By differentiating $w_0(t)$ with respect to time we obtain

$$\frac{dw_0(t)}{dt} = \mathcal{P}e^{t\mathcal{L}}\mathcal{P}\mathcal{L}\mathcal{Q}\mathcal{L}u_0 + w_1(t),$$

where

$$w_1(t) = \int_0^t \mathcal{P}e^{s\mathcal{L}}\mathcal{P}\mathcal{L}e^{(t-s)\mathcal{Q}\mathcal{L}}(\mathcal{Q}\mathcal{L})^2 u_0 ds.$$

By iterating this procedure n times we obtain

$$\frac{dw_{n-1}(t)}{dt} = \mathcal{P}e^{t\mathcal{L}}\mathcal{P}\mathcal{L}(\mathcal{Q}\mathcal{L})^{n-1}u_0 + w_n(t), \tag{3.16}$$

where

$$w_n(t) = \int_0^t \mathcal{P}e^{s\mathcal{L}}\mathcal{P}\mathcal{L}e^{(t-s)\mathcal{Q}\mathcal{L}}(\mathcal{Q}\mathcal{L})^{n+1}u_0 ds. \tag{3.17}$$

The hierarchy of equations (3.16)-(3.17) is equivalent to the following infinite-dimensional system of PDEs

$$\begin{cases} \dfrac{dw_0(t)}{dt} & = \mathcal{P}e^{t\mathcal{L}}\mathcal{P}\mathcal{L}\mathcal{Q}\mathcal{L}u_0 + w_1(t) \\ \dfrac{dw_1(t)}{dt} & = \mathcal{P}e^{t\mathcal{L}}\mathcal{P}\mathcal{L}\mathcal{Q}\mathcal{L}\mathcal{Q}\mathcal{L}u_0 + w_2(t) \\ & \vdots \\ \dfrac{dw_{n-1}(t)}{dt} & = \mathcal{P}e^{t\mathcal{L}}\mathcal{P}\mathcal{L}(\mathcal{Q}\mathcal{L})^n u_0 + w_n(t) \\ & \vdots \end{cases} \tag{3.18}$$

evolving from the initial condition $w_i(0) = 0$, $i = 1, 2, \ldots$ (see equation (3.17)). With such initial condition available, we can solve (3.18) with backward substitution, i.e., from the last equation to the first one, to obtain the following (exact) *Dyson series representation* of the memory integral (3.15)

$$\begin{aligned} w_0(t) = & \int_0^t \mathcal{P}e^{s\mathcal{L}}\mathcal{P}\mathcal{L}\mathcal{Q}\mathcal{L}u_0 ds + \int_0^t \int_0^{\tau_1} \mathcal{P}e^{s\mathcal{L}}\mathcal{P}\mathcal{L}\mathcal{Q}\mathcal{L}\mathcal{Q}\mathcal{L}u_0 ds d\tau_1 \\ & + \cdots + \int_0^t \int_0^{\tau_{n-1}} \cdots \int_0^{\tau_1} \mathcal{P}e^{s\mathcal{L}}\mathcal{P}\mathcal{L}(\mathcal{Q}\mathcal{L})^n u_0 ds d\tau_1 \ldots d\tau_{n-1} + \ldots . \end{aligned} \tag{3.19}$$

So far no approximation was introduced, i.e., the infinite-dimensional system (3.18) and the corresponding formal solution (3.19) are *exact*. To make progress in developing a computational scheme to estimate the memory integral (3.15), it is necessary to introduce approximations. The simplest of these rely on truncating the hierarchy (3.18) after n equations, while simultaneously introducing an approximation of the n-th order memory integral $w_n(t)$. We denote such an approximation as $w_n^{e_n}(t)$. The truncated system takes the form

$$\begin{cases} \dfrac{dw_0^n(t)}{dt} & = \mathcal{P}e^{t\mathcal{L}}\mathcal{P}\mathcal{L}\mathcal{Q}\mathcal{L}u_0 + w_1^n(t), \\ \dfrac{dw_1^n(t)}{dt} & = \mathcal{P}e^{t\mathcal{L}}\mathcal{P}\mathcal{L}\mathcal{Q}\mathcal{L}\mathcal{Q}\mathcal{L}u_0 + w_2^n(t), \\ & \vdots \\ \dfrac{dw_{n-1}^n(t)}{dt} & = \mathcal{P}e^{t\mathcal{L}}\mathcal{P}\mathcal{L}(\mathcal{Q}\mathcal{L})^n u_0 + w_n^{e_n}(t). \end{cases} \tag{3.20}$$

The notation $w_j^n(t)$ $(j = 0, .., n-1)$ emphasizes that the solution to (3.20) is, in general, different from the solution to (3.18). The initial condition of the system can be set as $w_i^n(0) = 0$, for all $i = 0, \ldots, n-1$. By using backward substitution, this yields the following formal solution

$$\begin{aligned} w_0^n(t) = &\int_0^t \mathcal{P}e^{s\mathcal{L}}\mathcal{P}\mathcal{L}\mathcal{Q}\mathcal{L}u_0 ds + \int_0^t \int_0^{\tau_1} \mathcal{P}e^{s\mathcal{L}}\mathcal{P}\mathcal{L}\mathcal{Q}\mathcal{L}\mathcal{Q}\mathcal{L}u_0 ds d\tau_1 \\ &+ \cdots + \int_0^t \int_0^{\tau_{n-1}} \cdots \int_0^{\tau_1} \mathcal{P}e^{s\mathcal{L}}\mathcal{P}\mathcal{L}(\mathcal{Q}\mathcal{L})^n u_0 ds d\tau_1 \ldots d\tau_{n-1} \\ &+ \int_0^t \int_0^{\tau_{n-1}} \cdots \int_0^{\tau_1} w_n^{e_n}(s) ds d\tau_1 \ldots d\tau_{n-1} \end{aligned} \tag{3.21}$$

representing an approximation of the memory integral (3.15). Note that, for a given system, such approximation depends only on the number of equations n in (3.20), and on the choice of approximation $w_n^{e_n}(t)$. In the present paper, we consider the following choices[2]

1. Approximation by truncation (H-model)

$$w_n^{e_n}(t) = 0. \tag{3.22}$$

[2]The quantities t_n and Δt_n appearing in (3.23) and (3.24) will be defined in Theorem 3.1.5 and Theorem 3.1.6, respectively.

2. Type-I finite memory approximation

$$w_n^{e_n}(t) = \int_{\max(0,t-\Delta t_n)}^{t} \mathcal{P}e^{s\mathcal{L}}\mathcal{P}\mathcal{L}e^{(t-s)\mathcal{Q}\mathcal{L}}(\mathcal{Q}\mathcal{L})^{n+1}u_0 ds. \tag{3.23}$$

3. Type-II finite memory approximation

$$w_n^{e_n}(t) = \int_{\min(t,t_n)}^{t} \mathcal{P}e^{s\mathcal{L}}\mathcal{P}\mathcal{L}e^{(t-s)\mathcal{Q}\mathcal{L}}(\mathcal{Q}\mathcal{L})^{n+1}u_0 ds. \tag{3.24}$$

4. H_t-model

$$w_n^{e_n}(t) = t\mathcal{P}e^{t\mathcal{L}}\mathcal{P}\mathcal{L}(\mathcal{Q}\mathcal{L})^{n+1}u_0. \tag{3.25}$$

The first approximation is a truncation of the hierarchy obtained by assuming that $w_n(t) = 0$. Such approximation was originally proposed by Stinis in [86], and we shall call it the H-model. The Type-I finite memory approximation (FMA) is obtained by applying the short memory approximation to the n-th order memory integral $w_n(t)$. The Type-II finite memory approximation (FMA) is a modified version of the Type-I, with a larger memory band. The H_t- model approximation is based on replacing the n-th order memory integral $w_n(t)$ with a classical t-model. Note that in this setting the classical t-model approximation proposed by Chorin and Stinis [18] is equivalent to a zeroth-order H_t-model approximation.

Hereafter, we present a thorough mathematical analysis that aims at estimating the error $\|w_0(t) - w_0^n(t)\|$, where $w_0(t)$ is full memory at time t (see (3.15) or (3.19)), while $w_0^n(t)$ is the solution of the truncated hierarchy (3.20), with $w_n^{e_n}(t)$ given by (3.22), (3.23), (3.24) or (3.25). With such error estimates available, we can infer whether the approximation of the full memory $w_0(t)$ with $w_0^n(t)$ is accurate and, more importantly, if the algorithm to approximate the memory integral converges. To the best of our knowledge, this is the first time a rigorous convergence analysis is performed on various approximations of the MZ memory integral. It turns out that the distance $\|w_0(t) - w_0^n(t)\|$ can be controlled through the construction of the hierarchy under some constraint on the initial condition.

The H-Model

Setting $w_n^{e_n}(t) = 0$ in (3.20) yields an approximation by truncation, which we will refer to as the H-model (hierarchical model). Such model was originally proposed by Stinis in [86]. Hereafter we provide error estimates and convergence results for this model. In particular, we derive an upper bound for the error $\|w_0(t) - w_0^n(t)\|$, and sufficient conditions for convergence of the reduced-order dynamical system. Such conditions are problem dependent, i.e., they involve the Liouvillian $\mathcal{L}$, the initial condition u_0, and the projection $\mathcal{P}$.

Theorem 3.1.4. (Accuracy of the H-model) *Let $e^{t\mathcal{L}}$ and $e^{t\mathcal{L}\mathcal{Q}}$ be strongly continuous semigroups with upper bounds $\|e^{t\mathcal{L}}\| \leq Me^{t\omega}$ and $\|e^{t\mathcal{L}\mathcal{Q}}\| \leq M_{\mathcal{Q}}e^{t\omega_{\mathcal{Q}}}$, and let $T > 0$ be a fixed integration time. For some fixed n, let*

$$\alpha_j = \frac{\|(\mathcal{L}\mathcal{Q})^{j+1}\mathcal{L}u_0\|}{\|(\mathcal{L}\mathcal{Q})^j\mathcal{L}u_0\|}, \quad 1 \leq j \leq n. \tag{3.26}$$

Then, for any $1 \leq p \leq n$ and all $t \in [0, T]$, we have

$$\|w_0(t) - w_0^p(t)\| \leq M_3^p(t) \leq M_3^p(T),$$

where

$$M_3^p(t) = C_1 A_1 A_2 \frac{t^{p+1}}{(p+1)!}\prod_{j=1}^{p}\alpha_j, \qquad C_1 = \|\mathcal{L}\mathcal{Q}\mathcal{L}u_0\| M M_{\mathcal{Q}},$$

and

$$A_1 = \max_{s\in[0,T]} e^{s(\omega-\omega_{\mathcal{Q}})} = \begin{cases} 1 & \omega \le \omega_{\mathcal{Q}} \\ e^{T(\omega-\omega_{\mathcal{Q}})} & \omega \ge \omega_{\mathcal{Q}} \end{cases} \quad A_2 = \max_{s\in[0,T]} e^{s\omega_{\mathcal{Q}}} = \begin{cases} 1 & \omega_{\mathcal{Q}} \le 0 \\ e^{T\omega_{\mathcal{Q}}} & \omega_{\mathcal{Q}} \ge 0 \end{cases}. \tag{3.27}$$

Proof. We begin with the expression for the difference between the memory term w_0 and its approximation w_0^p

$$w_0(t) - w_0^p(t) = \int_0^t \int_0^{\tau_p} \cdots \int_0^{\tau_2} \int_0^{\tau_1} \mathcal{P}e^{s\mathcal{L}}\mathcal{P}e^{(\tau_1-s)\mathcal{L}\mathcal{Q}}(\mathcal{L}\mathcal{Q})^{n+1}\mathcal{L}u_0 ds d\tau_1 \cdots d\tau_p. \tag{3.28}$$

Since $e^{t\mathcal{L}}$ and $e^{t\mathcal{L}\mathcal{Q}}$ are strongly continuous semigroups we have $\|e^{t\mathcal{L}}\| \le Me^{\omega t}$ and $\|e^{t\mathcal{L}\mathcal{Q}}\| \le M_{\mathcal{Q}}e^{\omega_{\mathcal{Q}} t}$. By using Cauchy's formula for repeated integration, we bound the norm of the error (3.28) as

$$\begin{aligned} \|w_0(t) - w_0^p(t)\| &\le \int_0^t \frac{(t-\sigma)^{p-1}}{(p-1)!} \int_0^\sigma \|\mathcal{P}e^{s\mathcal{L}}\mathcal{P}e^{(\sigma-s)\mathcal{L}\mathcal{Q}}(\mathcal{L}\mathcal{Q})^{p+1}\mathcal{L}u_0\| ds d\sigma \\ &\le \|\mathcal{P}\|^2 M M_{\mathcal{Q}} \|(\mathcal{L}\mathcal{Q})^{p+1}\mathcal{L}u_0\| \int_0^t \frac{(t-\sigma)^{p-1}}{(p-1)!} \int_0^\sigma e^{s\omega} e^{(\sigma-s)\omega_{\mathcal{Q}}} ds d\sigma \\ &\le C_1 \left(\prod_{j=1}^p \alpha_j\right) \underbrace{\int_0^t \frac{(t-\sigma)^{p-1}}{(p-1)!} \int_0^\sigma e^{s\omega} e^{(\sigma-s)\omega_{\mathcal{Q}}} ds d\sigma}_{f_p(t,\omega,\omega_{\mathcal{Q}})} \\ &= C_1 \left(\prod_{j=1}^p \alpha_j\right) f_p(t,\omega,\omega_{\mathcal{Q}}), \end{aligned} \tag{3.29}$$

where $C_1 = \|\mathcal{P}\|^2 \|\mathcal{L}\mathcal{Q}\mathcal{L}u_0\| M M_{\mathcal{Q}}$ as before. The function $f_p(t,\omega,\omega_{\mathcal{Q}})$, may be bounded from above as

$$
\begin{aligned}
f_p(t, \omega, \omega_{\mathcal{Q}}) &\leq A_1 A_2 \int_0^t \frac{(t-\sigma)^{p-1}}{(p-1)!} \int_0^\sigma ds d\sigma \\
&= A_1 A_2 \frac{t^{p+1}}{(p+1)!}.
\end{aligned}
$$

Hence, we have

$$
\|w_0(t) - w_0^p(t)\| \leq C_1 A_1 A_2 \left(\prod_{j=1}^{p} \alpha_j \right) \frac{t^{p+1}}{(p+1)!} = M_3^p(t).
$$

□

Theorem 3.1.4 states that for a given dynamical system (represented by $\mathcal{L}$) and quantity of interest (represented by $\mathcal{P}$) the error bound $M_3^p(t)$ is strongly related to $\{\alpha_j\}$ which is ultimately determined by the initial condition x_0. It turns out that by bounding $\{\alpha_j\}$, we can control $M_3^p(t)$, and therefore the overall error $\|w_0(t) - w_0^p(t)\|$. The following corollaries discuss sufficient conditions such that the error $\|w_0(T) - w_0^n(T)\|$ decays as we increase the differentiation order n for fixed time $T > 0$.

Corollary 3.1.4.1. (Uniform convergence of the H-model) *If $\{\alpha_j\}$ in Theorem 3.1.4 satisfy*

$$
\alpha_j < \frac{j+1}{T}, \quad 1 \leq j \leq n, \tag{3.30}
$$

for any fixed time $T > 0$, then there exists a sequence of constants $\delta_1 > \delta_2 > \cdots > \delta_n$ such that

$$
\|w_0(T) - w_0^p(T)\| \leq \delta_p \qquad 1 \leq p \leq n.
$$

Proof. Evaluating (3.29) at any fixed (finite) time $T > 0$ yields

$$\|w_0(T) - w_0^p(T)\| \leq C_2 \left(\prod_{j=1}^{p} \alpha_i\right) f_p(T, \omega, \omega_{\mathcal{Q}}) \leq C_2 \left(\prod_{j=1}^{p} \alpha_j\right) \frac{T^{p+1}}{(p+1)!},$$

$$\|w_0(T) - w_0^{p+1}(T)\| \leq C_2 \left(\prod_{j=1}^{p+1} \alpha_j\right) \frac{T^{p+2}}{(p+2)!},$$

where $C_2 = C_2(T) = C_1 A_1 A_2$. If there exists $\delta_p \geq 0$ such that

$$\|w_0(T) - w_0^p(T)\| \leq C_2 \left(\prod_{j=1}^{p} \alpha_j\right) \frac{T^{p+1}}{(p+1)!} \leq \delta_p,$$

then there exist a δ_{p+1} such that

$$\|w_0(T) - w_0^{p+1}(T)\| \leq C_2 \left(\prod_{j=1}^{p} \alpha_j\right) \frac{T^{p+1}}{(p+1)!} \frac{\alpha_{p+1} T}{p+2} \leq \delta_{p+1} < \delta_p,$$

since $\alpha_{p+1} < (p+2)/T$. Moreover, the condition $\alpha_j < (j+1)/T$ holds for all $1 \leq j \leq n$. Therefore, we conclude that for any fixed time $T > 0$, there exists a sequence of constants $\delta_1 > \delta_2 > \cdots > \delta_n$ such that $\|w_0(T) - w_0^p(T)\| \leq \delta_p$, where $1 \leq p \leq n$.

□

Corollary 3.1.4.1 provides a sufficient condition for the error $\|w_0(t) - w_0^p(t)\|$ to decrease monotonically as we increase p in (3.20). A stronger condition that yields an asymptotically decaying error bound is given by the following Corollary.

Corollary 3.1.4.2. (Asymptotic convergence of the H-model) *If α_j in Theorem 3.1.4 satisfies*

$$\alpha_j < C, \quad 1 \leq j < +\infty \tag{3.31}$$

for some positive constant C, then for any fixed time $T > 0$, and arbitrary $\delta > 0$, there exists a constant $1 \leq p < +\infty$ such that for all $n > p$,

$$\|w_0(T) - w_0^n(T)\| \leq \delta.$$

Proof. By introducing the condition $\alpha_j < C$ in the proof of Theorem 3.1.4 we obtain

$$\|w_0(T) - w_0^p(T)\| \leq C_2 \left(\prod_{j=1}^{p} \alpha_j \right) \frac{T^{p+1}}{(p+1)!} \leq C_2 T \frac{(CT)^p}{(p+1)!} \qquad \text{for all } 1 < p < +\infty.$$

The limit

$$\lim_{p \to +\infty} C_2 T \frac{(CT)^p}{(p+1)!} = 0$$

allows us to conclude that there exists a constant $1 < p < +\infty$ such that for all $n > p$, $\|w_0(T) - w_0^n(T)\| \leq \delta$. □

An interesting consequence of Corollary 3.1.4.2 is the existence of a *convergence barrier*, i.e., a "hump" in the error plot $\|w_0(T) - w_0^p(T)\|$ versus p generated by the H-model. While Corollary 3.1.4.2 only shows that behavior for an upper bound of the error, not directly the error itself, the feature is often found in the actual errors associated with numerical methods based on these ideas. The following Corollary shows that the requirements on $\{\alpha_j\}$ can be dropped (we still need $\alpha_j < +\infty$) if we consider relatively short integration times T.

Corollary 3.1.4.3. (Short-time convergence of the H-model) *For any integer n for which $\alpha_j < \infty$ for $1 \leq j \leq n$, and any sequence of constants*

$\delta_1 > \delta_2 > \cdots > \delta_n > 0$, there exists a fixed time $T > 0$ such that

$$\|w_0(T) - w_0^p(T)\| \leq \delta_p$$

for $1 \leq p \leq n$.

Proof. Since $\alpha_j < +\infty$, we can choose $C = \max_{1 \leq j \leq n} \alpha_j$. By following the same steps we used in the proof of Theorem 3.1.4, we conclude that, for

$$T \leq \frac{1}{C} \min_{1 \leq p \leq n} \left[\frac{C(p+1)!}{C_2} \delta_p \right]^{\frac{1}{p+1}},$$

the errors satisfy

$$\|w_0(T) - w_0^p(T)\| \leq C_2 \left(\prod_{j=1}^{p} \alpha_j \right) \frac{T^{p+1}}{(p+1)!} \leq \frac{C_2}{C} \frac{(CT)^{p+1}}{(p+1)!} \leq \delta_p$$

as desired, for all $1 \leq p \leq n$.

□

Corollary 3.1.4.1 and Corollary 3.1.4.2 provide sufficient conditions for the error $\|w_0(T) - w_0^n(T)\|$ generated by the H-model to decay as we increase the truncation order n. However, we still need to answer the important question of whether the H-model actually provides accurate results for a given nonlinear dynamics ($\mathcal{L}$), quantity of intererest ($\mathcal{P}$) and initial state x_0. Corollary 3.1.4.3 provides a partial answer to this question by showing that, at least in the short time period, condition (3.30) is always satisfied (assuming that $\{\alpha_j\}$ are finite). This guarantees the short-time convergence of the H-model for any reasonably smooth nonlinear dynamical system and almost any observable. However, for longer integration times T, convergence of the H-model for arbitrary nonlinear dynamical systems cannot be established in general, which means that we need to proceed on a case-by-case basis by applying Theorem 3.1.4 or by checking whether the hypotheses of Corollary 3.1.4.1 or Corollary 3.1.4.2 are satisfied[3]. On the other hand, convergence of the H-model can be established for any finite integration time in the case of linear dynamical systems, as we will show in Chapter 5.

Type-I Finite Memory Approximation (FMA)

The Type-I finite memory approximation is obtained by solving the system (3.20) with $w_n^{e_n}(t)$ given by (3.23). As before, we first derive an upper bound for $\|w_0(t) - w_0^n(t)\|$ and then discuss sufficient conditions for convergence. Such conditions basically control the growth of an upper bound on $\|w_0(t) - w_0^n(t)\|$.

Theorem 3.1.5. (Accuracy of the Type-I FMA) *Let $e^{t\mathcal{L}}$ and $e^{t\mathcal{L}\mathcal{Q}}$ be strongly continuous semigroups and let $T > 0$ be a fixed integration time. If*

$$\alpha_j = \frac{\|(\mathcal{L}\mathcal{Q})^{j+1}\mathcal{L}u_0\|}{\|(\mathcal{L}\mathcal{Q})^{j}\mathcal{L}u_0\|}, \quad 1 \leq j \leq n, \tag{3.32}$$

then for each $1 \leq p \leq n$ and for $\Delta t_p \leq t \leq T$

$$\|w_0(t) - w_0^p(t)\| \leq M_4^p(t),$$

where

$$M_4^p(t) = C_1 A_1 A_2 \left(\prod_{i=1}^{p} \alpha_i\right) \frac{(t-\Delta t_p)^{p+1}}{(p+1)!},$$

and C_1, A_1, A_2 are as in Theorem 3.1.4.

[3] The implementation of the H-model requires computing $(\mathcal{L}\mathcal{Q})^n\mathcal{L}x_0$ to high-order in n. This is not straightforward in nonlinear dynamical systems. However, such terms can be easily and effectively computed for linear dynamical systems. This yields a fast and practical memory approximation scheme for linear systems.

Proof. The error at the p-th level is of the form

$$w_p(t) - w_p^{e_p}(t) = \int_0^{\max(0,t-\Delta t_p)} \mathcal{P}e^{s\mathcal{L}}\mathcal{P}\mathcal{L}e^{(t-s)\mathcal{Q}\mathcal{L}}(\mathcal{Q}\mathcal{L})^{p+1}u_0 ds$$

and the error at the zeroth level is

$$\begin{aligned}
&w_0(t) - w_0^p(t)\\
&= \int_0^t \int_0^{\tau_p} \cdots \int_0^{\tau_2} \left[w_n(\tau_1) - w_n^{e_p}(\tau_1)\right] d\tau_1 \cdots d\tau_p\\
&= \int_0^t \int_0^{\tau_p} \cdots \int_0^{\tau_2} \int_0^{\max(0,\tau_1-\Delta t_p)} \mathcal{P}e^{s\mathcal{L}}\mathcal{P}e^{(\tau_1-s)\mathcal{L}\mathcal{Q}}(\mathcal{L}\mathcal{Q})^{p+1}\mathcal{L}u_0 ds d\tau_1 \cdots d\tau_p\\
&= \int_{\Delta t_p}^t \int_{\Delta t_p}^{\tau_p} \cdots \int_{\Delta t_p}^{\tau_2} \int_0^{\tau_1-\Delta t_p} \mathcal{P}e^{s\mathcal{L}}\mathcal{P}e^{(\tau_1-s)\mathcal{L}\mathcal{Q}}(\mathcal{L}\mathcal{Q})^{p+1}\mathcal{L}u_0 ds d\tau_1 \cdots d\tau_p\\
&= \int_{\Delta t_p}^t \int_{\Delta t_p}^{\tau_p} \cdots \int_0^{\tau_2-\Delta t_p} \int_0^{\tilde{\tau}_1} \mathcal{P}e^{s\mathcal{L}}\mathcal{P}e^{(\tilde{\tau}_1+\Delta t_p-s)\mathcal{L}\mathcal{Q}}(\mathcal{L}\mathcal{Q})^{p+1}\mathcal{L}u_0 ds d\tilde{\tau}_1 \cdots d\tau_p\\
&\ \vdots\\
&= \int_0^{\max(0,t-\Delta t_p)} \int_0^{\tilde{\tau}_p} \cdots \int_0^{\tilde{\tau}_2} \int_0^{\tilde{\tau}_1} \mathcal{P}e^{s\mathcal{L}}\mathcal{P}e^{(\tilde{\tau}_1+\Delta t_p-s)\mathcal{L}\mathcal{Q}}(\mathcal{L}\mathcal{Q})^{p+1}\mathcal{L}u_0 ds d\tilde{\tau}_1 \cdots d\tilde{\tau}_p.
\end{aligned}$$

The norm of this error may be bounded as

$$\begin{aligned}
&\|w_0(t) - w_0^p(t)\|\\
&\leq \int_0^{\max(0,t-\Delta t_p)} \int_0^{\tilde{\tau}_p} \cdots \int_0^{\tilde{\tau}_2} \int_0^{\tilde{\tau}_1} \left\|\mathcal{P}e^{s\mathcal{L}}\mathcal{P}e^{(\tilde{\tau}_1+\Delta t_p-s)\mathcal{L}\mathcal{Q}}(\mathcal{L}\mathcal{Q})^{p+1}\mathcal{L}u_0\right\| ds d\tilde{\tau}_1 \cdots d\tilde{\tau}_p\\
&\leq C_1 \left(\prod_{j=1}^p \alpha_j\right) \int_0^{\max(0,t-\Delta t_p)} \int_0^{\tilde{\tau}_p} \cdots \int_0^{\tilde{\tau}_2} \int_0^{\tilde{\tau}_1} e^{s(\omega-\omega_{\mathcal{Q}})}e^{(\tilde{\tau}_1+\Delta t_p)\omega_{\mathcal{Q}}} ds d\tilde{\tau}_1 \cdots d\tilde{\tau}_p\\
&\leq C_1 \left(\prod_{j=1}^p \alpha_j\right) f_p(t, \Delta t_p, \omega, \omega_{\mathcal{Q}}),
\end{aligned}$$

where

$$f_p(t, \Delta t_p, \omega, \omega_{\mathcal{Q}}) = \int_0^{\max(0,t-\Delta t_p)} \int_0^{\tilde{\tau}_p} \cdots \int_0^{\tilde{\tau}_2} \int_0^{\tilde{\tau}_1} e^{s(\omega-\omega_{\mathcal{Q}})}e^{(\tilde{\tau}_1+\Delta t_p)\omega_{\mathcal{Q}}} ds d\tilde{\tau}_1 \cdots d\tilde{\tau}_p.$$

If we bound f_p as

$$f_p(t, \Delta t_p, \omega, \omega_{\mathcal{Q}}) \leq A_1 A_2 \int_0^{\max(0, t-\Delta t_p)} \int_0^{\tilde{\tau}_p} \cdots \int_0^{\tilde{\tau}_2} \int_0^{\tilde{\tau}_1} ds d\tilde{\tau}_1 \cdots d\tilde{\tau}_p$$
$$= \begin{cases} 0 & 0 \leq t \leq \Delta t_p \\ A_1 A_2 \dfrac{(t-\Delta t_p)^{p+1}}{(p+1)!} & t \geq \Delta t_p \end{cases}$$

where A_1, A_2 are defined in (3.27), then we have that

$$\|w_0(t) - w_0^p(t)\| \leq C_1 A_1 A_2 \left(\prod_{j=1}^{p} \alpha_j \right) \frac{(t-\Delta t_p)^{p+1}}{(p+1)!} = M_4^p(t).$$

□

We notice that if the effective memory band at each level decreases as we increase the differentiation order p, then we can control the error $\|w_0(t) - w_0^n(t)\|$. The following corollary provides a sufficient condition that guarantees this sort of control of the error.

Corollary 3.1.5.1. (Uniform convergence of the Type-I FMA) *If α_j in Therorem 3.1.5 satisfy*

$$\alpha_j < (j+1) \left[\frac{\delta j!}{C_1 A_1 A_2 \left(\prod_{k=1}^{j-1} \alpha_k \right)} \right]^{-\frac{1}{j}} \qquad 1 \leq j \leq n \tag{3.37}$$

then for any $T > 0$ and $\delta > 0$, there exists an ordered sequence $\Delta t_n < \Delta t_{n-1} < \cdots < \Delta t_1 < T$ such that

$$\|w_0(T) - w_0^p(T)\| \leq \delta, \quad 1 \leq p \leq n,$$

and which satisfies

$$\Delta t_p \leq T - \left[\frac{\delta(p+1)!}{C_1 A_1 A_2 \left(\prod_{j=1}^{p} \alpha_j\right)} \right]^{\frac{1}{p+1}}. \tag{3.38}$$

Proof. For $1 \leq p \leq n$ we set

$$\|w_0(t) - w_0^p(t)\| \leq C_1 A_1 A_2 \left(\prod_{j=1}^{p} \alpha_j \right) \frac{(t - \Delta t_p)^{p+1}}{(p+1)!} \leq \delta.$$

This yields the following requirement on Δt_p

$$\Delta t_p \geq T - \left[\frac{\delta(p+1)!}{C_1 A_1 A_2 \left(\prod_{j=1}^{p} \alpha_j\right)} \right]^{\frac{1}{p+1}}. \tag{3.39}$$

Since hypothesis (3.37) holds, it is easy to check that the lower bound on each Δt_p satisfies

$$T - \left[\frac{\delta(p+1)!}{C_1 A_1 A_2 \left(\prod_{j=1}^{p} \alpha_j\right)} \right]^{\frac{1}{p+1}} < T - \left[\frac{\delta p!}{C_1 A_1 A_2 \left(\prod_{j=1}^{p-1} \alpha_j\right)} \right]^{\frac{1}{p}} \qquad \Delta t_p > \Delta t_{p-1}.$$

Therefore, by using the equality in (3.39) to define a sequence of Δt_n, we find that it is a decreasing time sequence $0 < \Delta t_n < \Delta t_{n-1} < \cdots < \Delta t_1 < T$ such that $\|w_0(T) - w_0^n(T)\| \leq \delta$ holds for all $t \in [0, T]$ and which satisfies (3.38). □

Remark The sufficient condition provided in Corollary 3.1.5.1 guarantees *uniform convergence* of the Type-I finite memory approximation. If we replace condition (3.37) with

$$\alpha_j < C, \qquad \text{for all} \qquad 1 \leq j < +\infty,$$

where C is a positive constant (independent on T), then we obtain asymptotic convergence. In other words, for each $\delta > 0$, there exists an integer p such that for all $n > p$ we have $\|w_0(t) - w_0^n(t)\| < \delta$. This result is based on the limit

$$\lim_{p\to+\infty} \frac{\delta(p+1)!}{C_1 A_1 A_2 \left(\prod_{j=1}^{p} \alpha_j\right)} > \lim_{p\to+\infty} \frac{\delta(p+1)!}{C_1 A_1 A_2 C^p} = +\infty$$

which guarantees the existence of an integer p for which the upper bound on Δt_p is smaller or equal to zero. In such case, the Type I FMA degenerates to the H-model, for which Corollary 3.1.4.2 holds.

Type-II Finite Memory Approximation

The Type-II finite memory approximation is obtained by solving the system (3.20) with $w_n^{e_n}(t)$ given in (3.24). We first derive an upper bound for $\|w_0(t) - w_0^n(t)\|$ and then discuss sufficient conditions for convergence.

Theorem 3.1.6. (Accuracy of the Type-II FMA) *Let $e^{t\mathcal{L}}$ and $e^{t\mathcal{L}\mathcal{Q}}$ be strongly continuous semigroups with upper bounds $\|e^{t\mathcal{L}}\| \leq M e^{t\omega}$ and $\|e^{t\mathcal{L}\mathcal{Q}}\| \leq M_{\mathcal{Q}} e^{t\omega_{\mathcal{Q}}}$. If*

$$\alpha_j = \frac{\|(\mathcal{L}\mathcal{Q})^{j+1}\mathcal{L}u_0\|}{\|(\mathcal{L}\mathcal{Q})^{j}\mathcal{L}u_0\|}, \quad 1 \leq j \leq n, \tag{3.40}$$

then for $1 \leq p \leq n$

$$\|w_0(t) - w_0^p(t)\| \leq M_5^p(t),$$

where

$$M_5^p(t) = C_1 \left(\prod_{j=1}^{p} \alpha_j\right) f_p(\omega_{\mathcal{Q}}, t) h(\omega - \omega_{\mathcal{Q}}, t_p),$$

$$f_p(\omega_{\mathcal{Q}}, t) = \int_0^t \frac{(t-\sigma)^{p-1}}{(p-1)!} e^{\sigma\omega_{\mathcal{Q}}} d\sigma, \qquad h(\omega - \omega_{\mathcal{Q}}, t_p) = \int_0^{t_p} e^{s(\omega-\omega_{\mathcal{Q}})} ds,$$

and $C_1 = MM_{\mathcal{Q}} \|\mathcal{P}\|^2 \|(\mathcal{L}\mathcal{Q})^{p+1}\mathcal{L}u_0\|$.

Proof. By following the same procedure as in the proof of the Theorem 3.1.4 we obtain

$$w_0(t) - w_0^p(t) = \int_0^t \int_0^{\tau_p} \cdots \int_0^{\tau_2} \int_0^{\min(\tau_1, t_p)} \mathcal{P} e^{s\mathcal{L}} \mathcal{P} e^{(\tau_1 - s)\mathcal{L}\mathcal{Q}} (\mathcal{L}\mathcal{Q})^{p+1} \mathcal{L} u_0 \, ds d\tau_1 \cdots d\tau_p.$$

By applying Cauchy's formula for repeated integration, this expression may be simplified to

$$w_0(t) - w_0^p(t) = \int_0^t \frac{(t-\sigma)^{p-1}}{(p-1)!} \int_0^{\min(\sigma, t_p)} \mathcal{P} e^{s\mathcal{L}} \mathcal{P} e^{(\sigma - s)\mathcal{L}\mathcal{Q}} (\mathcal{L}\mathcal{Q})^{p+1} \mathcal{L} u_0 \, ds d\sigma.$$

Thus,

$$\begin{aligned}
\|w_0(t) - w_0^p(t)\| &\leq \int_0^t \frac{(t-\sigma)^{p-1}}{(p-1)!} \int_0^{\min(\sigma, t_p)} \left\| \mathcal{P} e^{s\mathcal{L}} \mathcal{P} e^{(\sigma - s)\mathcal{L}\mathcal{Q}} (\mathcal{L}\mathcal{Q})^{p+1} \mathcal{L} u_0 \right\| ds d\sigma \\
&\leq MM_{\mathcal{Q}} \|\mathcal{P}\|^2 \|(\mathcal{L}\mathcal{Q})^{p+1}\mathcal{L}u_0\| \int_0^t \frac{(t-\sigma)^{p-1}}{(p-1)!} \int_0^{t_p} e^{s\omega} e^{(\sigma - s)\omega_{\mathcal{Q}}} ds d\sigma \\
&\leq C_1 \left(\prod_{j=1}^p \alpha_j \right) \left(\int_0^t \frac{(t-\sigma)^{p-1}}{(p-1)!} e^{\sigma\omega_{\mathcal{Q}}} d\sigma \right) \left(\int_0^{t_p} e^{s(\omega - \omega_{\mathcal{Q}})} ds \right) \\
&= C_1 \left(\prod_{j=1}^p \alpha_j \right) f_p(\omega_{\mathcal{Q}}, t) h(\omega - \omega_{\mathcal{Q}}, t_p) = M_5^p(t),
\end{aligned}$$

where $C_1 = MM_{\mathcal{Q}} \|\mathcal{P}\|^2 \|(\mathcal{L}\mathcal{Q})^{p+1}\mathcal{L}u_0\|$,

$$f_p(\omega_{\mathcal{Q}}, t) = \int_0^t \frac{(t-\sigma)^{p-1}}{(p-1)!} e^{\sigma\omega_{\mathcal{Q}}} d\sigma = \begin{cases} \dfrac{t^p}{p!} & \omega_{\mathcal{Q}} = 0 \\ \dfrac{1}{\omega_{\mathcal{Q}}^p} \left[e^{t\omega_{\mathcal{Q}}} - \displaystyle\sum_{k=0}^{p-1} \frac{(t\omega_{\mathcal{Q}})^k}{k!} \right] & \omega_{\mathcal{Q}} \neq 0 \end{cases} \tag{3.41}$$

and

$$h(\omega - \omega_{\mathcal{Q}}, t_p) := \int_0^{t_p} e^{s(\omega - \omega_{\mathcal{Q}})} ds = \begin{cases} t_p & \omega = \omega_{\mathcal{Q}} \\ \dfrac{e^{t_p(\omega-\omega_{\mathcal{Q}})} - 1}{\omega - \omega_{\mathcal{Q}}} & \omega \neq \omega_{\mathcal{Q}} \end{cases}$$

are both strictly increasing functions of t and t_p, respectively. □

Corollary 3.1.6.1. (Uniform convergence of the Type-II FMA) *If α_j in Theorem 3.1.6 satisfy*

$$\alpha_j < \frac{j}{T} \quad (\omega_{\mathcal{Q}} = 0) \qquad \text{or} \qquad \alpha_j < \omega_{\mathcal{Q}} \frac{e^{T\omega_{\mathcal{Q}}} - \sum_{k=0}^{j-2} \frac{(T\omega_{\mathcal{Q}})^k}{k!}}{e^{T\omega_{\mathcal{Q}}} - \sum_{k=0}^{j-1} \frac{(T\omega_{\mathcal{Q}})^k}{k!}} \quad (\omega_{\mathcal{Q}} \neq 0) \tag{3.42}$$

for $1 \leq j \leq n$, then for any arbitrarily small $\delta > 0$, there exists an ordered sequence $0 < t_0 < t_1 < \cdots < t_n \leq T$ such that

$$\|w_0(T) - w_0^p(T)\| \leq \delta, \quad 1 \leq p \leq n$$

and which satisfies

$$t_j \geq \begin{cases} \dfrac{j!\delta}{C_1 \left(\prod_{i=1}^{j} \alpha_i\right) T^j} & \omega_{\mathcal{Q}} = 0, \\ \dfrac{\omega_{\mathcal{Q}}^j \delta}{C_1 \left(\prod_{i=1}^{j} \alpha_i\right) \left[e^{T\omega_{\mathcal{Q}}} - \sum_{k=0}^{j-1} \frac{(T\omega_{\mathcal{Q}})^k}{k!}\right]} & \omega_{\mathcal{Q}} \neq 0, \end{cases}$$

when $\omega = \omega_Q$*, and*

$$t_j \geq \begin{cases} \dfrac{1}{\omega} \ln\left[1 + \dfrac{j!\omega\delta}{C_1\left(\prod_{i=1}^{j}\alpha_i\right)T^j}\right] & \omega_Q = 0, \\ \dfrac{1}{\omega - \omega_Q} \ln\left[1 + \dfrac{(\omega - \omega_Q)\omega_Q^j\delta}{C_1\left(\prod_{i=1}^{j}\alpha_i\right)\left[e^{T\omega_Q} - \sum_{k=0}^{j-1}\frac{(T\omega_Q)^k}{k!}\right]}\right] & \omega_Q \neq 0, \end{cases}$$

when $\omega \neq \omega_Q$*.*

Proof. We now consider separately the two cases where $\omega = \omega_Q$ and where $\omega \neq \omega_Q$. If $\omega = \omega_Q$, then

$$\begin{aligned} \|w_0(t) - w_0^p(t)\| &\leq C_1\left(\prod_{i=1}^{p}\alpha_i\right)\int_0^t\int_0^{\tau_p}\cdots\int_0^{\tau_2}\int_0^{t_p} e^{\tau_1\omega_Q}e^{s(\omega-\omega_Q)}dsd\tau_1\cdots d\tau_p \\ &= t_pC_1\left(\prod_{i=1}^{p}\alpha_i\right)\int_0^t\int_0^{\tau_p}\cdots\int_0^{\tau_2} e^{\tau_1\omega_Q}d\tau_1\cdots d\tau_p \\ &= t_pC_1\left(\prod_{i=1}^{p}\alpha_i\right)f_p(\omega_Q, t), \end{aligned}$$

where $f_p(\omega_Q, t)$ is defined in (3.41). To ensure that $\|w_0(t) - w_0^p(t)\| \leq \delta$ for all $0 \leq t \leq T$, we can take

$$t_pC_1\left(\prod_{i=1}^{p}\alpha_i\right)f_p(\omega_Q, T) = \max_{t\in[0,T]} t_pC_1\left(\prod_{i=1}^{p}\alpha_i\right)f_p(\omega_Q, t) \leq \delta,$$

so that

$$t_p \leq \frac{\delta}{C_1\left(\prod_{i=1}^{p}\alpha_i\right)f_p(\omega_Q, T)} = \begin{cases} \dfrac{p!\delta}{C_1\left(\prod_{i=1}^{p}\alpha_i\right)T^p} & \omega_Q = 0, \\ \dfrac{\omega_Q^p\delta}{C_1\left(\prod_{i=1}^{p}\alpha_i\right)\left[e^{T\omega_Q} - \sum_{k=0}^{p-1}\frac{(T\omega_Q)^k}{k!}\right]} & \omega_Q \neq 0. \end{cases}$$

On the other hand, if $\omega \neq \omega_{\mathcal{Q}}$ then

$$\begin{aligned}
\|w_0(t) - w_0^p(t)\| &\leq C_1 \left(\prod_{i=1}^{p} \alpha_i\right) \int_0^t \int_0^{\tau_p} \cdots \int_0^{\tau_2} \int_0^{t_p} e^{\tau_1 \omega_{\mathcal{Q}}} e^{s(\omega - \omega_{\mathcal{Q}})} ds d\tau_1 \cdots d\tau_p \\
&= \frac{e^{t_p(\omega-\omega_{\mathcal{Q}})} - 1}{\omega - \omega_{\mathcal{Q}}} C_1 \left(\prod_{i=1}^{p} \alpha_i\right) \int_0^t \int_0^{\tau_p} \cdots \int_0^{\tau_2} e^{\tau_1 \omega_{\mathcal{Q}}} d\tau_1 \cdots d\tau_p \\
&= \frac{e^{t_p(\omega-\omega_{\mathcal{Q}})} - 1}{\omega - \omega_{\mathcal{Q}}} C_1 \left(\prod_{i=1}^{p} \alpha_i\right) f_p(\omega_{\mathcal{Q}}, t).
\end{aligned}$$

To ensure that $\|w_0(t) - w_0^p(t)\| \leq \delta$ for all $0 \leq t \leq T$, we can take

$$\frac{e^{t_p(\omega-\omega_{\mathcal{Q}})} - 1}{\omega - \omega_{\mathcal{Q}}} C_1 \left(\prod_{i=1}^{p} \alpha_i\right) f_p(\omega_{\mathcal{Q}}, T) = \max_{t \in [0,T]} \frac{e^{t_p(\omega-\omega_{\mathcal{Q}})} - 1}{\omega - \omega_{\mathcal{Q}}} C_1 \left(\prod_{i=1}^{p} \alpha_i\right) f_p(\omega_{\mathcal{Q}}, t) \leq \delta.$$

Let us now consider the two cases $\omega > \omega_{\mathcal{Q}}$ and $\omega < \omega_{\mathcal{Q}}$ separately. When $\omega > \omega_{\mathcal{Q}}$, we have

$$e^{t_p(\omega-\omega_{\mathcal{Q}})} \leq 1 + \frac{(\omega - \omega_{\mathcal{Q}})\delta}{C_1 \left(\prod_{i=1}^{p} \alpha_i\right) f_p(\omega_{\mathcal{Q}}, T)},$$

and

$$t_p \leq \begin{cases} \dfrac{1}{\omega} \ln \left[1 + \dfrac{p!\omega\delta}{C_1 \left(\prod_{i=1}^{p} \alpha_i\right) T^p}\right] & \omega_{\mathcal{Q}} = 0, \\ \dfrac{1}{\omega - \omega_{\mathcal{Q}}} \ln \left[1 + \dfrac{(\omega - \omega_{\mathcal{Q}})\omega_{\mathcal{Q}}^p \delta}{C_1 \left(\prod_{i=1}^{p} \alpha_i\right) \left[e^{T\omega_{\mathcal{Q}}} - \sum_{k=0}^{p-1} \frac{(T\omega_{\mathcal{Q}})^k}{k!}\right]}\right] & \omega_{\mathcal{Q}} \neq 0. \end{cases}$$

On the other hand, when $\omega < \omega_{\mathcal{Q}}$, we have

$$\frac{1 - e^{-t_p(\omega_{\mathcal{Q}}-\omega)}}{\omega_{\mathcal{Q}} - \omega} C_1 \left(\prod_{i=1}^{p} \alpha_i\right) f_p(\omega_{\mathcal{Q}}, T) = \max_{t \in [0,T]} \frac{1 - e^{-t_p(\omega_{\mathcal{Q}}-\omega)}}{\omega_{\mathcal{Q}} - \omega} C_1 \left(\prod_{i=1}^{p} \alpha_i\right) f_p(\omega_{\mathcal{Q}}, t) \leq \delta,$$

so that

$$e^{-t_p(\omega_Q-\omega)} \geq 1 - \frac{(\omega_Q-\omega)\delta}{C_1\left(\prod_{i=1}^{p}\alpha_i\right) f_p(\omega_Q,T)}$$

i.e.,

$$-t_p(\omega_Q-\omega) \geq \ln\left[1 - \frac{(\omega_Q-\omega)\delta}{C_1\left(\prod_{i=1}^{p}\alpha_i\right) f_p(\omega_Q,T)}\right].$$

Hence,

$$t_p \leq \begin{cases} \dfrac{1}{\omega}\ln\left[1 - \dfrac{p!(-\omega)\delta}{C_1\left(\prod_{i=1}^{p}\alpha_i\right)T^p}\right] & \omega_Q = 0, \\ -\dfrac{1}{\omega_Q-\omega}\ln\left[1 - \dfrac{(\omega_Q-\omega)\omega_Q^p\delta}{C_1\left(\prod_{i=1}^{p}\alpha_i\right)\left[e^{T\omega_Q} - \sum_{k=0}^{p-1}\frac{(T\omega_Q)^k}{k!}\right]}\right] & \omega_Q \neq 0. \end{cases}$$

For all the four cases, if $\omega_Q = 0$ then we have condition $\alpha_p < p/T$, and the upper bound of the time sequence satisfies:

$$\frac{p!\delta}{C_1\left(\prod_{i=1}^{p}\alpha_i\right)T^p} < \frac{(p-1)!\delta}{C_1\left(\prod_{i=1}^{p-1}\alpha_i\right)T^{p-1}}, \qquad p \geq 2.$$

If $\omega_Q \neq 0$ then we have the condition

$$\alpha_p < \omega_Q\left(e^{T\omega_Q} - \sum_{k=0}^{p-2}\frac{(T\omega_Q)^k}{k!}\right)\left(e^{T\omega_Q} - \sum_{k=0}^{p-1}\frac{(T\omega_Q)^k}{k!}\right)^{-1}$$

and the upper bound of the time sequence satisfies

$$\frac{\omega_Q^{p-1}}{\left(e^{T\omega_Q} - \sum_{k=0}^{p-2}\frac{(T\omega_Q)^k}{k!}\right)\prod_{i=1}^{p-1}\alpha_i} < \frac{\omega_Q^{p}}{\left(e^{T\omega_Q} - \sum_{k=0}^{p-1}\frac{(T\omega_Q)^k}{k!}\right)\prod_{i=1}^{p}\alpha_i}, \qquad p \geq 2.$$

Therefore, there always exists a increasing time sequence $0 < t_1 < \cdots < t_n$ such that $\|w_0(t) - w_0^p(t)\| \leq \delta$ for all $0 \leq t \leq T$. And since we have proved that this δ-bound on the error holds for all t_n upper bounded as in the two cases above, there exists such an increasing time sequence $0 < t_1 < \cdots < t_n$ with t_n *lower*-bounded by the same quantities. Indeed, because of the coarseness of the approximations applied in the proof, there may exist such a time sequence with significantly larger t_i. □

Remark If we replace (3.42) with the stronger condition

$$\begin{cases} \alpha_j < \dfrac{j}{\epsilon T}, & \omega_{\mathcal{Q}} = 0 \\ \alpha_j < \dfrac{1}{\epsilon T}, & \omega_{\mathcal{Q}} \neq 0 \end{cases} \qquad 1 \leq j < \infty, \tag{3.43}$$

where ϵ is some arbitrary constant satisfying $\epsilon > 1$, then we have

$$\lim_{j\to+\infty} t_j \geq \lim_{j\to+\infty} \begin{cases} \dfrac{j!\delta}{C_1\left(\prod_{i=1}^{j}\alpha_i\right)T^j} \geq \dfrac{\delta}{C_1}\epsilon^j = +\infty & \omega_{\mathcal{Q}} = 0, \\ \dfrac{\omega_{\mathcal{Q}}^j\delta}{C_1\left(\prod_{i=1}^{j}\alpha_i\right)\left[e^{T\omega_{\mathcal{Q}}} - \sum_{k=0}^{j-1}\frac{(T\omega_{\mathcal{Q}})^k}{k!}\right]} = +\infty & \omega_{\mathcal{Q}} \neq 0. \end{cases}$$

for $\omega = \omega_{\mathcal{Q}}$ and

$$\lim_{j\to+\infty} t_j \geq \lim_{j\to+\infty} \begin{cases} \dfrac{j}{\omega}\ln\left[\dfrac{\delta}{C_1}\omega\epsilon\right] = +\infty & \omega_{\mathcal{Q}} = 0, \\ \dfrac{j}{\omega - \omega_{\mathcal{Q}}}\ln\left[\dfrac{(\omega - \omega_{\mathcal{Q}})T^j}{C_1 o(T^j)}\right] = +\infty & \omega_{\mathcal{Q}} \neq 0. \end{cases}$$

Hence, there exists a j such that the upper bound for t_j is greater than or equal to T. For such case, the Type II FMA degenerates to the truncation approximation (H-model), for which Corollary 3.1.4.2 grants us asymptotic convergence.

H_t-model

The H_t-model is obtained by solving the system (3.20) with $w_n^{e_n}(t)$ approximated using Chorin's t-model [18] (see equation (3.25)). Convergence analysis can be performed by using the mathematical methods we employed for the proofs of the H-model. Note that the classical t-model is equivalent to a zeroth-order H_t-model.

Theorem 3.1.7. (Accuracy of the H_t-model) *Let $e^{t\mathcal{L}}$ and $e^{t\mathcal{L}\mathcal{Q}}$ be strongly continuous semigroups with upper bounds $\|e^{t\mathcal{L}}\| \leq Me^{t\omega}$ and $\|e^{t\mathcal{L}\mathcal{Q}}\| \leq M_{\mathcal{Q}}e^{t\omega_{\mathcal{Q}}}$, and let $T > 0$ be a fixed integration time. For some fixed n, let*

$$\alpha_j = \frac{\|(\mathcal{L}\mathcal{Q})^{j+1}\mathcal{L}u_0\|}{\|(\mathcal{L}\mathcal{Q})^{j}\mathcal{L}u_0\|}, \quad 1 \leq j \leq n. \tag{3.44}$$

Then, for any $1 \leq p \leq n$ and all $t \in [0, T]$, we have

$$\|w_0(t) - w_0^p(t)\| \leq M_6^p(t) \leq M_6^p(T),$$

where

$$M_6^p(t) = C_4\left(\prod_j^p \alpha_j\right)\frac{t^{p+1}}{(p+1)!}, \quad C_4 = \left[C_1A_1A_2 + \frac{C_1}{M_{\mathcal{Q}}A_3}\right], \quad A_3 = \begin{cases} 1 & \omega \leq 0, \\ e^{T\omega} & \omega > 0 \end{cases},$$

and C_1, A_1, A_2 are the same as before.

Proof. For p-th order H_t-model, the difference between the memory term w_0 and its approximation w_0^p is

$$w_0(t) - w_0^p(t) = \int_0^t\int_0^{\tau_p}\cdots\int_0^{\tau_2}\int_0^{\tau_1} \mathcal{P}e^{s\mathcal{L}}\mathcal{P}e^{(\tau_1-s)\mathcal{L}\mathcal{Q}}(\mathcal{L}\mathcal{Q})^{p+1}\mathcal{L}u_0 ds \\ - \tau_1\mathcal{P}e^{\tau_1\mathcal{L}}\mathcal{P}(\mathcal{L}\mathcal{Q})^{p+1}\mathcal{L}u_0 d\tau_1\cdots d\tau_p. \tag{3.45}$$

Using Cauchy's formula for repeated integration, we can bound the norm of the second term in (3.45) as

$$\left\| \int_0^t \frac{(t-\sigma)^{p-1}}{(p-1)!} \sigma \mathcal{P} e^{\sigma \mathcal{L}} \mathcal{P} (\mathcal{L}\mathcal{Q})^{p+1} \mathcal{L} x_0 d\sigma \right\| \leq \int_0^t \frac{(t-\sigma)^{p-1}}{(p-1)!} \| \sigma \mathcal{P} e^{\sigma \mathcal{L}} \mathcal{P} (\mathcal{L}\mathcal{Q})^{p+1} \mathcal{L} x_0 \| d\sigma$$
$$\leq M \|\mathcal{P}\|^2 \|(\mathcal{L}\mathcal{Q})^{p+1} \mathcal{L} u_0\| \underbrace{\int_0^t \frac{(t-\sigma)^{p-1}}{(p-1)!} \sigma e^{\sigma\omega} d\sigma}_{g_p(t,\omega)}$$
$$= \frac{C_1}{M_{\mathcal{Q}}} \left(\prod_{j=1}^{p} \alpha_j \right) g_p(t,\omega), \tag{3.46}$$

where $C_1 = \|\mathcal{P}\|^2 \|\mathcal{L}\mathcal{Q}\mathcal{L} u_0\| M M_{\mathcal{Q}}$ as before. The function $g_p(t,\omega)$, may be bounded from above as

$$g_p(t,\omega) \leq A_3 \int_0^t \frac{(t-\sigma)^{p-1}}{(p-1)!} \sigma d\sigma = A_3 \frac{t^{p+1}}{(p+1)!}, \quad A_3 = \max_{s\in[0,T]} e^{s\omega} = \begin{cases} 1 & \omega \leq 0 \\ e^{T\omega} & \omega > 0 \end{cases}.$$

By applying the triangle inequality to (3.45), and taking (4.16) into account, we obtain

$$\|w_0(t) - w_0^p(t)\| \leq C_1 A_1 A_2 \left(\prod_{j=1}^{p} \alpha_j \right) \frac{t^{p+1}}{(p+1)!} + \frac{C_1}{M_{\mathcal{Q}}} A_3 \left(\prod_{j=1}^{p} \alpha_j \right) \frac{t^{p+1}}{(p+1)!} = M_6^p(t).$$

□

One can see that the upper bounds $M_6^p(t)$ and $M_3^p(t)$ (see Theorem 3.1.4) share the same structure, the only difference being the constant out front. Hence by changing C_2 to C_4, we can prove of a series of corollaries similar to 3.1.4.1, 3.1.4.2, and 3.1.4.3. In summary, what holds for the H-model also holds for the H_t-model. For the sake of brevity, we omit the statement and proofs of those corollaries.

3.1.5 Some special dynamical Systems

Linear systems The upper bounds we obtained above are not easily computable for general nonlinear systems and infinite-rank projections, e.g., Chorin's projection (2.7). However, if the dynamical system is linear, then such upper bounds are explicitly computable and convergence of the H-model can be established for linear phase space functions in any finite integration time T. To this end, consider the linear system $\dot{x} = Ax$ with random initial condition $x(0)$ sampled from the joint probability density function

$$\rho_0(x_0) = \delta(x_{01} - x_1(0)) \prod_{j=2}^{N} \rho_{0j}(x_{0j}). \tag{3.47}$$

In other words, the initial condition for the quantity of interest $u(x) = x_1(t)$ is set to be deterministic, while all other variables $x_2, \ldots, x_N$ are zero-mean and statistically independent at $t = 0$[4]. Here we assume for simplicity that ρ_{0j} $(j = 2, .., N)$ are i.i.d. standard normal distributions. Observe that the Liouville operator associated with the linear system $\dot{x} = Ax$ is

$$\mathcal{L} = \sum_{i=1}^{N} \sum_{j=1}^{N} A_{ij} x_j \frac{\partial}{\partial x_i}, \tag{3.48}$$

where A_{ij} are the entries of the matrix A. If we choose observable $u = x_1(t)$, then Chorin's projection operator (2.9) yields the evolution equation for the conditional expectation $\mathbb{E}[x_1(t)|x_1(0)]$, i.e., the *conditional mean path* (2.11), which can be

[4]These choices for ρ_0 are merely for convenience in demonstrating important features. With a more general choice of ρ_0, it is convenient to represent $\mathcal{L}$, $\mathcal{P}$, and $\mathcal{Q}$ in terms of an orthonormal basis for V with respect to the ρ_0 inner product. Then, e.g., operator norms within the invariant subspace reduce to matrix norms of the associated matrix.

explicitly written as

$$\frac{d}{dt}\mathbb{E}[x_1|x_1(0)] = A_{11}\mathbb{E}[x_1|x_1(0)] + w_0(t), \tag{3.49}$$

where $A_{11} = \mathcal{P}\mathcal{L}x_1(0)$ is the first entry of the matrix A, w_0 represents the memory integral (3.15). Next, we explicitly compute the upper bounds for the memory growth and the error in the H-model for this system. To this end, we first notice that the domain of the Liouville operator can be restricted to the linear space

$$V = \text{span}\{x_1, \ldots, x_N\}. \tag{3.50}$$

In fact, V is invariant under $\mathcal{L}$, $\mathcal{P}$ and $\mathcal{Q}$, i.e., $\mathcal{L}V \subseteq V$, $\mathcal{P}V \subseteq V$ and $\mathcal{Q}V \subseteq V$. These operators have the following matrix representations

$$\mathcal{L} \simeq A^T \simeq \begin{bmatrix} a_{11} & b^T \\ a & M_{11}^T \end{bmatrix}, \qquad \mathcal{P} \simeq \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix}, \qquad \mathcal{Q} \simeq \begin{bmatrix} 0 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix},$$

where M_{11} is the minor of the matrix of A obtained by removing the first column and the first row, while

$$a = [A_{12} \cdots A_{1N}]^T, \qquad b^T = [A_{21} \cdots A_{N1}]. \tag{3.51}$$

Therefore,

$$\mathcal{L}\mathcal{Q} \simeq \begin{bmatrix} 0 & b^T \\ 0 & M_{11}^T \end{bmatrix}, \qquad \mathcal{L}(\mathcal{Q}\mathcal{L})^n x_1(0) \simeq \begin{bmatrix} b^T \left(M_{11}^T\right)^{n-1} a \\ \left(M_{11}^T\right)^n a \end{bmatrix}. \tag{3.52}$$

At this point, we set $x_{01} = x_1(0)$ and

$$q(t, x_{01}, \tilde{x}_0) = \int_0^t e^{s\mathcal{L}} \mathcal{P}\mathcal{L} e^{(t-s)\mathcal{Q}\mathcal{L}} \mathcal{Q}\mathcal{L} x_{01} ds.$$

Since $\tilde{x}_0 = (x_2(0), ..., x_N(0))$ is random, $q(t, x_{01}, \tilde{x}_0)$ is a random variable. By using Jensen's inequality $[\mathbb{E}(X)]^2 \leq \mathbb{E}[X^2]$, we have the following L^∞ estimate

$$\|(\mathcal{P}q)(t, x_{01})\|_{L^\infty} \leq \|q(t, x_{01}, \cdot)\|_{L^2_{\rho_0}}. \tag{3.53}$$

On the other hand, we have

$$\|e^{t\mathcal{L}}\|_{L^2_{\rho_0}(V)} \leq \|e^{t\mathcal{L}}\|_{L^2_{\rho_0}} \leq e^{t\omega}, \quad \omega = -\frac{1}{2} \inf \operatorname{div}_{\rho_0} F. \tag{3.54}$$

For linear dynamical systems, both $\|\cdot\|_{L^2_{\rho_0}(V)}$ and $\|\cdot\|_{L^2_{\rho_0}}$ upper bounds can be used to estimate the norm of the semigroup $e^{t\mathcal{L}}$. However, for the semigroup $e^{t\mathcal{L}\mathcal{Q}}$, we can only obtain the explicit form of the $\|\cdot\|_{L^2_{\rho_0}(V)}$ bound, which is given by the following perturbation theorem [28] (see also appendix 3.A):

$$\|e^{t\mathcal{L}\mathcal{Q}}\|_{L^2_{\rho_0}(V)} \leq e^{t\omega_{\mathcal{Q}}} \tag{3.55}$$

where

$$\omega_{\mathcal{Q}} = \omega + \sqrt{A_{11}^2 + \sum_{i=2}^{N} A_{1i}^2 \frac{\langle x_i^2(0)\rangle_{\rho_0}}{x_1^2(0)}} \geq \omega + \|\mathcal{L}\mathcal{P}\|_{L^2_{\rho_0}(V)}. \tag{3.56}$$

Memory growth It is straightforward at this point to compute the upper bound of the memory growth we obtained in Theorem 3.1.1. Since $\|\mathcal{P}\|_{L^2_{\rho_0}} = \|\mathcal{Q}\|_{L^2_{\rho_0}} = 1$ ($\mathcal{P}$ and $\mathcal{Q}$ are orthogonal projections relative to ρ_0), we have the following result

$$|w_0(t)| \leq \|\mathcal{L}\mathcal{Q}\mathcal{L}x_1(0)\|\frac{e^{t\omega}-e^{t\omega_{\mathcal{Q}}}}{\omega-\omega_{\mathcal{Q}}} = \sqrt{(b^Ta)^2x_1^2(0)+\left\|\Lambda_{x_{i+1}(0)}M_{11}^Ta\right\|_2^2}\frac{e^{t\omega}-e^{t\omega_{\mathcal{Q}}}}{\omega-\omega_{\mathcal{Q}}}, \tag{3.57}$$

where $\Lambda_{x_{i+1}(0)}$ is a $N-1\times N-1$ diagonal matrix with $\Lambda_{ii}=\langle x_{i+1}(0)\rangle_{\rho_0}$, and $\|\cdot\|_2$ is the vector 2-norm.

Accuracy of the H-model We are interested in computing the upper bound of the approximation error generated by the H-model (see section 3.1.4 - Theorem 3.1.4). By using the matrix representation of $\mathcal{L}$, $\mathcal{P}$ and $\mathcal{Q}$, the n-th order H-model MZ equation (3.49) for linear system can be explicitly written as

$$\begin{cases} \dfrac{d}{dt}\mathbb{E}[x_1|x_1(0)] = A_{11}\mathbb{E}[x_1|x_1(0)]+w_0^n(t) \qquad \text{(MZ equation)},\\ \dfrac{dw_j^n(t)}{dt} = b^T(M_{11}^T)^ja^T\mathbb{E}[x_1|x_1(0)]+w_{j+1}^n(t), \qquad j=0,1,\dots,n-1,\\ \dfrac{dw_n^n(t)}{dt} = b^T(M_{11}^T)^na^T\mathbb{E}[x_1|x_1(0)], \end{cases} \tag{3.58}$$

where M_{11}, a and b are defined as before (see equation (3.51)). The upper bound for the memory term approximation error is explicitly obtained as[5]

$$\begin{aligned} |w_0(t)-w_0^n(t)| &\leq A_1A_2\|\mathcal{L}(\mathcal{Q}\mathcal{L})^nx_1(0)\|\frac{t^{n+1}}{(n+1)!}\\ &= A_1A_2\sqrt{\left[b^T\left(M_{11}^T\right)^na\right]^2x_1^2(0)+\left\|\Lambda_{x_{i+1}(0)}\left(M_{11}^T\right)^{n+1}a\right\|_2^2}\frac{t^{n+1}}{(n+1)!} \end{aligned} \tag{3.59}$$

[5]The error bound for $|w_0(t)-w_0^n(t)|$ used here is slightly different from the one we obtained in Theorem 3.1.4. Instead of bounding the quotient $\alpha_n=|\mathcal{L}(\mathcal{Q}\mathcal{L})^{n+1}u_0\|/\|\mathcal{L}(\mathcal{Q}\mathcal{L})^nu_0\|$, here we choose to bound $\|\mathcal{L}(\mathcal{Q}\mathcal{L})^nu_0\|$ directly, which yields the estimate (3.59).

where A_1, A_2 are defined in (3.27), while ω and $\omega_{\mathcal{Q}}$ are given in (3.54) and (3.55), respectively. Note that for each fixed integration time T, the upper bound (3.59) goes to zero as we send n to infinity, i.e.,

$$\lim_{n\to+\infty} |w_0(T) - w_0^n(T)| = 0.$$

This means that the H-model converges for all linear dynamical systems with observables in the linear space (3.50).

Hamiltonian Systems with Mori's projection The semigroup estimates we obtained in section 3.1.1 allow us to compute explicitly an *a priori* estimate of the memory kernel in the Mori-Zwanzig equation if we employ *finite-rank* projections. In this section we outline the procedure to obtain such estimate for Hamiltonian dynamical systems. We begin by recalling that, in general, Hamiltonian systems are necessarily divergence-free, i.e.,

$$\text{div}_{\rho_{eq}}(F) = 0. \tag{3.60}$$

Here, $F(x)$ is the velocity field at the right hand side of (2.1), while $\rho_{eq} = e^{-\beta\mathcal{H}}/Z$ is the canonical Gibbs distribution[6]. By combining the estimate (3.7) with the divergence-free constraint (3.60), we find that the Koopman semigroup of a Hamil-

[6]Equation (3.60) is obtained by noticing that

$$\begin{aligned}\text{div}_{\rho_{eq}}(F) =& e^{\beta\mathcal{H}}\nabla\cdot\left(e^{-\beta\mathcal{H}}F\right),\\ =& e^{\beta\mathcal{H}}\sum_{i=1}^{N}\left(\frac{\partial}{\partial q_i}\left[e^{-\beta\mathcal{H}}\frac{\partial\mathcal{H}}{\partial p_i}\right]-\frac{\partial}{\partial p_i}\left[e^{-\beta\mathcal{H}}\frac{\partial\mathcal{H}}{\partial q_i}\right]\right),\\ =&0.\end{aligned}$$

tonian dynamical system is a *contraction* in the L^2_{eq} norm, i.e.,

$$\left\| e^{t\mathcal{L}} \right\|_{L^2_{eq}} \leq 1. \tag{3.61}$$

We recall that the GLE for the time autocorrelation function $C_{ij}(t)$ is given by

$$\frac{dC_{ij}}{dt} = \sum_{k=1}^{M} \Omega_{ik} C_{kj} - \sum_{k=1}^{M} \int_0^t K_{ik}(t-s) C_{kj}(s) ds. \tag{3.62}$$

If we employ a one-dimensional Mori's basis, i.e., $M = 1$, then we obtain the simplified equation

$$\frac{dC(t)}{dt} = \Omega_1 C(t) - \int_0^t K(t-s) C(s) ds. \tag{3.63}$$

where $C(t) = \langle u_1(0), u_1(t) \rangle_{eq}$. The main difficulty in solving the GLE (3.62) (or (3.63)) lies in computing the memory kernels $K_{ij}(t)$ (or $K(t)$ in the case of (3.63)). Hereafter we prove that such memory kernels can be uniformly bounded by a computable quantity that *depends only on the initial condition.* For the sake of simplicity, we will focus on the one-dimensional GLE (3.63).

Theorem 3.1.8. *The memory kernel $K(t)$ in the one-dimensional GLE* (3.62) *is uniformly bounded by $\|\dot{u}_1(0)\|^2_{\rho_{eq}} / \|u_1(0)\|^2_{\rho_{eq}}$, i.e.*

$$|K(t)| \leq \frac{\|\dot{u}_1(0)\|^2_{L^2_{\rho_{eq}}}}{\|u_1(0)\|^2_{L^2_{eq}}} \qquad \forall t \geq 0. \tag{3.64}$$

Proof. From the second-fluctuation dissipation theorem, we get that the memory kernel $K(t)$ satisfies

$$
\begin{aligned}
|K(t)| &= \left|\frac{\langle e^{t\mathcal{QL}}\mathcal{QL}u_1(0), \mathcal{QL}u_1(0)\rangle_{eq}}{\langle u_1(0), u_1(0)\rangle_{eq}}\right| \\
&\leq \|e^{t\mathcal{QL}}\mathcal{Q}\|_{L^2_{\rho_{eq}}} \frac{\|\mathcal{L}u_1(0)\|^2_{L^2_{\rho_{eq}}}}{\|u_1(0)\|^2_{L^2_{eq}}} = \|e^{t\mathcal{QL}}\mathcal{Q}\|_{L^2_{\rho_{eq}}} \frac{\|\dot{u}_1(0)\|^2_{L^2_{\rho_{eq}}}}{\|u_1(0)\|^2_{L^2_{eq}}}
\end{aligned}
$$

On the other hand, by using the numerical abscissa (3.5) and formula (3.8), we see that the semigroup $e^{t\mathcal{QLQ}}$ is contractive, i.e. $\|e^{t\mathcal{QLQ}}\|_{L^2_{\rho_{eq}}} \leq 1$. Since $\mathcal{Q}$ is an orthogonal projection with respect to ρ_{eq}, we have $\|e^{t\mathcal{QL}}\mathcal{Q}\|_{L^2_{\rho_{eq}}} = \|\mathcal{Q}e^{t\mathcal{QLQ}}\|_{L^2_{\rho_{eq}}} \leq \|\mathcal{Q}\|_{L^2_{\rho_{eq}}}\|e^{t\mathcal{QLQ}}\|_{L^2_{\rho_{eq}}} \leq 1$. This yields

$$
|K(t)| \leq \|e^{t\mathcal{QL}}\mathcal{Q}\|_{L^2_{\rho_{eq}}} \frac{\|\dot{u}_1(0)\|^2_{L^2_{\rho_{eq}}}}{\|u_1(0)\|^2_{L^2_{eq}}} \leq \frac{\|\dot{u}_1(0)\|^2_{L^2_{\rho_{eq}}}}{\|u_1(0)\|^2_{L^2_{eq}}}.
$$

□

Theorem 3.1.8 provides an *a priori* and easily computable upper bound for the memory kernel defining the dynamics of any quantity of interest u_1 that is initially in the Gibbs canonical ensemble $\rho_{eq} = e^{-\beta\mathcal{H}}/Z$. In section 3.2, we will calculate the upper bound (3.64) analytically and compare it with the exact memory kernel we obtain in prototype linear and nonlinear Hamiltonian systems.

Remark. We emphasized in section 3.1.1 that the semigroup estimate for $e^{t\mathcal{QLQ}}$ is not necessarily tight. In the context of high-dimensional Hamiltonian systems (e.g., molecular dynamics) it is often empirically assumed that the semigroup $e^{t\mathcal{QL}}\mathcal{Q}$ is dissipative, i.e. $\|e^{t\mathcal{QLQ}}\| \leq e^{t\omega_0}$, where $\omega_0 < 0$. In this case, the memory kernel turns out to be uniformly bounded by an exponentially decaying function since

$$|K(t)| \leq \|e^{t\mathcal{QL}}\mathcal{Q}\|_{L^2_{\rho_{eq}}} \frac{\|\mathcal{L}u_1\|^2_{L^2_{\rho_{eq}}}}{\|u_1\|^2_{L^2_{eq}}} \leq e^{t\omega_0} \frac{\|\dot{u}_1\|^2_{L^2_{\rho_{eq}}}}{\|u_1\|^2_{L^2_{eq}}}.$$

3.2 Numerical Examples

In this section, we provide simple numerical examples of the MZ memory approximation methods we discussed throughout the paper. Specifically, we study Hamiltonian systems (linear and nonlinear) with finite-rank projections (Mori's projection), and non-Hamiltonian systems with infinite-rank projections (Chorin's projection). In both cases we demonstrate the accuracy of the a priori memory estimation method we developed in 3.1.5. We also compute the solution to the MZ equation for non-Hamiltonian systems with the t-model, the H-model and the H_t-model.

3.2.1 Hamiltonian Dynamical Systems with Finite-Rank Projections

In this section we consider dimension reduction in linear and nonlinear Hamiltonian dynamical systems with finite-rank projection. In particular, we consider the Mori projection and study the MZ equation for the temporal auto-correlation function of a scalar quantity of interest.

Harmonic Chains of Oscillators

Consider a one-dimensional chain of harmonic oscillators. This is a simple but illustrative example of a linear Hamiltonian dynamical system which has been widely studied in statistical mechanics, mostly in relation with the microscopic theory of Brownian motio [33]. The Hamiltonian of the system can be written as

$$\mathcal{H}(p,q) = \frac{1}{2m}\sum_{i=1}^{N} p_i^2 + \frac{k}{2}\sum_{\substack{i,j=0\\ i<j}}^{N+1} (q_i - q_j)^2, \tag{3.65}$$

where q_i and p_i are, respectively, the displacement and momentum of the i-th particle, m is the mass of the particles (assumed constant throughout the network), and k is the elasticity constant that modulates the intensity of the quadratic interactions. We set fixed boundary conditions at the endpoints of the chain, i.e., $q_0(t) = q_{N+1}(t) = 0$ and $p_0(t) = p_{N+1}(t) = 0$ (particles are numbered from left to right) and $m = k = 1$. The Hamilton's equations are

$$\frac{dq_i}{dt} = \frac{\partial \mathcal{H}}{\partial p_i}, \qquad \frac{dp_i}{dt} = -\frac{\partial \mathcal{H}}{\partial q_i}, \tag{3.66}$$

which can be written in a matrix-vector form as

$$\begin{bmatrix} \dot{p} \\ \dot{q} \end{bmatrix} = \begin{bmatrix} 0 & kB - kD \\ I/m & 0 \end{bmatrix} \begin{bmatrix} p \\ q \end{bmatrix} \tag{3.67}$$

where B is the adjacency matrix of the chain and D is the degree matrix (see [8]). Note that (5.86) is a linear dynamical system. We are interested in the velocity auto-correlation function of a tagged oscillator, say the one at location $j = 1$. Such auto-correlation function is defined as

$$C_{p_1}(t) = \frac{\langle p_1(0)p_1(t)\rangle_{eq}}{\langle p_1(0)p_1(0)\rangle_{eq}}, \tag{3.68}$$

where the average is with respect to the Gibbs canonical distribution $\rho_{eq} = e^{-\beta\mathcal{H}}/Z$. It was shown in [33] that $C_{p_1}(t)$ can be obtained analytically by employing Lee's continued fraction method . The result is the well-known $J_0 - J_4$

solution

$$C_{p_1}(t) = J_0(2t) - J_4(2t), \tag{3.69}$$

where $J_i(t)$ is the i-th Bessel function of the first kind. On the other hand, the Mori-Zwanzig equation derived by the following Mori's projection

$$\mathcal{P}(\cdot) = \frac{\langle (\cdot), p_1(0)\rangle_{eq}}{\langle p_1(0), p_1(0)\rangle_{eq}} p_1(0) \tag{3.70}$$

yields the following GLE for $C_{p_1}(t)$

$$\frac{dC_{p_1}(t)}{dt} = \Omega_{p_1} C_{p_1}(t) - \int_0^t K(s) C_{p_1}(t-s) ds. \tag{3.71}$$

Here,

$$\Omega_{p_1} = \frac{\langle \mathcal{L} p_1(0), p_1(0)\rangle_{eq}}{\langle p_1(0), p_1(0)\rangle_{eq}} = 0$$

since $\langle p_i(0), q_j(0)\rangle_{eq} = 0$, while $K(t)$ is the MZ memory kernel. For the $J_0 - J_4$ solution, it is possible to derive the memory kernel $K(t)$ analytically. To this end, we simply insert (3.69) into (3.71) and apply the Laplace transform

$$\mathscr{L}[\cdot](s) = \int_0^\infty (\cdot) e^{-st} dt$$

to obtain

$$\hat{K}(s) = -s + \frac{1}{\hat{C}(s)}, \tag{3.72}$$

where $\hat{C}(s) = \mathscr{L}[C_{p_1}(t)]$ and $\hat{K}(s) = \mathscr{L}[K(t)]$. The inverse Laplace transform of (3.72) can be computed analytically as

$$K(t) = \frac{J_1(2t)}{t} + 1. \tag{3.73}$$

With $K(t)$ available, we can verify the memory estimated we derived in Theorem 3.1.8. To this end,

$$|K(t)| \leq \frac{\|\dot{p}_1(0)\|^2_{L^2_{\rho_{eq}}}}{\|p_1(0)\|^2_{L^2_{\rho_{eq}}}} = \frac{\|q_2(0) - 2q_1(0)\|^2_{L^2_{\rho_{eq}}}}{\|p_1(0)\|^2_{L^2_{\rho_{eq}}}} = 2. \tag{3.74}$$

Here we used the exact solution of the velocity auto-correlation function and displacement auto-correlation function of the fixed-end harmonic chain given by (see [33])

$$\begin{aligned}
\langle p_i(0), p_j(0)\rangle_{eq} &= \frac{k_B T}{\pi}\int_0^\pi \sin(ix)\sin(jx)dx \\
\langle q_i(0), q_j(0)\rangle_{eq} &= \frac{k_B T}{\pi}\int_0^\pi \frac{\sin(ix)\sin(jx)}{4\sin^2(x/2)}dx.
\end{aligned}$$

In Figure 5.5 we plot the absolute value of the memory kernel $K(t)$ together with the theoretical bound (3.74). It is seen that the upper bound we obtain in this case is of the same order of magnitude as the memory kernel.

Hald System

In this section, we study the Hamiltonian system studied by Chorin *et. al.* in [16, 18]. The Hamiltonian function is defined as

$$\mathcal{H}(p, q) = \frac{1}{2}(q_1^2 + p_1^2 + q_2^2 + p_2^2 + q_1^2 q_2^2), \tag{3.75}$$

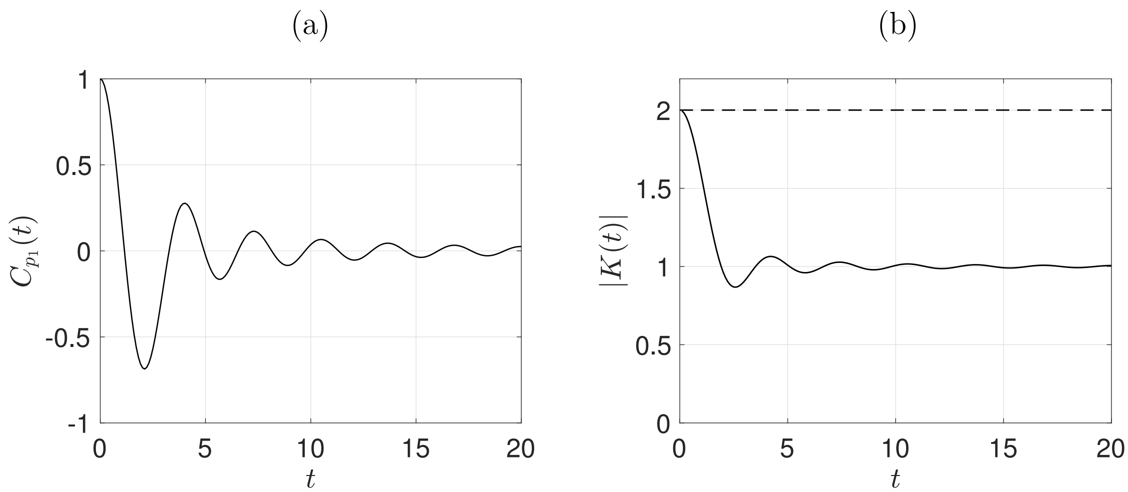

Figure 3.1: Harmonic chain of oscillators. (a) Velocity auto-correlation function $C_{p_1}(t)$ and (b) memory kernel $K(t)$ of the corresponding MZ equation. It is seen that our theoretical estimate (3.74) (dashed line) correctly bounds the MZ memory kernel. Note that the upper bound we obtain is of the same order of magnitude as the memory kernel.

while the corresponding Hamilton's equations of motion are

$$\begin{cases} \dot{q}_1 &= p_1 \\ \dot{p}_1 &= -q_1(1+q_2^2) \\ \dot{q}_2 &= p_2 \\ \dot{p}_2 &= -q_2(1+q_1^2) \end{cases} \tag{3.76}$$

We assume that the initial state is distributed according to canonical Gibbs distribution $\rho_{eq} = e^{-\mathcal{H}(p,q)}/Z$. The partition function Z is given by

$$Z = e^{1/4}(2\pi)^{3/2} K_0\left(\frac{1}{4}\right), \tag{3.77}$$

where $K_0(t)$ is the modified Bessel function of the second kind. We aim to study the properties of the autocorrelation function of the first component q_1, which is defined as

$$C_{q_1}(t) = \frac{\langle q_1(0), q_1(t)\rangle_{eq}}{\langle q_1(0), q_1(0)\rangle_{eq}}.$$

Obviously, $C_{q_1}(0) = 1$. The evolution equation for $C_{q_1}(t)$ is obtained by using the MZ formulation with the Mori's projection

$$\mathcal{P}(\cdot) = \frac{\langle (\cdot), q_1(0)\rangle_{eq}}{\langle q_1(0), q_1(0)\rangle_{eq}} q_1(0). \tag{3.78}$$

This yields the GLE

$$\frac{dC_{q_1}(t)}{dt} = \Omega_{q_1} C_{q_1}(t) - \int_0^t K(s) C_{q_1}(t-s) ds. \tag{3.79}$$

The streaming term $\Omega_{q_1} C_{q_1}(t)$ is again identically zero since

$$\Omega_{q_1} = \frac{\langle \mathcal{L} q_1(0), q_1(0)\rangle_{eq}}{\langle q_1(0), q_1(0)\rangle_{eq}} = 0.$$

Theorem 3.1.8 provides the following computable upper bound for the modulus of $K(t)$

$$|K(t)| \leq \frac{\|\dot{q}_1(0)\|^2_{L^2_{\rho_{eq}}}}{\|q_1(0)\|^2_{L^2_{\rho_{eq}}}} = \frac{\|p_1(0)\|^2_{L^2_{\rho_{eq}}}}{\|q_1(0)\|^2_{L^2_{\rho_{eq}}}} = \frac{e^{1/4} K_0\left(1/4\right)}{\sqrt{\pi} U\left(1/2, 0, 1/2\right)} \approx 1.39786, \tag{3.80}$$

where $U(a, b, y)$ is the confluent hypergeometric function of the second kind. In Figure 3.2 we plot the correlation function $C_{q_1}(t)$ we obtain numerically with Markov Chain Monte Carlo (MCMC), together with the memory kernel $K(t)$ we

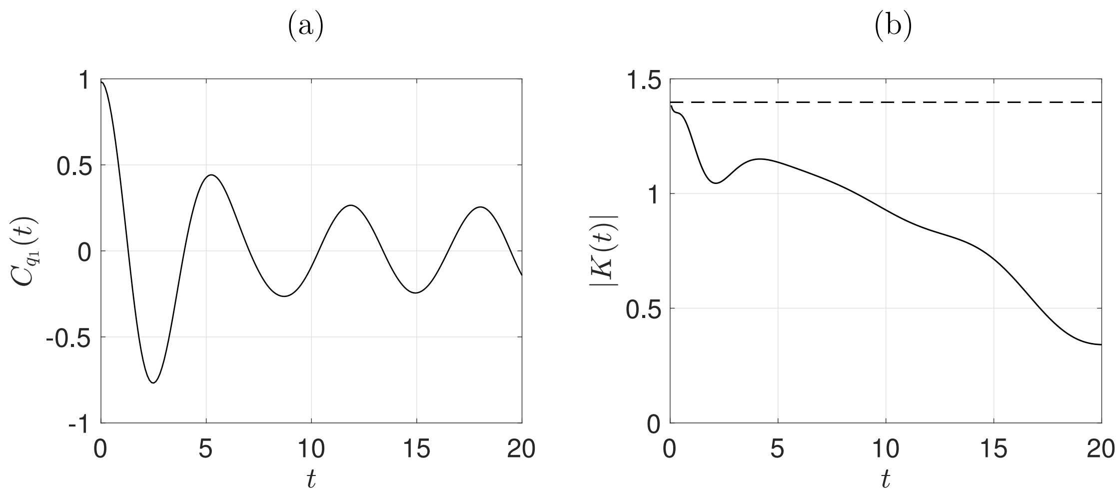

Figure 3.2: Hald Hamiltonian system (3.76). (a) Autocorrelation function of the displacement $q_1(t)$ and (b) memory kernel of the governing MZ equation. Here $C_{q_1}(t)$ is computed by Markov chain Monte-Carlo (MCMC) while $K(t)$ is determined by inverting numerically the Laplace transform in (3.81) with the Talbot algorithm. It is seen that the theoretical upper bound (3.80) (dashed line) is of the same order of magnitude as the memory kernel.

computed numerically based on $C_{q_1}(t)$[7].

3.2.2 Non-Hamiltonian Systems with Infinite-Rank Projections

In this section we study the accuracy of the t-model, the H-model and the H_t model in predicting scalar quantities of interest in non-Hamiltonian systems. In particular, we consider the MZ formulation with Chorin's projection operator. For

[7] The memory kernel $K(t)$ plotted in Figure 3.2(b) is computed by inverting numerically the Laplace transform of (3.79), i.e.,

$$K(t) = \mathscr{L}^{-1}\left[-s + \frac{1}{\hat{C}(s)}\right], \tag{3.81}$$

where $\hat{C}(s) = \mathscr{L}[C_{q_1}(t)]$. In practice, we replaced the numerical solution $C_{q_1}(t)$ within the time interval $[0, 20]$ with a high-order polynomial interpolant at Gauss-Chebyshev-Lobatto nodes (in $[0, 20]$), computed $\hat{C}(s)$ analytically (Laplace transform of a polynomial), and then computed the inverse Laplace transform (3.81) numerically with the Talbot algorithm [?].

the particular case of linear dynamical systems we also compute the theoretical upper bounds we obtained in Section 3.1.5 for the memory growth and the error in the H-model, and compare such bounds with exact results.

Linear Dynamical Systems

We begin by considering a low-dimensional linear dynamical system $\dot{x} = Ax$ evolving from a random initial state with density $\rho_0(x)$ to verify the MZ memory estimates we obtained in Section 3.1.5. For simplicity, we choose A to be negative definite

$$A = e^C B e^{-C}, \qquad B = \begin{bmatrix} -\frac{1}{8} & 0 & 0 \\ 0 & -\frac{2}{3} & 0 \\ 0 & 0 & -\frac{1}{2} \end{bmatrix}, \qquad C = \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 1 \\ 0 & -1 & 0 \end{bmatrix}. \tag{3.82}$$

In this case, the origin of the phase space is a stable node and it is easy to estimate $\left\|e^{t\mathcal{L}}\right\|_{\rho_0}$.[8] We set $x_1(0) = 1$ and $x_2(0), x_3(0)$ independent standard normal random variables. In this setting, the semigroup estimates (3.54) and (3.55) are explicit

$$\|e^{t\mathcal{L}}\| \le e^{t\omega}, \quad \omega = -\frac{1}{2}\overrightarrow{\mathcal{T}}(A) = 0.6458,$$
$$\|e^{t\mathcal{L}\mathcal{Q}}\| \le e^{t\omega_{\mathcal{Q}}}, \quad \omega_{\mathcal{Q}} = \omega + \sqrt{A_{11}^2 + \sum_{i=2}^{N} \frac{A_{1i}^2}{x_1^2(0)}} = 1.1621.$$

[8] For general matrices A, it is more difficult to estimate $\|e^{t\mathcal{L}}\|_{L^2_{\rho_0}}$. However, since $\mathcal{L}$ is a bounded linear operator in the subspace V where the quantity of interest lives, we can use the norm $\|e^{t\mathcal{L}}\|_{L^2_{\rho_0}(V)}$, which is explicitly computable.

Therefore, we obtain the following explicit upper bounds for the memory integral and the error of the H-model (see equations (3.57) and (3.59))

$$|w_0(t)| \leq 0.1964 \left(e^{1.1621t} - e^{0.6458t}\right), \tag{3.83}$$

$$|w_0(t) - w_0^n(t)| \leq e^{1.1624t} \sqrt{\left(b^T \left(M_{11}^T\right)^n a_1\right)^2 x_1^2(0) + \left\|\left(M_{11}^T\right)^{n+1} a\right\|_2^2} \frac{t^{n+1}}{(n+1)!}. \tag{3.84}$$

Next, we compare these error bounds with numerical results obtained by solving numerically the H-model (3.58). For example, the second-order H-model reads

$$\begin{cases} \dfrac{d}{dt}\mathbb{E}[x_1(t)|x_1(0)] = -0.4560\mathbb{E}[x_1(t)|x_1(0)] + w_0^2(t), \\ \dfrac{dw_0^2(t)}{dt} = 0.0586\mathbb{E}[x_1(t)|x_1(0)] + w_1^2(t), \\ \dfrac{dw_1^2(t)}{dt} = -0.0192\mathbb{E}[x_1(t)|x_1(0)]. \end{cases} \tag{3.85}$$

In Figure 3.3 we demonstrate convergence of the H-model to the benchmark solution computed by Monte-Carlo simulation as we increase the H-model differentiation order. In Figure 3.4 we plot the bound on the memory growth (equation (3.83)) and the bound in the memory error (equation (3.84)) together with exact results.

Remark The results we just obtained can be obviously extended to higher-dimensional linear dynamical systems. In Figure 3.5 we plot the benchmark conditional mean path we obtained through Monte Carlo simulation together with the solution of the H-model (3.58) for the 100-dimensional linear dynamical system

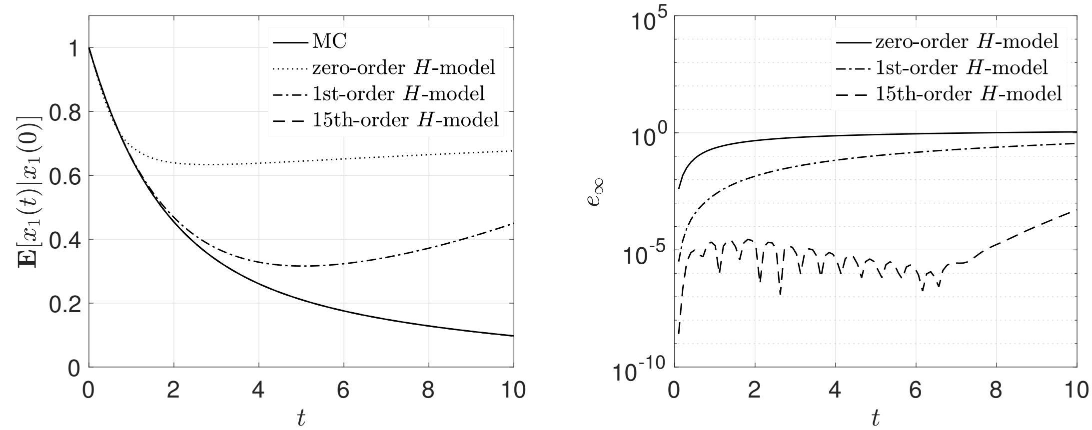

Figure 3.3: Convergence of the H-model for the linear dynamical system with matrix (3.82). The benchmark solution is computed with Monte-Carlo (MC) simulation. Also, the zero-order H-model represents the Markovian approximation to the MZ equation, i.e. the MZ equation without the memory term.

defined by the matrix ($N = 100$)

$$A = \begin{bmatrix} -1 & 1 & \dots & (-1)^N \\ 1 & & & \\ \vdots & & B & \\ 1 & & & \end{bmatrix}, \tag{3.86}$$

where $B = e^C \Lambda e^{-C}$ and

$$\Lambda = \begin{bmatrix} -\frac{1}{8} & 0 & \cdots & 0 \\ 0 & -\frac{2}{9} & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & -\frac{N-1}{N+6} \end{bmatrix}, \qquad C = \begin{bmatrix} 0 & 1 & & 0 \\ -1 & 0 & \ddots & \\ & \ddots & \ddots & 1 \\ 0 & & -1 & 0 \end{bmatrix}.$$

It is seen that the H-model converges as we increase the differentiation order in any finite time interval, in agreement with the theoretical prediction of section 3.1.5.

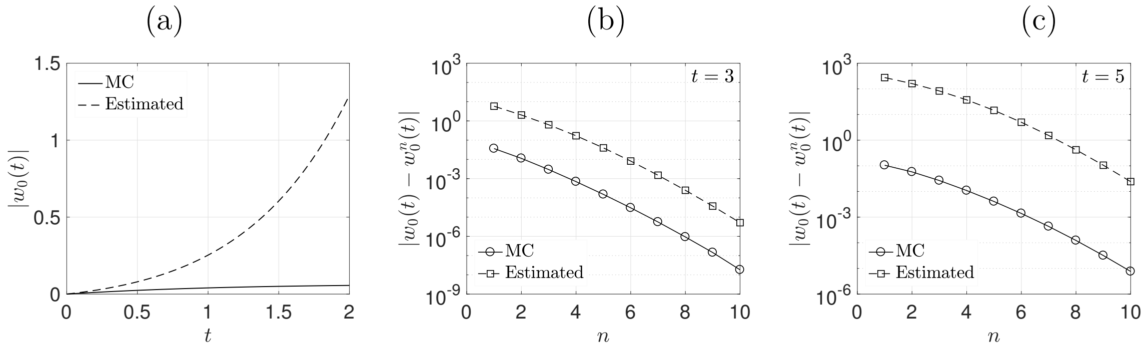

Figure 3.4: Linear dynamical system with matrix (3.82). In (a) we plot the memory term $w_0(t)$ we obtain from Monte Carlo simulation together with the estimated upper bound (3.83). In (b) and (c) we plot H-model approximation error $|w_0(T) - w_0^n(T)|$ together with the upper bound (3.84) for different differentiation orders n and at different times t.

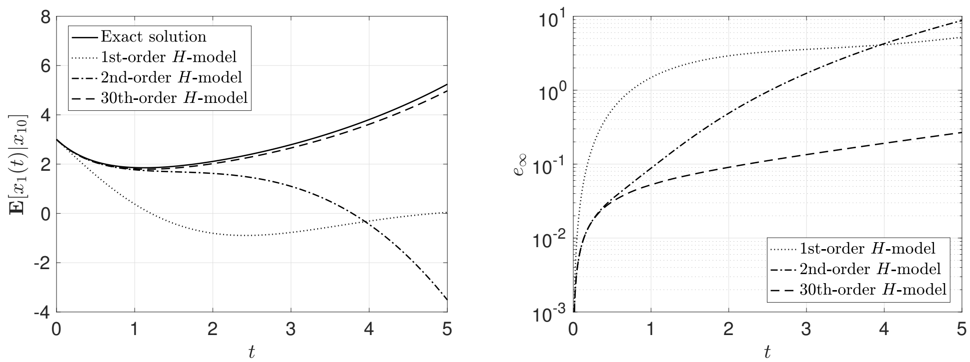

Figure 3.5: Linear dynamical system with matrix A (3.86). Convergence of the H-model to the conditional mean path solution $\mathbb{E}[x_1(t)|x_1(0)]$. The initial condition is set as $x_1(0) = 3$, while $\{x_2(0), \ldots, x_{100}(0)\}$ are i.i.d. Normals.

Nonlinear Dynamical Systems

The hierarchical memory approximation method we discussed in section 3.1.4 can be applied to nonlinear dynamical systems in the form (2.1). As we will see, if we employ the H_t-model then the nonlinearity introduces a closure problem that needs to be addressed properly.

Lorenz-63 System Consider the classical Lorenz-63 model

$$\begin{cases} \dot{x}_1 = \sigma(x_2 - x_1) \\ \dot{x}_2 = x_1(r - x_3) - x_2 \\ \dot{x}_3 = x_1 x_2 - \beta x_3 \end{cases} \tag{3.87}$$

where $\sigma = 10$ and $\beta = 8/3$. The phase space Liouville operator for this ODE is

$$\mathcal{L} = \sigma(x_2 - x_1)\frac{\partial}{\partial x_1} + (x_1(r - x_3) - x_2)\frac{\partial}{\partial x_2} + (x_1 x_2 - \beta x_3)\frac{\partial}{\partial x_3}.$$

We choose the resolved variables to be $\hat{x} = \{x_1, x_2\}$ and aim at formally integrating out $\tilde{x} = x_3$ by using the Mori-Zwanzig formalism. To this end, we set $x_3(0) \sim \mathcal{N}(0,1)$ and consider the zeroth-order H_t-model (t-model)

$$\begin{cases} \dfrac{dx_{1m}}{dt} = \sigma(x_{1m} - x_{2m}), \\ \dfrac{dx_{2m}}{dt} = -x_{2m} + r x_{1m} - t x_{1m}^2 x_{2m}, \end{cases} \tag{3.88}$$

where $x_{1m}(t) = \mathbb{E}[x_1(t)|x_1(0), x_2(0)]$ and $x_{2m}(t) = \mathbb{E}[x_2(t)|x_1(0), x_2(0)]$ are *conditional mean paths*. To obtain this system we introduced the following mean field closure approximation

$$\begin{aligned} t\mathcal{P}e^{t\mathcal{L}}\mathcal{P}\mathcal{L}\mathcal{Q}\mathcal{L}x_2(0) =& - t\mathbb{E}[x_1(t)^2 x_2(t)|x_1(0), x_2(0)], \\ \simeq& - t\mathbb{E}[x_1(t)|x_1(0), x_2(0)]^2\mathbb{E}[x_2(t)|x_1(0), x_2(0)], \\ =& - t x_{1m}^2 x_{2m}. \end{aligned} \tag{3.89}$$

Higher-order H_t-models can be derived based on (3.89). As is well known, if $r < 1$, the fixed point $(0,0,0)$ is a global attractor and exponentially stable. In this case, the t-model (zeroth-order H_t-model) yields accurate prediction of the conditional

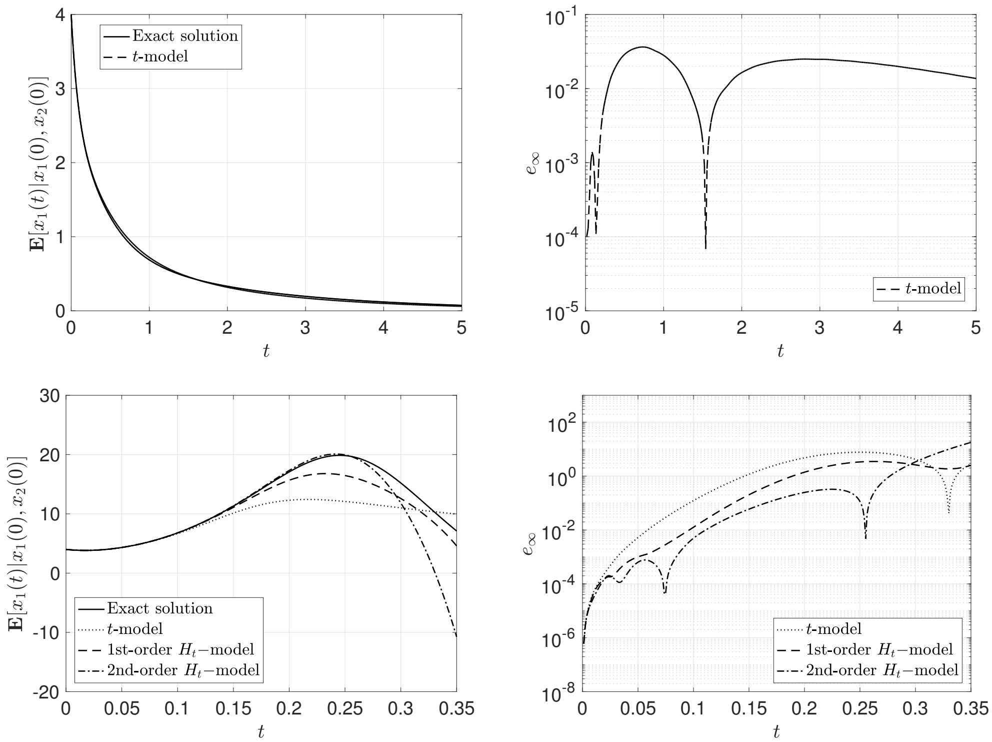

Figure 3.6: Accuracy of the H_t model in representing the conditional mean path in the Lorenz-63 system (3.87). It is seen that if $r = 0.5$ (first row), then the zeroth-order H_t-model, i.e., the t-model, is accurate for long integration times. On the other hand, if we consider the chaotic regime at $r = 28$ (second row) then we see that the t-model and its high-order extension (H_t-model) are accurate only for relatively short time.

mean path for long time (see Figure 3.6). On the other hand, if we consider the chaotic regime at $r = 28$ then the t-model and its higher-order extension, i.e., the H_t-model, are accurate only for relatively short time. This is in agreement with our theoretical predictions. In fact, different from linear systems where the hierarchical representation of the memory integral can be proven to be convergent for long time, in nonlinear systems the memory hierarchy is, in general, provably convergent only in a short time period (Theorem 3.1.7 and Corollary 3.1.4.3). This doesn't mean that the H-model or the H_t-model are not accurate for nonlinear systems. It just means that the accuracy depends on the system, the quantity of interest, and the initial condition.

Modified Lorenz-96 system. As an example of a high dimensional nonlinear dynamical system, we consider the following modified Lorenz-96 system [47, 60]

$$\begin{cases} \dot{x}_1 = -x_1 + x_1 x_2 + F \\ \dot{x}_2 = -x_2 + x_1 x_3 + F \\ \quad \vdots \\ \dot{x}_i = -x_i + (x_{i+1} - x_{i-2}) x_{i-1} + F \\ \quad \vdots \\ \dot{x}_N = x_N - x_{N-2} x_{N-1} + F \end{cases} \tag{3.90}$$

where F is constant. As is well known, depending on the values of N and F this system can exhibit a wide range of behaviors [47]. Suppose we take the resolved variables to be $\hat{x} = \{x_1, x_2\}$. Correspondingly, the unresolved ones, i.e., those we aim at integrating through the MZ framework, are $\tilde{x} = \{x_3, \dots, x_N\}$, which we set to be independent standard normal random variables. By using the mean field approximation (3.89), we obtain the following zeroth-order H_t-model (t-model) of the modified Lorenz-96 system is (3.90)

$$\begin{cases} \dot{x}_{1m} = -x_{1m} + x_{1m} x_{2m} + F, \\ \dot{x}_{2m} = -x_{2m} + F + t(x_{1m}^2 x_{2m} - x_{1m} F). \end{cases} \tag{3.91}$$

In Figure 3.7 we study the accuracy of the H_t-model in representing the conditional mean path for with $F = 5$ and $N = 100$. It is seen that the the H_t-model converges only for short time (in agreement with the theoretical predictions) and it provides results that are more accurate that the classical t-model.

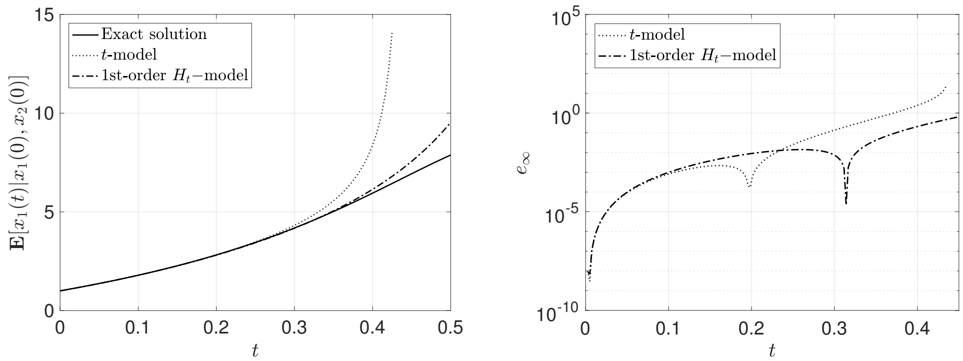

Figure 3.7: Accuracy of the H_t-model in representing the conditional mean path in the Lorenz-96 system (3.87). Here we set $F = 5$ and $N = 100$. It is seen that the H_t-model converges only for short time and provides results that are more accurate that the classical t-model.

3.3 Summary

In this chapter, through a detailed study of the evolution operator $e^{t\mathcal{L}}$ and $e^{t\mathcal{L}\mathcal{Q}}$, we deduce conditions for accuracy and convergence of different approximations of the memory integral in the Mori-Zwanzig (MZ) equation. To the best of our knowledge, this is the first time such rigorous analysis is presented. In particular, we studied the short memory approximation, the t-model and various hierarchical memory approximation techniques. We also derived useful upper bounds for the MZ memory integral, which allowed us to estimate *a priori* the contribution of the MZ memory to the dynamics. Such upper bounds are easily computable for systems with finite-rank projections – such as Mori's projection. Furthermore, if the system is Hamiltonian then the MZ memory bounds we obtained are tight. This can be attributed to the fact that the Koopman operator of a Hamiltonian system is always a contraction relative to the $L^2_{\rho_{eq}}$ norm weighted by the Gibbs canonical distribution ρ_{eq}. In the more general case of dissipative systems, however, we do not have any guarantee that the MZ memory bounds we obtained are tight. Indeed, for dissipative systems the numerical abscissa in equation (3.5) (see also equations (3.10a)-(3.10b)) can be positive, yielding an exponential growth of the memory bounds. In the presence of infinite-rank projections – such as Chorin's projection – we found that it is extremely difficult to obtain explicitly computable upper bounds for the MZ memory or any of its approximations discussed in this paper. The main difficulties are related the lack of an explicitly computable upper bound for the operator semigroup $e^{t\mathcal{L}\mathcal{Q}}$, which involves the unbounded generator $\mathcal{L}\mathcal{Q}$ (see section 3.1.1 and appendix 3.A). To demonstrate the accuracy of the approximation methods and error bounds we developed and discussed throughout the paper, we provided numerical examples involving Hamiltonian dynamical systems (linear and nonlinear), and more general dissipative nonlinear systems evolving from random initial states. The results we obtained are found to be in agreement with our theoretical predictions.

Appendix 3.A Semigroup Bounds via Function Decomposition

In looking for the *numerical abscissa* [20] (i.e., the logarithmic norm) of $\mathcal{L}\mathcal{Q}$, we seek to bound

$$\sup_{\mathcal{D}(\mathcal{L}\mathcal{Q})\ni x\neq 0} \operatorname{Re}\frac{\langle x, \mathcal{L}\mathcal{Q}x\rangle_\sigma}{\langle x, x\rangle_\sigma}.$$

Notice that, if $\mathcal{P}\mathcal{L}\mathcal{Q}=0$, then $\mathcal{L}\mathcal{Q}=\mathcal{Q}\mathcal{L}\mathcal{Q}$, so that we have the previously proven bound

$$\sup \operatorname{Re}\frac{\langle x, \mathcal{Q}\mathcal{L}\mathcal{Q}x\rangle_\sigma}{\langle x, x\rangle_\sigma} \leq -\frac{1}{2}\inf \operatorname{div}_\sigma \mathbf{F}.$$

In this section, we consider what happens when $\mathcal{PLQ} \neq 0$. To that end, let us note that $x \in \mathcal{D}(\mathcal{LQ})$ may be decomposed as

$$x = \mathcal{Q}x + \alpha\mathcal{PLQ}x + \mathcal{P}y$$

where $\alpha \in \mathbb{C}$ and $\mathcal{P}y$ is orthogonal to $\mathcal{PLQ}x$. In other words, we define $\mathcal{P}y$ as

$$\mathcal{P}y = \mathcal{P}x - \frac{\langle \mathcal{PLQ}x, \mathcal{P}x\rangle_\sigma}{\langle \mathcal{PLQ}x, \mathcal{PLQ}x\rangle_\sigma}\mathcal{PLQ}x,$$

and define α as

$$\alpha = \frac{\langle \mathcal{PLQ}x, \mathcal{P}x\rangle_\sigma}{\langle \mathcal{PLQ}x, \mathcal{PLQ}x\rangle_\sigma}.$$

Then

$$\begin{aligned}
\operatorname{Re}\frac{\langle x, \mathcal{LQ}x\rangle_\sigma}{\langle x, x\rangle_\sigma} &= \operatorname{Re}\frac{\langle \mathcal{Q}x + \alpha\mathcal{PLQ}x + \mathcal{P}y, \mathcal{LQ}x\rangle_\sigma}{\langle x, x\rangle_\sigma}, \\
&= \frac{\operatorname{Re}\langle \mathcal{Q}x, \mathcal{LQ}x\rangle_\sigma + \operatorname{Re}(\alpha)\|\mathcal{PLQ}x\|_\sigma^2}{\|\mathcal{Q}x\|_\sigma^2 + |\alpha|^2\|\mathcal{PLQ}x\|_\sigma^2 + \|\mathcal{P}y\|_\sigma^2}, \\
&\leq \max\left[0,\ \frac{\operatorname{Re}\langle \mathcal{Q}x, \mathcal{LQ}x\rangle_\sigma + \operatorname{Re}(\alpha)\|\mathcal{PLQ}x\|_\sigma^2}{\|\mathcal{Q}x\|_\sigma^2 + |\alpha|^2\|\mathcal{PLQ}x\|_\sigma^2}\right].
\end{aligned}$$

Since we assume $\mathcal{PLQ} \neq 0$, there exists x such that $\mathcal{PLQ}x \neq 0$ and then, for any α such that

$$\operatorname{Re}(\alpha) \geq \frac{\operatorname{Re}\langle \mathcal{Q}x, \mathcal{LQ}x\rangle_\sigma}{\|\mathcal{PLQ}x\|_\sigma^2},$$

we have

$$\operatorname{Re}\langle \mathcal{Q}x, \mathcal{LQ}x\rangle_\sigma + \operatorname{Re}(\alpha)\|\mathcal{PLQ}x\|_\sigma^2 \geq 0,$$

so that, for $\mathcal{PLQ} \neq 0$,

$$\sup_{\mathcal{D}(\mathcal{LQ})\ni x\neq 0} \operatorname{Re}\frac{\langle x, \mathcal{LQ}x\rangle_\sigma}{\langle x, x\rangle_\sigma} = \sup_{\mathcal{D}(\mathcal{LQ})\ni x\neq 0} \frac{\operatorname{Re}\langle \mathcal{Q}x, \mathcal{LQ}x\rangle_\sigma + \operatorname{Re}(\alpha)\|\mathcal{PLQ}x\|_\sigma^2}{\|\mathcal{Q}x\|_\sigma^2 + |\alpha|^2\|\mathcal{PLQ}x\|_\sigma^2}.$$

Now, fix any $\mathcal{Q}x \in \mathcal{D}(\mathcal{L}) \neq 0$ and consider the expression

$$\frac{\operatorname{Re}\langle \mathcal{Q}x, \mathcal{LQ}x\rangle_\sigma + \operatorname{Re}(\alpha)\|\mathcal{PLQ}x\|_\sigma^2}{\|\mathcal{Q}x\|_\sigma^2 + |\alpha|^2\|\mathcal{PLQ}x\|_\sigma^2} = \frac{\xi + \operatorname{Re}(\alpha)\beta^2}{1 + |\alpha|^2\beta^2}$$

where

$$\xi = \xi(\mathcal{Q}x) = \operatorname{Re}\frac{\langle \mathcal{Q}x, \mathcal{LQ}x\rangle_\sigma}{\|\mathcal{Q}x\|_\sigma^2}, \qquad \beta = \beta(\mathcal{Q}x) = \frac{\|\mathcal{PLQ}x\|_\sigma}{\|\mathcal{Q}x\|_\sigma}.$$

Then, for this fixed $\mathcal{Q}x$,

$$\frac{\xi + \operatorname{Re}(\alpha)\beta^2}{1 + |\alpha|^2\beta^2} \leq \max_{a\in\mathbb{R}} \frac{\xi + a\beta^2}{1 + a^2\beta^2}.$$

Differentiating w.r.t. a and setting equal to zero, we find that the latter expression is extremized when

$$0 = \beta^2(1 + a^2\beta^2) - 2a\beta^2(\xi + a\beta^2),$$

i.e., when

$$\beta^2 a^2 + 2\xi a - 1 = 0, \qquad a = \frac{-\xi \pm \sqrt{\xi^2 + \beta^2}}{\beta^2}.$$

Since $\beta^2 > 0$, $\dfrac{\xi + a\beta^2}{1 + a^2\beta^2}$ is maximized at $\hat{a} = \dfrac{-\xi + \sqrt{\xi^2 + \beta^2}}{\beta^2}$. Then

$$\xi + \hat{a}\beta^2 = \sqrt{\xi^2 + \beta^2}, \qquad 1 + \hat{a}^2\beta^2 = 2\,(1 - \xi\hat{a}) = 2\left(\frac{\xi^2 + \beta^2 - \xi\sqrt{\xi^2 + \beta^2}}{\beta^2}\right),$$

so that

$$\max_{a\in\mathbb{R}} \frac{\xi + a\beta^2}{1+a^2\beta^2} = \frac{\xi + \hat{a}\beta^2}{1+\hat{a}^2\beta^2} = \frac{1}{2}\frac{\beta^2}{\sqrt{\xi^2+\beta^2}-\xi} = \frac{1}{2}\left(\sqrt{\xi^2+\beta^2}+\xi\right).$$

Therefore,

$$\sup_{\mathcal{D}(\mathcal{L}\mathcal{Q})\ni x\neq 0} \operatorname{Re}\frac{\langle x, \mathcal{L}\mathcal{Q}x\rangle_\sigma}{\langle x, x\rangle_\sigma} = \sup_{\mathcal{D}(\mathcal{L})\ni(\mathcal{Q}x)\neq 0} \frac{1}{2}\left[\sqrt{\xi^2(\mathcal{Q}x)+\beta^2(\mathcal{Q}x)}+\xi(\mathcal{Q}x)\right].$$

When $\mathcal{P}\mathcal{L}\mathcal{Q}$ is unbounded, which is the typical case when $\mathcal{P}$ is an infinite-rank projection, such as most conditional expectations, there is unlikely to be a finite numerical abscissa for $\mathcal{L}\mathcal{Q}$. In particular, notice that if $\operatorname{div}_\sigma(F)$ is a bounded function (bounded both above and below), then $\xi(\mathcal{Q}x)$ is bounded for all $\mathcal{Q}x$ while $\beta(\mathcal{Q}x)$ is unbounded, in which case

$$\sup_{\mathcal{D}(\mathcal{L}\mathcal{Q})\ni x\neq 0} \operatorname{Re}\frac{\langle x, \mathcal{L}\mathcal{Q}x\rangle_\sigma}{\langle x, x\rangle_\sigma} = \infty.$$

It follows [89] that in these cases, $\|e^{t\mathcal{L}\mathcal{Q}}\|_\sigma$ has infinite slope at $t=0$, and therefore there is no finite ω such that $\|e^{t\mathcal{L}\mathcal{Q}}\|_\sigma \leq e^{\omega t}$ for all $t \geq 0$ (see [20]). Assuming still that $\mathcal{L}\mathcal{Q}$ generates a strongly continuous semigroup, we must look to bound the semigroup as $\|e^{t\mathcal{L}\mathcal{Q}}\|_\sigma \leq Me^{\omega t}$, where

$$\omega > \omega_0 = \lim_{t\to\infty} \frac{\ln\|e^{t\mathcal{L}\mathcal{Q}}\|_\sigma}{t}$$

and

$$M \geq M(\omega) = \sup\{\|e^{t\mathcal{L}\mathcal{Q}}\|_\sigma e^{-\omega t} : t \geq 0\}.$$

On the other hand, if $\mathcal{PLQ}$ is a bounded operator, for example when $\mathcal{P}$ is a finite-rank projection (e.g., Mori's projection), there exists a finite value for the numerical abscissa. Indeed, in this case, since ζ is bounded by $\omega = -\inf \operatorname{div}_\sigma(F)$, the numerical abscissa of $\mathcal{LQ}$ may be bounded as

$$\omega_{\mathcal{LQ}} = \sup_{\mathcal{D}(\mathcal{LQ})\ni x\neq 0} \operatorname{Re} \frac{\langle x, \mathcal{LQ}x\rangle_\sigma}{\langle x, x\rangle_\sigma} \leq \frac{1}{2}\left[\sqrt{\omega^2 + \|\mathcal{PLQ}\|_\sigma^2} + \omega\right]. \tag{3.98}$$

Alternatively, in the case of finite rank $\mathcal{P}$, the operator $\mathcal{LQ}$ may be thought of as a bounded perturbation of $\mathcal{L}$, i.e. $\mathcal{LQ} = \mathcal{L} - \mathcal{LQ}$, and the numerical abscissa of $\mathcal{LQ}$ can be bounded using the bounded perturbation theorem [28, III.1.3], obtaining

$$\omega_{\mathcal{LQ}} \leq \omega + \|\mathcal{LP}\|_\sigma. \tag{3.99}$$

Either of these bounds for $\omega_{\mathcal{LQ}}$ can be used to bound the semigroup norm

$$\begin{aligned} \|e^{t\mathcal{LQ}}\|_\sigma &\leq e^{\omega_{\mathcal{LQ}}t} \leq e^{\frac{1}{2}\left(\sqrt{\omega^2+\|\mathcal{PLQ}\|_\sigma^2}+\omega\right)t} \\ \|e^{t\mathcal{LQ}}\|_\sigma &\leq e^{\omega_{\mathcal{LQ}}t} \leq e^{(\omega+\|\mathcal{LP}\|_\sigma)t}. \end{aligned}$$

Which of these two estimates gives the tighter bound will generally depend on the values of $\|\mathcal{PLQ}\|_\sigma$ and $\|\mathcal{LP}\|_\sigma$. It may be noted, however, that, when σ is invariant, $\omega = 0$ and $\mathcal{L}$ is skew-adjoint, so that

$$\|\mathcal{PLQ}\|_\sigma = \|\mathcal{QL}^\dagger\mathcal{P}\|_\sigma = \|\mathcal{QLP}\|_\sigma \leq \|\mathcal{LP}\|_\sigma$$

and therefore the bound in (3.98) is half that of (3.99).

Appendix 3.B Estimate of the MZE in PDF Space

It was shown in [22] that the Banach dual of (2.4) defines an evolution in the probability density function space. Specifically, the joint probability density function of the state vector $u(t)$ that solves equation (2.1) is pushed forward by the Frobenius-Perron operator $\mathcal{F}(t,0)$ (Banach dual of the Koopman operator (2.3))

$$p(x,t) = \mathcal{F}(t,s)p(x,s), \qquad \mathcal{F}(t,s) = e^{(t-s)\mathcal{M}}, \tag{3.101}$$

where

$$\mathcal{M}(x)p(x,t) = -\nabla \cdot (F(x)p(x,t)). \tag{3.102}$$

By introducing a projection $\mathcal{P}$ in the space of probability density functions[9] and its complement $\mathcal{Q} = \mathcal{I} - \mathcal{P}$, it is easy to show that the projected density $\mathcal{P}p$ satisfies the MZ equation [96]

$$\frac{\partial \mathcal{P}p(t)}{\partial t} = \mathcal{P}\mathcal{M}\mathcal{P}p(t) + \mathcal{P}e^{t\mathcal{Q}\mathcal{M}}\mathcal{Q}p(0) + \int_0^t \mathcal{P}\mathcal{M}e^{(t-s)\mathcal{Q}\mathcal{M}}\mathcal{Q}\mathcal{M}\mathcal{P}p(s)ds. \tag{3.103}$$

In the next sections we perform an analysis of different types of approximations of the MZ memory integral

$$\int_0^t \mathcal{P}\mathcal{M}e^{(t-s)\mathcal{Q}\mathcal{M}}\mathcal{Q}\mathcal{M}\mathcal{P}p(s)ds. \tag{3.104}$$

The main objective of such analysis is to establish rigorous error bounds for widely used approximation methods, and also propose new provably convergent approximation schemes.

[9] With some abuse of notation we denote the projections $\mathcal{P}$ and $\mathcal{Q}$ in the PDF space with the same letter we used for projections in the phase space.

3.B.1 Analysis of the Memory Integral

In this section, we develop the analysis of the memory integral arising in the PDF formulation of the MZ equation. The starting point is the definition (3.104). As before, we begin with the following estimate of upper bound estimation of the integral

Theorem 3.B.1. (Memory growth) *Let $e^{t\mathcal{M}\mathcal{Q}}$ and $e^{t\mathcal{M}}$ be strongly continuous semigroups with upper bounds $\|e^{t\mathcal{M}}\| \leq Me^{t\omega}$ and $\|e^{t\mathcal{M}\mathcal{Q}}\| \leq M_{\mathcal{Q}}e^{t\omega_{\mathcal{Q}}}$, and let $T > 0$ be a fixed integration time. Then for any $0 \leq t \leq T$ we have*

$$\left\| \int_0^t \mathcal{P}\mathcal{M}e^{(t-s)\mathcal{Q}\mathcal{M}}\mathcal{Q}\mathcal{M}\mathcal{P}p(s)ds \right\| \leq N_0(t),$$

where

$$N_0(t) = \begin{cases} tC_4, & \omega_{\mathcal{Q}} = 0; \\ \dfrac{C_4}{\omega_{\mathcal{Q}}}(e^{t\omega_{\mathcal{Q}}} - 1), & \omega_{\mathcal{Q}} \neq 0; \end{cases}$$

and $C_4 = \max\limits_{0\leq s\leq T} \|\mathcal{P}\|\|\mathcal{M}\mathcal{P}\mathcal{M}p(s)\|$. Moreover, $N(t)$ satisfies $\lim\limits_{t\to 0} N(t) = 0$.

Proof. Consider

$$\begin{aligned} \left\| \int_0^t \mathcal{P}\mathcal{M}e^{(t-s)\mathcal{Q}\mathcal{M}}\mathcal{Q}\mathcal{M}\mathcal{P}p(s)ds \right\| &= \left\| \int_0^t \mathcal{P}e^{(t-s)\mathcal{M}\mathcal{Q}}\mathcal{M}\mathcal{Q}\mathcal{M}\mathcal{P}p(s)ds \right\| \\ &\leq C_4 M_{\mathcal{Q}} \int_0^t e^{(t-s)\omega_{\mathcal{Q}}}ds \\ &= \begin{cases} tC_4, & \omega_{\mathcal{Q}} = 0 \\ \dfrac{C_4}{\omega_{\mathcal{Q}}}(e^{t\omega_{\mathcal{Q}}} - 1), & \omega_{\mathcal{Q}} \neq 0 \end{cases} \end{aligned}$$

where $C_4 = \max\limits_{0\leq s\leq T} \|\mathcal{P}\|\|\mathcal{M}\mathcal{Q}\mathcal{M}\mathcal{P}p(s)\|$.

□

Theorem 3.B.2. (Memory approximation via the t-model) *Let $e^{t\mathcal{M}\mathcal{Q}}$ and $e^{t\mathcal{M}}$ be strongly continuous semigroups with bounds $\|e^{t\mathcal{M}}\| \leq Me^{t\omega}$ and $\|e^{t\mathcal{M}\mathcal{Q}}\| \leq M_{\mathcal{Q}}e^{t\omega_{\mathcal{Q}}}$, and let $T > 0$ be a fixed integration time. If the function $k(s,t) = \mathcal{P}\mathcal{M}e^{(t-s)\mathcal{Q}\mathcal{M}}\mathcal{Q}\mathcal{M}\mathcal{P}p(s)$ (integrand of the memory term) is at least twice differentiable respect to s for all $t \geq 0$, then*

$$\left\| \int_0^t \mathcal{P}\mathcal{M}e^{(t-s)\mathcal{Q}\mathcal{M}}\mathcal{Q}\mathcal{M}\mathcal{P}p(s)ds - t\mathcal{P}\mathcal{M}\mathcal{Q}\mathcal{M}\mathcal{P}p(t) \right\| \leq N_1(t)$$

where $N_1(t)$ is defined as

$$N_1(t) = \begin{cases} t(M_{\mathcal{Q}}+1)C_4, & \omega_{\mathcal{Q}} = 0 \\ \dfrac{C_2 M_{\mathcal{Q}}}{\omega_{\mathcal{Q}}}(e^{t\omega_{\mathcal{Q}}} - 1) + tC_4, & \omega_{\mathcal{Q}} \neq 0 \end{cases}$$

and C_4 is as in Theorem 3.B.1.

Proof.

$$\begin{aligned} &\left\| \int_0^t \mathcal{P}\mathcal{M}e^{(t-s)\mathcal{Q}\mathcal{M}}\mathcal{Q}\mathcal{M}\mathcal{P}p(s)ds - t\mathcal{P}\mathcal{M}\mathcal{Q}\mathcal{M}\mathcal{P}p(t) \right\| \\ &\leq C_4 M_{\mathcal{Q}} \int_0^t e^{(t-s)\omega_{\mathcal{Q}}} ds + C_4 t \\ &= \begin{cases} t(M_{\mathcal{Q}}+1)C_4, & \omega_{\mathcal{Q}} = 0 \\ \dfrac{C_4 M_{\mathcal{Q}}}{\omega_{\mathcal{Q}}}(e^{t\omega_{\mathcal{Q}}} - 1) + tC_4, & \omega_{\mathcal{Q}} \neq 0 \end{cases} \end{aligned}$$

□

3.B.2 Hierarchical Memory Approximation in PDF Space

The hierarchical memory approximation methods we discussed in section 3.1.4 can be also developed in the PDF space. To this end, let us first define

$$v_0(t) = \int_0^t \mathcal{P}\mathcal{M}e^{(t-s)\mathcal{Q}\mathcal{M}}\mathcal{Q}\mathcal{M}\mathcal{P}p(s)ds. \tag{3.105}$$

By repeatedly differentiating $v_0(t)$ with respect to time (assuming $v_0(t)$ smooth enough) we obtain the hierarchy of equations

$$\frac{\partial}{\partial t}v_{i-1}(t) = \mathcal{P}\mathcal{M}(\mathcal{Q}\mathcal{M})^i\mathcal{P}p(t) + v_i(t) \qquad i = 1, \ldots, n$$

where,

$$v_i(t) = \int_0^t \mathcal{P}\mathcal{M}e^{(t-s)\mathcal{Q}\mathcal{M}}(\mathcal{Q}\mathcal{M})^{i+1}\mathcal{P}p(s)ds.$$

By following closely the discussion in section 3.1.4 we introduce the hierarchy of memory equations

$$\begin{cases} \dfrac{dv_0^n(t)}{dt} = \mathcal{P}\mathcal{M}\mathcal{Q}\mathcal{M}\mathcal{P}p(t) + v_1^n(t) \\ \dfrac{dv_1^n(t)}{dt} = \mathcal{P}\mathcal{M}(\mathcal{Q}\mathcal{M})^2\mathcal{P}p(t) + v_2^n(t) \\ \quad\vdots \\ \dfrac{dv_{n-1}^n(t)}{dt} = \mathcal{P}\mathcal{M}(\mathcal{Q}\mathcal{M})^n\mathcal{P}p(t) + v_n^{e_n}(t) \end{cases} \tag{3.106}$$

and approximate the last term in such hierarchy in different ways. Specifically, we consider

$$
\begin{aligned}
v_n^{e_n}(t) &= \int_t^t \mathcal{P}\mathcal{M}e^{(t-s)\mathcal{Q}\mathcal{M}}(\mathcal{Q}\mathcal{M})^{n+1}\mathcal{P}p(s)ds = 0 && (H\text{-model}),\\
v_n^{e_n}(t) &= \int_{\max(0,t-\Delta t)}^t \mathcal{P}\mathcal{M}e^{(t-s)\mathcal{Q}\mathcal{M}}(\mathcal{Q}\mathcal{M})^{n+1}\mathcal{P}p(s)ds && (\text{Type I FMA}),\\
v_n^{e_n}(t) &= \int_{\min(t,t_n)}^t \mathcal{P}\mathcal{M}e^{(t-s)\mathcal{Q}\mathcal{M}}(\mathcal{Q}\mathcal{M})^{n+1}\mathcal{P}p(s)ds && (\text{Type II FMA}),\\
v_n^{en}(t) &= t\mathcal{P}\mathcal{M}(\mathcal{Q}\mathcal{M})^{n+1}\mathcal{P}p(t) && (H_t\text{-model}).
\end{aligned}
$$

Hereafter we establish the accuracy of the approximation schemes resulting from the substitution of each $v_n^{e_n}(t)$ above into (3.106).

Theorem 3.B.3. (Accuracy of the H-model) *Let $e^{t\mathcal{M}}$ and $e^{t\mathcal{M}\mathcal{Q}}$ be strongly continuous semigroups, $T > 0$ a fixed integration time, and*

$$
\beta_i = \frac{\sup_{s\in[0,T]} \|(\mathcal{M}\mathcal{Q})^{i+1}\mathcal{M}\mathcal{P}p(s)\|}{\sup_{s\in[0,T]} \|(\mathcal{M}\mathcal{Q})^{i}\mathcal{M}\mathcal{P}p(s)\|}, \qquad 1 \le i \le n. \tag{3.107}
$$

Then for $1 \le q \le n$ we have

$$
\|v_0(t) - v_0^q(t)\| \le N_2^q(t),
$$

where

$$
N_2^q(t) = A_2 C_4 \left(\prod_{i=1}^q \beta_i\right) \frac{t^{q+1}}{(q+1)!},
$$

$A_2 = \max_{s\in[0,T]} e^{s\omega_{\mathcal{Q}}}$*, and C_4 is as in Theorem 3.B.1.*

Proof. The error at the n-th level can be bounded as

$$\begin{aligned}\|v_0(t) - v_0^q(t)\| &\leq \int_0^t \int_0^{\tau_q} \cdots \int_0^{\tau_2} \|\mathcal{P}\mathcal{M}e^{(\tau_1 - s)\mathcal{Q}\mathcal{M}}(\mathcal{Q}\mathcal{M})^{p+1}\mathcal{P}p(s)\| ds d\tau_1 \ldots d\tau_q \\ &\leq A_2 C_4 \left(\prod_{i=1}^{q} \beta_i\right) \frac{t^{q+1}}{(q+1)!},\end{aligned}$$

where

$$A_2 = \max_{s \in [0,T]} e^{s\omega_{\mathcal{Q}}} = \begin{cases} 1 & \omega_{\mathcal{Q}} \leq 0 \\ e^{T\omega_{\mathcal{Q}}} & \omega_{\mathcal{Q}} \geq 0 \end{cases}. \tag{3.108}$$

Let

$$\beta_i = \frac{\sup_{s \in [0,T]} \|(\mathcal{M}\mathcal{Q})^{i+1}\mathcal{M}\mathcal{P}p(s)\|}{\sup_{s \in [0,T]} \|(\mathcal{M}\mathcal{Q})^{i}\mathcal{M}\mathcal{P}p(s)\|}, \tag{3.109}$$

under the assumption that these quantities are finite. Then we have

$$\|v_0(t) - v_0^q(t)\| \leq A_2 C_4 \left(\prod_{i=1}^{q} \beta_i\right) \frac{t^{q+1}}{(q+1)!}.$$

□

Corollary 3.B.3.1. (Uniform convergence of the H-model) *If β_i in Theorem 3.B.3 satisfy*

$$\beta_i < \frac{i+1}{T}, \quad 1 \leq i \leq n$$

for any fixed time T, then there exits a sequence $\delta_1 > \delta_2 > \cdots > \delta_n$ such that

$$\|v_0(T) - v_0^q(T)\| \leq \delta_q,$$

where $1 \leq q \leq n$.

Corollary 3.B.3.2. (Asymptotic convergence of the H-model) *If* β_i *in Theorem 3.B.3 satisfy*

$$\beta_i < C, \quad 1 \leq i < +\infty$$

for some constant C, *then for any fixed time* T *and arbitrary* $\delta > 0$, *there exits an integer* q *such that for all* $n > q$,

$$\|v_0(T) - v_0^n(T)\| \leq \delta.$$

The proofs of the Corollary 3.B.3.1 and 3.B.3.2 closely follow the proofs of Corollary 3.1.4.1 and 3.1.4.2. Therefore we omit details here.

Theorem 3.B.4. (Accuracy of Type-I FMA) *Let* $e^{t\mathcal{M}}$ *and* $e^{t\mathcal{MQ}}$ *be strongly continuous semigroups,* $T > 0$ *a fixed integration time, and let*

$$\beta_i = \frac{\sup\limits_{s\in[0,T]} \|(\mathcal{MQ})^{i+1}\mathcal{MP}p(s)\|}{\sup\limits_{s\in[0,T]} \|(\mathcal{MQ})^{i}\mathcal{MP}p(s)\|}, \qquad 1 \leq i \leq n. \tag{3.110}$$

Then for $1 \leq q \leq n$

$$\|v_0(t) - v_0^q(t)\| \leq N_3^q(t),$$

where

$$N_3^q(t) = A_2 C_4 \left(\prod_{i=1}^{q} \beta_i\right) \frac{(t - \Delta t_q)^{q+1}}{(q+1)!},$$

and C_4 *is as in Theorem 3.B.1.*

Proof. The proof is very similar with the proof of Theorem 3.1.5. We begin with the estimate of $v_0(t) - v_0^q(t)$

$$\int_0^{\max(0,t-\Delta t_q)} \int_0^{\tilde{\tau}_q} \cdots \int_0^{\tilde{\tau}_2} \int_0^{\tilde{\tau}_1} \mathcal{P} e^{(\tilde{\tau}_1 + \Delta t_q - s)\mathcal{MQ}} (\mathcal{MQ})^{q+1} \mathcal{MQ} p(s) ds d\tilde{\tau}_1 \cdots d\tilde{\tau}_q,$$

$$= \begin{cases} 0 & 0 \le t \le \Delta t_q \\ \displaystyle\int_0^{t-\Delta t_q} \int_0^{\sigma} \frac{(t-\Delta t_q - \sigma)^{q-1}}{(q-1)!} \mathcal{P} e^{(\sigma + \Delta t_q - s)\mathcal{MQ}} (\mathcal{MQ})^{q+2} p(s) ds d\sigma & t \ge \Delta t_q. \end{cases}$$

This can be bounded by following the technique in the proof of Theorem 3.B.3. This yields

$$\|v_0(t) - v_0^q(t)\| \le \begin{cases} 0 & 0 \le t \le \Delta t_q \\ A_2 C_4 \left(\displaystyle\prod_{i=1}^{q} \beta_i \right) \dfrac{(t-\Delta t_q)^{q+1}}{(q+1)!} & t \ge \Delta t_q \end{cases}. \tag{3.112a}$$

□

Corollary 3.B.4.1. (Uniform convergence of Type-I FMA) *If β_i in Theorem 3.B.4 satisfy*

$$\beta_i < (i+1) \left[\frac{\delta i!}{C_4 A_2 \left(\prod_{k=1}^{i} \beta_k \right)} \right]^{-\frac{1}{i}} \tag{3.113}$$

then for any $\delta > 0$, there exists an ordered time sequence $\Delta t_n < \Delta t_{n-1} < \cdots < \Delta t_1 < T$ such that

$$\|w_0(T) - w_0^q(T)\| \le \delta, \quad 1 \le q \le n$$

and which satisfies

$$\Delta t_q \le T - \left[\frac{\delta (q+1)!}{C_4 A_2 \left(\prod_{i=1}^{q} \beta_i \right)} \right]^{\frac{1}{q+1}}.$$

The proof is very similar with the proof of Corollary 3.1.5.1 and therefore we omit it.

Theorem 3.B.5. (Accuracy of Type-II FMA) *Let $e^{t\mathcal{M}}$ and $e^{t\mathcal{M}\mathcal{Q}}$ be strongly continuous semigroups and $T > 0$ a fixed integration time. Set*

$$\beta_i = \frac{\sup_{s\in[0,T]} \|(\mathcal{M}\mathcal{Q})^{i+1}\mathcal{M}\mathcal{P}p(s)\|}{\sup_{s\in[0,T]} \|(\mathcal{M}\mathcal{Q})^{i}\mathcal{M}\mathcal{P}p(s)\|}, \qquad 1 \leq i \leq n. \tag{3.114}$$

Then for $1 \leq q \leq n$

$$\|v_0(t) - v_0^q(t)\| \leq N_4^q(t),$$

where

$$N_4^q(t) = \frac{C_4}{\omega_\mathcal{Q}}\left[1 - e^{-t_q\omega_\mathcal{Q}}\right]\left(\prod_{i=1}^{q}\beta_i\right) f_q(\omega_\mathcal{Q}, t),$$

$f_q(\omega_\mathcal{Q}, t)$ is defined in (3.41) and C_4 is as in Theorem 3.B.1.

Proof. The proof is very similar with the proof of Theorem 3.1.6. Hereafter we provide the proof for the case when $\omega_\mathcal{Q} > 0$. Other cases can be easily obtained by using the same method. First of all, we have error estimation

$$\begin{aligned}
\|v_0(t) - v_0^q(t)\| &\leq \int_0^t\int_0^{\tau_q}\cdots\int_0^{\tau_2}\int_0^{t_q} \|\mathcal{P}\mathcal{M}e^{(\tau_1-s)\mathcal{Q}\mathcal{M}}(\mathcal{Q}\mathcal{M})^{q+1}\mathcal{P}p(s)\| ds d\tau_1\cdots d\tau_q \\
&\leq C_4\left(\prod_{i=1}^{q}\beta_i\right)\int_0^t\int_0^{\tau_q}\cdots\int_0^{\tau_2}\int_0^{t_q} e^{(\tau_1-s)\omega_\mathcal{Q}} ds d\tau_1\cdots d\tau_q \\
&= \frac{C_4}{\omega_\mathcal{Q}}\left[1 - e^{-t_q\omega_\mathcal{Q}}\right]\left(\prod_{i=1}^{q}\beta_i\right) f_q(\omega_\mathcal{Q}, t),
\end{aligned}$$

where $f_q(\omega_\mathcal{Q}, t)$ is as in (3.41).

□

Corollary 3.B.5.1. (Uniform convergence of Type-II FMA) *If β_i in Theorem 3.B.5 satisfy*

$$\begin{cases} \beta_i < \dfrac{i}{T}, & \omega_{\mathcal{Q}} = 0 \\ \beta_i < \omega_{\mathcal{Q}} \dfrac{e^{T\omega_{\mathcal{Q}}} - \sum_{k=0}^{i-2} \frac{(T\omega_{\mathcal{Q}})^k}{k!}}{e^{T\omega_{\mathcal{Q}}} - \sum_{k=0}^{i-1} \frac{(T\omega_{\mathcal{Q}})^k}{k!}}, & \omega_{\mathcal{Q}} \neq 0 \end{cases}$$

for all $t \in [0, T]$, then for arbitrarily small $\delta > 0$ there exists an ordered time sequence $0 < t_0 < t_1 < \ldots t_n < T$ such that

$$\|v_0(T) - v_0^q(T)\| \leq \delta, \quad 1 \leq q \leq n$$

which satisfies

$$t_q \geq \frac{1}{\omega_{\mathcal{Q}}} \ln \left[1 - \frac{\delta \omega_{\mathcal{Q}}^{q+1}}{C_4 \left(\prod_{i=1}^q \beta_i \right) \left[e^{T\omega_{\mathcal{Q}}} - \sum_{k=0}^{q-1} \frac{(T\omega_{\mathcal{Q}})^k}{k!} \right]} \right].$$

Proof. To ensure that $\|v_0(t) - v_0^q(t)\| \leq \delta$ for all $0 \leq t \leq T$, we can take (for $\omega_{\mathcal{Q}} > 0$)

$$\frac{C_4}{\omega_{\mathcal{Q}}} \left[1 - e^{-t_q \omega_{\mathcal{Q}}} \right] \left(\prod_{i=1}^q \beta_i \right) f_q(\omega_{\mathcal{Q}}, T) = \max_{t \in [0,T]} \frac{C_2}{\omega_{\mathcal{Q}}} \left[1 - e^{-t_q \omega_{\mathcal{Q}}} \right] \left(\prod_{i=1}^q \beta_i \right) f_q(\omega_{\mathcal{Q}}, t) \leq \delta.$$

Therefore

$$e^{-t_q \omega_{\mathcal{Q}}} \geq 1 - \frac{\delta \omega_{\mathcal{Q}}}{C_4 \left(\prod_{i=1}^q \beta_i \right) f_q(\omega_{\mathcal{Q}}, T)} \qquad \text{i.e.,}$$

$$t_q \leq \frac{1}{\omega_{\mathcal{Q}}} \ln \left[1 - \frac{\delta \omega_{\mathcal{Q}}^{q+1}}{C_4 \left(\prod_{i=1}^q \beta_i \right) \left[e^{T\omega_{\mathcal{Q}}} - \sum_{k=0}^{q-1} \frac{(T\omega_{\mathcal{Q}})^k}{k!} \right]} \right].$$

Since for $\omega_{\mathcal{Q}} > 0$, we have have condition

$$\beta_i(t) < \omega_{\mathcal{Q}} \frac{e^{T\omega_{\mathcal{Q}}} - \sum_{k=0}^{i-2} \frac{(T\omega_{\mathcal{Q}})^k}{k!}}{e^{T\omega_{\mathcal{Q}}} - \sum_{k=0}^{i-1} \frac{(T\omega_{\mathcal{Q}})^k}{k!}}.$$

Thus, there exists an ordered time sequence $0 < t_1 < \cdots < t_n$ such that $\|v_0(T) - v_0^n(T)\| \leq \delta$. As in Theorem 3.1.6, this δ-bound on the error holds for all t_n (with upper bound as above), which implies the existence of such an increasing time sequence $0 < t_1 < \cdots < t_n$ with t_n bounded from below by the same quantities. □

Chapter 4

Analysis of the MZE for stochastic system

In this chapter, we focus on the analysis of the effective Mori-Zwanzig equation (EMZE) derived in Section 2.3. Similar to the case for classical stochastic system, the memory integral will be our main concentration since it encoded the interaction between the complicated orthogonal dynamics and the steaming term. In Chapter 3, we have shown that the semigroup properties of the evolution operator $e^{t\mathcal{L}}$ and $e^{t\mathcal{L}\mathcal{Q}}$ are closely related with the convergence of series expansion of the Mori-Zwanzig equation and upper bounds of the memory kernel. This inspired us to apply the estimate of semigroup $e^{t\mathcal{K}}$ for EMZE to study the orthogonal dynamics and the corresponding memory kernel. In mathematical physics, the dynamical properties of $e^{t\mathcal{K}}$ happen to be an important research topic with special emphasis on its convergence to equilibrium as $t \to \infty$. In the past twenty years, there are significant developments in this field. In particular, the study of hypoellipticity and hypocoercivity [26, 24, 25, 38, 39, 98, 72] enables us to establish prior estimates on the *rates of convergence* for both linear and nonlinear kinetic equations. In this chapter, we give a first attempt to use the developed analytic techniques for semigroup $e^{t\mathcal{K}}$ on the study of the Mori-Zwanzig equation. In particular, we show that the classical hypoelliptic analysis mainly developed by Eckmann & Hairer [26, 24, 25], Hèrau & Nier [39] and Nier & Helffer [38] can be used to locate the spectrum of $\mathcal{QKQ}$ inside a cusp-like region of the complex plane. Hence the analytic method used to deduce the exponential convergence to equilibrium for Fokker-Planck equation can applied to get the exponential convergence of the orthogonal dynamics and the MZ memory kernel. We will use the Langevin dynamics as an example to derive the explicit convergence estimates.

4.1 Abstract analysis

In this section, we analyze the effective Mori-Zwanzig equation using Hörmander analysis established in a series of paper [26, 24, 25, 38, 39, 98, 72]. The approach was mainly developed by Hérau and Nier [39], Eckmann and Hairer [26, 24, 25], Helffer and Nier [38] in the study of linear hypoelliptic equation, which can be seen as an extension of Kohn's method for Hörmander operators. The following notations are used in Section 4.1 and 4.2:

Notations

1. C and C_α are used to represent constants, where C_α refers to a *specific* constant and C, when it appears in an inequality, always means *some constant* such that the inequality holds.

2. $\langle\cdot,\cdot\rangle$ is the inner product in flat Hilbert space $L^2(\mathbb{R}^n)$. $\|\cdot\|$ denotes $L^2(\mathbb{R}^n)$ norm for function inputs and $l^2(\mathbb{R}^n)$ norm for vector inputs.

3. The same notation is used to represent the closable operator and its closure in some function space.

4.1.1 Estimate of $\mathcal{K}$

Following the same notation in [25, 106], we begin our discussion by reviewing some of known results on the Kolmogorov operator $\mathcal{K}$. In [25], Eckmann and Hairer considered the following Hörmander-type operator

$$\mathcal{K} = \sum_{i=1}^{m} X_i^* X_i + X_0 + f. \tag{4.1}$$

where X_i, $0 \leq i \leq m$ denotes the first order differential operator, X_i^* is the formal adjoint of X_i in L^2, f is a function with at most polynomial growth. It is shown in [26, 24, 25, 38, 39] that for many SDE (2.21), the generator of the Markovian semigroup $\mathcal{M}(t, 0)$ can be written in the form (4.1). To derive the spectral estimate for operator $\mathcal{K}$ with polynomially bounded coefficients, we define the following class of functions and operators:

Definition 1 (Eckmann and Hairer [25])**.** For $N \in \mathbb{R}$, we define the set $\mathbf{Pol}_0^N$ of polynomially growing functions by

$$\mathbf{Pol}_0^N = \left\{ f \in \mathcal{C}^\infty(\mathbb{R}^n) | \forall \alpha, \sup_{x \in \mathbb{R}^n} (1 + \|x\|)^{-N} |\partial^\alpha f(x)| \leq C_\alpha \right\},$$

where α denotes the multi-index of arbitrary order. Similarly, we define sets $\mathbf{Pol}_k^N$ of $k-$th order differential operators with elements written as

$$G = G_0(x) + \sum_{j=1}^{n} \sum_{i=1}^{k} G_j^i(x) \partial_j^i, \qquad G_j^i(x) \in \mathbf{Pol}_0^N$$

It is easy to verify that if $X \in \mathbf{Pol}_k^N$ and $Y \in \mathbf{Pol}_l^M$, then the commutator $[X, Y] \in \mathbf{Pol}_{k+l-1}^{N+M}$. The following *non-degenerate* condition is needed to get the hypoellipticity of $\mathcal{K}$.

Definition 2 (Eckmann and Hairer [25])**.** A family $\{A_i\}_{i=1}^m$ of vector fields in $\mathbb{R}^n$ with $A_i = \sum_{j=1}^n A_{ij}\partial_j$ is called *non-degenerate* if there exist constants N and C such that for all $x \in \mathbb{R}^n$ and every $v \in \mathbb{R}^n$ one has bound

$$\|v\|^2 \leq C(1+\|x\|^2)^N \sum_{i=1}^m \langle A_i(x), v\rangle^2$$

with $\langle A_i(x), v\rangle = \sum_{j=1}^n A_{i,j}(x)v_j$.

The Lie algebra generated by the vector field X_i fulfilling the non-degenerate condition leads to the hypoellipticity of $\mathcal{K}$.

Proposition 1 (Eckmann and Hairer [25])**.** *If the vector field $\{X_i\}_{i=0}^m$ satisfies*

1. *The vector field X_j with $j = 0, \cdots, m$ belong to $\boldsymbol{Pol}_1^N$ and the function f belongs to $\boldsymbol{Pol}_0^N$.*

2. *There exits a finite number M such that the family of vector fields consisting of $\{X_i\}_{i=1}^m$, $\{[X_i, X_j]\}_{i,j=0}^m$, $\{[X_i,[X_j,X_k]]\}_{i,j,k=0}^m$ and so on up to the commutators of rank M is non-degenerate.*

then the corresponding operator and $\mathcal{K} = \sum_{i=1}^m X_i^T X_i + X_0 + f$ and $\partial_t + \mathcal{K}$ are hypoelliptic operator (see definition in [26, 43]).

Condition 1 and 2 are called the *poly-Hörmander condition.* The hypoellipticity of the operator $\partial_t + \mathcal{K}$ guaranteed the smoothness of the transitional probability density associated with the Markovian semigroup $\mathcal{M}(t,0)$ [24]. Moreover, the developed Hörmander analysis enables us to get a rather detailed estimate for the spectrum of $\mathcal{K}$. Before we get into this part, we first review some properties of the operator $\mathcal{K}$. As a differential operator with C^∞, tempered (i.e. with all its derivatives polynomially bounded) coefficients, $\mathcal{K}$ and its formal adjoint $\mathcal{K}^*$ are well-defined operators with domain $\mathscr{S}(\mathbb{R}^n)$, where $\mathscr{S}(\mathbb{R}^n)$ is the Schwarz space which is dense in $L^p(\mathbb{R}^n), 1 \leq p < \infty$. On the other hand, $\mathcal{K}$ and $\mathcal{K}^*$ are both closable operator on $\mathscr{S}(\mathbb{R}^n)$, hence we always make a formal calculus in $L^p(\mathbb{R}^n)$, i.e. every estimate obtained in this section is supposed to hold in $\mathscr{S}(\mathbb{R}^n)$. The actual operator are the closure of the operators defined in $\mathscr{S}(\mathbb{R}^n)$. We now introduce a family of weighted Sobolev space $\mathcal{S}^{\alpha,\beta}$, with $\alpha, \beta \in \mathbb{R}$. $\mathcal{S}^{\alpha,\beta}$ is defined on the space of tempered distribution $\mathscr{S}'(\mathbb{R}^n)$ as

$$\mathcal{S}^{\alpha,\beta} = \{u \in \mathscr{S}'(\mathbb{R}^n) | \Lambda^\alpha \bar{\Lambda}^\beta u \in L^2(\mathbb{R}^n)\}.$$

where operator $\Lambda^2 = 1 - \sum_{i=1}^n \partial_i = 1 - \Delta$, $\bar{\Lambda}^2 = 1 + \|x\|^2$. For $\alpha \in \mathbb{R}$, Λ^α is interpreted as some pseudo-differential operator (see its definition and properties in [26, 25, 39]). $\mathcal{S}^{\alpha,\beta}$ is equipped with the scalar product

$$\langle f, g \rangle_{\alpha,\beta} = \langle \Lambda^\alpha \bar{\Lambda}^\beta f, \Lambda^\alpha \bar{\Lambda}^\beta g \rangle_{L^2}$$

and the corresponding Sobolev norm $\|\cdot\|_{\alpha,\beta}$. With the above definitions, Eckmann and Hairer [25] were enable to get the following important estimates on the operator $\mathcal{K}$

Theorem 4.1.1 (Eckmann and Hairer [25]). *Let $\mathcal{K} \in \boldsymbol{Pol}_2^N$ is of the form (4.1) and satisfies the poly-Hörmander condition. Moreover, the closure of $\mathcal{K}$ is a maximal-accretive operator in $L^2(\mathbb{R}^n)$ and for every $\epsilon > 0$, there exit constants $\delta > 0$ and $C > 0$ such that*

$$\|u\|_{\delta,\delta} \leq C(\|u\|_{0,\epsilon} + \|\mathcal{K}u\|) \tag{4.2}$$

for every $u \in \mathscr{S}(\mathbb{R}^n)$, Moreover, if there exist constants $\delta > 0$ and $C > 0$ such *that*

$$\|u\|_{0,\epsilon} \leq C(\|u\| + \|\mathcal{K}_y u\|), \tag{4.3}$$

then $\mathcal{K}$ has compact resolvent when considered as an operator acting on $L^2(\mathbb{R}^n)$, whose spectrum $\sigma(\mathcal{K})$ (in $L^2(\mathbb{R}^n)$) is contained in the cusp (see Figure 4.1):

$$\mathcal{S}_{\mathcal{K}} = \{z \in \mathbb{C} |\, \mathrm{Re}\, z \geq 0, |\, \mathrm{Im}\, z| < (8C_1)^{M/2}(1 + \mathrm{Re}\, z)^M\} \tag{4.4}$$

for some positive constant C_1 and integer M.

Theorem 4.1.1 was proved in [26] using a modified version of the standard Hörmander's analysis. In particular, the following estimate was used to get the cusp shape of the spectrum:

$$\frac{1}{4}|z+1|^{2/M}\|u\|^2 \leq C_1([1 + \mathrm{Re}\, z]^2\|u\|^2 + \|(\mathcal{K} - z)u\|^2), \qquad \forall\, \mathrm{Re}\, z \geq 0 \tag{4.5}$$

Estimate (4.5) is the key result of the whole analysis and its importance cannot be overemphasized. For a specific $\mathcal{K}$ corresponding to the Langevin dynamics. Hèrau, Nier and Helffer [39, 38] used (4.5) to get the exponential convergence of the semigroup $e^{-t\mathcal{K}}$ for any initial condition $u_0 \in L^2(\mathbb{R}^n)$. This can be done for generally defined $\mathcal{K}$ (4.1) by following the exact same procedure in [39, 38]. We do not claim any originality of the following theorem, but for the sake completeness, we still list the details of proof here.

Theorem 4.1.2. *We assume that $\mathcal{K}$ satisifies all the conditions listed in Theorem 4.1.1. If the spectrum of $\mathcal{K}$ in $L^2(\mathbb{R}^n)$ satisfies*

$$\sigma(\mathcal{K}) \cap i\mathbb{R} = \sigma_{dis}(\mathcal{K}) \cap i\mathbb{R} \subset \{0\}, \tag{4.6}$$

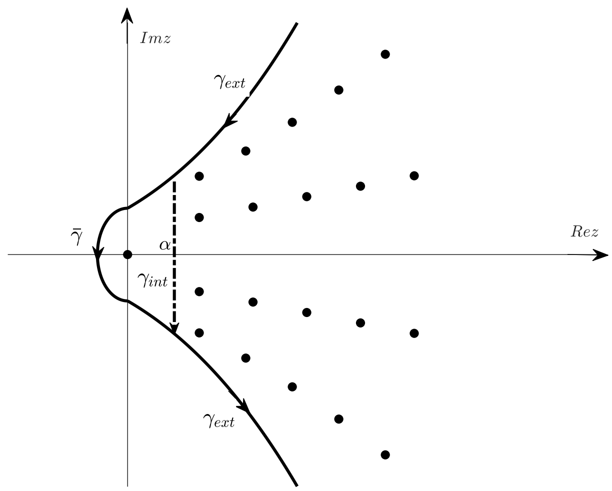

Figure 4.1: Cusp containing the spectrum of $\mathcal{K}$ and $\mathcal{QKQ}$.

and the possible eigenvalue 0 *has finite algebraic multiplicity, then for any* $0 < \alpha < \min(\operatorname{Re}\sigma(\mathcal{K})/\{0\})$*, there exits a positive constant* $C = C(\alpha)$ *such that the estimate*

$$\|e^{-t\mathcal{K}}u_0 - \sqcap_0 u_0\| \leq Ce^{-\alpha t}\|u_0\| \tag{4.7}$$

holds for all $u_0 \in L^2(\mathbb{R}^n)$ *and for all* $t > 0$*, where* $\sqcap_0$ *is the spectral projection on the kernel of* $\mathcal{K}$*.*

Proof. First we note that for closed, maximal-accretive operator $\mathcal{K}$ with densely defined domain, the Lumer-Philips indicates operator $\mathcal{K}$ generates a contraction semigroup $e^{-t\mathcal{K}}$ in $L^2(\mathbb{R}^n)$. The Schwarz space $\mathscr{S}(\mathbb{R}^n)$ is proven [25, 39] to be the core of the closed operator $\mathcal{K}$. Hence by a simple approximation argument, one can show that the hypoelliptic estimate (4.5) holds for any $u \in L^2(\mathbb{R}^n)$.

According to Theorem 4.1.1, $\mathcal{K}$ only has discrete spectrum hence $\sigma(\mathcal{K}) = \sigma_{disk}(\mathcal{K})$. (4.6) requires that for $\mathcal{K}$ in $L^2(\mathbb{R}^n)$, there is no spectrum on $i\mathbb{R}$ except the possible eigenvalue 0 which is isolated and with finite algebraic multiplicity. With this condition, according to Theorem 6.1 in [38], we can get a weakly convergent (in the sense that $\langle e^{-t\mathcal{K}}u_0, \phi\rangle$ for $u_0 \in L^2(\mathbb{R}^n)$ and $\phi \in D(\mathcal{K}^*)$) Dunford integral representation of the semigroup $e^{-t\mathcal{K}}$, which is given by

$$e^{-t\mathcal{K}}u_0 - \sqcap_0 u_0 = \frac{1}{2\pi i}\int_{\partial\mathcal{S}'_{\mathcal{K}}} e^{-tz}(z-\mathcal{K})^{-1}u_0 dz, \tag{4.8}$$

where $\partial\mathcal{S}'_{\mathcal{K}} = \gamma_{int} \cup \gamma_{ext}$ (see Figure 4.1), $(z-\mathcal{K})^{-1}$ is the resolvent operator of $\mathcal{K}$. This formula allows us to transfer the semigroup estimation problem into the estimate of a complex integral. To proceed, we need get an upper bound for the resolvent $(z-\mathcal{K})^{-1}$. According to the definition of $\mathcal{S}_{\mathcal{K}}$, for all $z \notin \mathcal{S}_{\mathcal{K}}$, $\operatorname{Re} z \geq 0$, we have $|z+1|^{2/M} \geq (8C_1)(1+\operatorname{Re} z)^2$. Plugging this formula in (4.5), we get for all $u \in L^2(\mathbb{R}^n)$ that

$$\begin{aligned}&\frac{1}{4}|z+1|^{2/M}\|u\|^2 \leq C_1(\frac{1}{8C_1}|z+1|^{2/M}\|u\|^2 + \|(\mathcal{K}-z)u\|^2), \qquad \forall \operatorname{Re} z \geq 0, z \notin \mathcal{S}_{\mathcal{K}}\\ &\frac{1}{8}|z+1|^{2/M}\|u\|^2 \leq C_1\|(\mathcal{K}-z)u\|^2.\end{aligned}$$

Hence we conclude that $\|(\mathcal{K}-z)^{-1}\| \leq \sqrt{8C_1}|z+1|^{-1/M}$. We now rewrite Dunford integral (4.8) as

$$\begin{aligned}\frac{1}{2\pi i}\int_{\partial\mathcal{S}'_{\mathcal{K}}} e^{-tz}(z-\mathcal{K})^{-1}u_0 dz &= \frac{1}{2\pi i}\int_{\gamma_{int}} e^{-tz}(z-\mathcal{K})^{-1}u_0 dz\\ &\quad + \frac{1}{2\pi i}\int_{\gamma_{ext}} e^{-tz}(z-\mathcal{K})^{-1}u_0 dz.\end{aligned}$$

Since $(K-z)^{-1}$ is a compact operator, for any $0 < \alpha < \min(\operatorname{Re}\sigma(\mathcal{K})/\{0\})$ there exits a constant $C_\alpha > 0$ such that $\|(\mathcal{K}-\alpha)u\| \geq C_\alpha\|u\|$. On the other hand, since

$\mathcal{K}$ is a real operator, for all $z = \alpha + i\nu \notin \sigma(\mathcal{K})$, $\nu \in \mathbb{R}$, we have

$$\begin{aligned} &\|(\mathcal{K} - \alpha + i\nu)u\|^2 = \|(\mathcal{K} - \alpha)u\|^2 + \nu^2\|u\|^2 \geq (C_\alpha^2 + \nu^2)\|u\|^2 \\ \Rightarrow &\|(\mathcal{K} - \alpha + i\nu)^{-1}u\| \leq \frac{1}{\sqrt{C_\alpha^2 + \nu^2}}\|u\|. \end{aligned} \tag{4.9}$$

This indicates in γ_{int} the resolvent $(\mathcal{K} - z)^{-1}$ is uniformally bounded by $1/C_\alpha$. Hence for the first term in the right hand side, we have

$$\left\|\frac{1}{2\pi i}\int_{\gamma_{int}} e^{-tz}(z - \mathcal{K})^{-1}u_0 dz\right\| \leq Ce^{-\alpha t}\|u_0\|. \tag{4.10}$$

For the region γ_{ext}, we have for $z = x + iy \in \gamma_{ext}$, $|\operatorname{Im} z| = (8C_1)^{M/2}(1 + \operatorname{Re} z)^M$. On the other hand, for $z \notin \mathcal{S}_\mathcal{K}$, the resolvent is bounded by $\|(\mathcal{K} - z)^{-1}\| \leq \sqrt{8C_1}|z + 1|^{-1/M}$. With these two estimates, for both $M = 1$ and $M \geq 2$, we can get that the second integral can be bounded as

$$\begin{aligned} &\left\|\frac{1}{2\pi i}\int_{\gamma_{ext}} e^{-tz}(z - \mathcal{K})^{-1}u_0 dz\right\| \\ &\leq C\|u_0\| \int_{(8C_1)^{M/2}(1+\alpha)^M}^{+\infty} \exp\{-t[(8C_1)^{-\frac{1}{2}} y^{\frac{1}{M}} - 1]\} y^{-\frac{1}{M}} dy \\ &= C\|u_0\| \int_\alpha^{+\infty} e^{-th}(h + 1)^{M-2} dh \\ &\leq Ce^{-\alpha t}\|u_0\| \left[\frac{1}{t} + \frac{1}{t^2} + \cdots \frac{1}{t^M}\right], \qquad t > 0. \end{aligned} \tag{4.11}$$

On the other hand, since $e^{-t\mathcal{K}}$ is a dissipative semigroup and $\sqcap_0$ is a projection operator into the kernel of $\mathcal{K}$, $\|e^{-t\mathcal{K}}u_0 - \sqcap_0 u_0\| = \|e^{-t\mathcal{K}}(u_0 - \sqcap_0 u_0)\| \leq \|u_0 - \sqcap_0 u_0\|$. Combining with estimate (4.10) and (4.11), one can find a constant $C = C(\alpha)$ such that

$$\|e^{-t\mathcal{K}}u_0 - \sqcap_0 u_0\| \leq Ce^{-\alpha t}\|u_0\|. \tag{4.12}$$

$\square$

Another direct consequence of the estimate (4.5) is the convergence of the formal power series expansion of the semigroup $e^{-t\mathcal{K}}$. This is an important feature for stochastic system with generator $\mathcal{K}$. It is still unclear whether the same result holds for classical systems [103].

Corollary 4.1.2.1. *Assuming that $\mathcal{K}$ satisfies all the conditions listed in Theorem 4.1.2, then for any observable $u = u(x(t))$ with initial condition $u_0 \in L^2(\mathbb{R}^n)$, the formal power series expansion of the semigroup*

$$e^{-(t+s)\mathcal{K}}u_0 = \sum_{n=0}^{+\infty} \frac{(-s)^n}{n!} e^{-t\mathcal{K}}\mathcal{K}^n u_0 \tag{4.13}$$

converges in operator norm for any $t > 0$, $s > t$.

Proof. For closed linear operator $\mathcal{K}$, using the resolvent equation $(z-\mathcal{K})^{-1}\mathcal{K} = z(z-\mathcal{K})^{-1} - \mathcal{I}$, Cauchy's integral theorem and Dunford integral representation (4.8), we get

$$\begin{aligned} e^{-t\mathcal{K}}\mathcal{K}u_0 - \sqcap_0(\mathcal{K}u_0) &= \frac{1}{2\pi i}\int_{\partial\mathcal{S}'_{\mathcal{K}}} e^{-tz}(z-\mathcal{K})^{-1}\mathcal{K}u_0 dz &(4.14)\\ &= \frac{1}{2\pi i}\int_{\partial\mathcal{S}'_{\mathcal{K}}} ze^{-tz}(z-\mathcal{K})^{-1}u_0 dz, &(4.15)\end{aligned}$$

which holds for any $t > 0$. When z locates in γ_{int}, $|z| \leq (\alpha^2 + (8C_1)^M(1+\alpha)^{2M})^{1/2}$, Combined with the uniform boundedness of the resolvent (4.9), we have the estimate for the first integral

$$\left\|\frac{1}{2\pi i}\int_{\gamma_{int}} e^{-tz}z(z-\mathcal{K})^{-1}u_0 dz\right\| \leq Ce^{-\alpha t}\|u_0\| \tag{4.16}$$

For the second integral in γ_{ext}, $|z| = \sqrt{x^2+y^2} = \sqrt{[(8C_1)^{-1/2}|y|^{1/M}-1]^2+y^2}$. To get the convergence estimate of the complex integral, we only need to consider the case when y is large enough such that $|y|^{2/M} < |y|^2$. [1] This leads to $|z| \le C|y|$. Plugging in this estimate in the second complex integral, we get

$$\begin{aligned}
&\left\| \frac{1}{2\pi i} \int_{\gamma_{ext}} z^n e^{-tz} (z-\mathcal{K})^{-1} u_0 dz \right\| \\
&\le C\|u_0\| \int_{(8C_1)^{M/2}(1+\alpha)^M}^{+\infty} \exp\{-t[(8C_1)^{-\frac{1}{2}} y^{\frac{1}{M}} - 1]\} y^{1-\frac{1}{M}} dy \\
&\le C\|u_0\| \int_{\alpha}^{+\infty} e^{-th}(h+1)^{2M-2} dh \\
&\le Ce^{-\alpha t}\|u_0\| \left[\frac{1}{t} + \frac{1}{t^2} + \cdots \frac{1}{t^{2M-1}} \right] = \mathcal{B}(t), \qquad t > 0. \qquad (4.17)
\end{aligned}$$

Combining (4.16) and (4.17), we can get that the Dunford integral (4.14) is bounded by $\mathcal{B}(t)$. Using triangle inequality, for any fixed $t > 0$, $n \in \mathbb{N}^+$, we have

$$\left\| e^{-\frac{t}{n}\mathcal{K}} \mathcal{K} u_0 \right\| - |\sqcap_0 (\mathcal{K}u_0)| \le \left\| e^{-\frac{t}{n}\mathcal{K}} \mathcal{K} u_0 - \sqcap_0(\mathcal{K}u_0) \right\| \le \mathcal{B}\left(\frac{t}{n}\right).$$

Using operator identity $e^{-t\mathcal{K}}\mathcal{K}^n = (e^{-t\mathcal{K}/n}\mathcal{K})^n$ we will get

$$\|e^{-t\mathcal{K}}\mathcal{K}^n u_0\| \le \left\| e^{-\frac{t}{n}\mathcal{K}} \mathcal{K} u_0 \right\|^n \le \left(\mathcal{B}\left(\frac{t}{n}\right) + |\sqcap_0 (\mathcal{K}u_0)| \right)^n.$$

According to the definition of $\mathcal{B}(t)$, one can verify that for any fixed $t > 0$,

$$\lim_{n\to+\infty} \frac{1}{n!} \left(\mathcal{B}\left(\frac{t}{n}\right) + |\sqcap_0 (\mathcal{K}u_0)| \right)^n = 0. \qquad (4.18)$$

Using the limit (4.18) and the Lagrangian remainder representation (see e.g. Page 104 [28]) for the residual of the truncated power series, one can get (4.13) converges in operator norm for any $t > 0$, $s > t$. □

[1] Otherwise since $|z|$ and the resolvent of $\mathcal{K}$ are uniformly bounded in any finite length segment of γ_{ext}, the second complex integral within this segment can be shown to be bounded as (4.16)

4.1.2 Estimate of $\mathcal{QKQ}$

In this section, we aim to get the hypoelliptic estimate of operator $\mathcal{QKQ}$. $\mathcal{QKQ}$ appears in the EMZE/GLE as the generator of the orthogonal propagator $e^{-t\mathcal{QKQ}}$. In principle, projection operator $\mathcal{P}$ can be chosen in an arbitrary way to derive the corresponding EMZE, we refer readers to [103, 16] for a detailed exploration. Here we are mainly concerned with compact, self-adjoint projection operator $\mathcal{P}$ in Hilbert space L^2. Mori's projection is one of the examples. With such $\mathcal{P}$, we can get the cusp shape spectral estimate of $\mathcal{QKQ}$.

Theorem 4.1.3. *Assuming that $\mathcal{K}$ satisfies all the conditions listed in Theorem 4.1.1, then for compact, self-adjoint projection operator $\mathcal{P} : L^2(\mathbb{R}^n) \to L^2(\mathbb{R}^n)$, the operator $\mathcal{QKQ}$ has compact resolvent, whose spectrum is contained in the exact same cusp-shape region*

$$\mathcal{S}_{\mathcal{QKQ}} = \{z \in \mathbb{C} | \operatorname{Re} z \geq 0, |\operatorname{Im} z| < (8C_{\mathcal{Q}})^{M_{\mathcal{Q}}/2}(1 + \operatorname{Re} z)^{M_{\mathcal{Q}}}\}$$

for some the positive constants $C_{\mathcal{Q}}$ and integer $M_{\mathcal{Q}}$.

Proof. We first verify that when a closely defined operator $\mathcal{K}$ is maximal-accretive, so does $\mathcal{QKQ}$. According to Lumer-Philips theorem [28], the adjoint of a maximal-accretive operator is accretive, therefore

$$\operatorname{Re}\langle \mathcal{K}f, f\rangle \geq 0, \qquad \forall f \in D(\mathcal{K})$$
$$\operatorname{Re}\langle \mathcal{K}^*f, f\rangle \geq 0, \qquad \forall f \in D(\mathcal{K}^*).$$

On the other hand, When $\mathcal{P}$ is a self-adjoint operator in $L^2(\mathbb{R}^n)$, $\mathcal{Q} = \mathcal{I} - \mathcal{P}$ is also a self-adjoint operator and we can get

$$\mathrm{Re}\langle \mathcal{QKQ}f, f\rangle = \mathrm{Re}\langle \mathcal{KQ}f, \mathcal{Q}f\rangle \geq 0, \qquad \forall f \in D(\mathcal{K})$$
$$\mathrm{Re}\langle (\mathcal{QKQ})^* f, f\rangle = \mathrm{Re}\langle \mathcal{K}^*\mathcal{Q}f, \mathcal{Q}f\rangle \geq 0, \qquad \forall f \in D(\mathcal{K}^*).$$

Hence $\mathcal{QKQ}$ and its adjoint operator $\mathcal{QK}^*\mathcal{Q}$ are both accretive. $\mathcal{QKQ}$ is also a closable operator defined in $D(\mathcal{K})$. This can be seen if we decompose $\mathcal{QKQ}$ as $\mathcal{QKQ} = \mathcal{K} - \mathcal{KP} - \mathcal{PKQ}$. When $\mathcal{K}$ is closed (using the formal calculus we mentioned before), $\mathcal{QKQ}$ is also a closed operator since $\mathcal{KP}$ and $\mathcal{PKQ}$ are both bounded operator (this will be proved immediately) (see [48]). Using the Lumer-Philips theorem again, we shall get that $\mathcal{QKQ}$ is also maximal-accertive, whose closure generates a contraction semigroup $e^{-t\mathcal{QKQ}}$ in $L^2(\mathbb{R}^n)$.

For the second step, we verify that when $\mathcal{K}$ satisfies the hypoelliptic estimate $\|u\|_{\delta,\delta} \leq C(\|u\| + \|\mathcal{K}u\|)$, so does $\mathcal{QKQ}$. i.e.

$$\|u\|_{\delta,\delta} \leq C(\|u\| + \|\mathcal{QKQ}u\|). \tag{4.19}$$

Using triangle inequality, we have

$$\|u\|_{\delta,\delta} \leq C(\|u\| + \|\mathcal{K}u\|) \leq C(\|u\| + \|\mathcal{KP}u\| + \|\mathcal{QKQ}u\| + \|\mathcal{PKQ}u\|)$$

To get (4.19), it is sufficient to show that $\mathcal{KP}$ and $\mathcal{PKQ}$ are bounded operator in $L^2(\mathbb{R}^n)$. Since the compact, self-adjoint operator is necessarily finite rank [42], $\mathcal{P}$ admits a canonical form (see [42]) $\mathcal{P}(\cdot) = \sum_{i=1}^n \lambda_i \langle (\cdot), \phi_i\rangle \varphi_i$, where $\{\phi_i\}_{i=1}^n$ and

$\{\varphi_i\}_{i=1}^n$ are linearly independent vector set in $L^2(\mathbb{R}^n)$. With this, we have

$$\begin{aligned}
\|\mathcal{K}\mathcal{P}u\| &= \left\|\sum_{i=1}^n \lambda_i\langle u,\phi_i\rangle\mathcal{K}\varphi_i\right\| \leq \sum_{i=1}^n |\lambda_i|\|\mathcal{K}\varphi_i\|\|\phi_i\|\|u\| \leq C\|u\| \\
\|\mathcal{P}\mathcal{K}\mathcal{Q}u\| &= \left\|\sum_{i=1}^n \lambda_i\langle \mathcal{K}\mathcal{Q}u,\phi_i\rangle\varphi_i\right\| = \left\|\sum_{i=1}^n \lambda_i\langle u,\mathcal{Q}\mathcal{K}^*\phi_i\rangle\varphi_i\right\| \\
&\leq \sum_{i=1}^n |\lambda_i|\|\mathcal{Q}\mathcal{K}^*\phi_i\|\|\varphi_i\|\|u\| \leq C\|u\|
\end{aligned}$$

On the other hand, since $\mathcal{Q}\mathcal{K}\mathcal{Q}$ is accretive which makes $(\mathcal{Q}\mathcal{K}\mathcal{Q}+1)$ invertible, we have

$$\begin{aligned}
&\|u\|_{\delta,\delta} \leq C(\|u\| + \|\mathcal{Q}\mathcal{K}\mathcal{Q}u\|) \leq \sqrt{2}C\|(\mathcal{Q}\mathcal{K}\mathcal{Q}+1)u\| \\
\Rightarrow &\|(\mathcal{Q}\mathcal{K}\mathcal{Q}+1)^{-1}u\|_{\delta,\delta} \leq \sqrt{2}C\|u\|
\end{aligned}$$

This indicates the operator $(\mathcal{Q}\mathcal{K}\mathcal{Q}+1)^{-1}$ is bounded operator from L^2 to $\mathcal{S}^{\delta,\delta}$. On the other hand, since $S^{\delta,\delta}$ is compactly embedded into L^2 (Lemma 3.2 [25]). The operator $(\mathcal{Q}\mathcal{K}\mathcal{Q}+1)^{-1}$ is compact therefore $\mathcal{Q}\mathcal{K}\mathcal{Q}$ has compact resolvent [48].

To prove that the discrete spectrum of $\mathcal{Q}\mathcal{K}\mathcal{Q}$ is contained in the cusp-like region $\mathcal{S}_{\mathcal{Q}\mathcal{K}\mathcal{Q}}$, we simply follows the procedure in [25]. For $\mathcal{K} \in \mathbf{Pol}_2^N$, one have for $\alpha = \max\{2, N\}$ the bound

$$\|(\mathcal{K}+1)u\| \leq C\|u\|_{\delta,\delta}, \qquad \forall u \in \mathscr{S}_n$$

then for all $u \in \mathscr{S}_n$, we also have

$$\begin{aligned}
\|(\mathcal{Q}\mathcal{K}\mathcal{Q}+1)u\| &\leq \|\mathcal{Q}\|(\|\mathcal{K}u\| + \|\mathcal{K}\mathcal{P}u\|) + \|u\| \\
&\leq C(\|\mathcal{K}u\| + \|u\|) \leq \sqrt{2}C\|(\mathcal{K}+1)u\| \leq C\|u\|_{\delta,\delta}
\end{aligned}$$

Recall that $\mathcal{Q}\mathcal{K}\mathcal{Q} : D(\mathcal{Q}\mathcal{K}\mathcal{Q}) \to L^2(\mathbb{R}^{2d})$ is a maximal accretive operator, by

Lemma 4.5 in [25], one can find for every $\delta > 0$ and integer $M_{\mathcal{Q}} > 0$ and a constant C such that

$$\langle u, [(\mathcal{QKQ}+1)^*(\mathcal{QKQ}+1)]^{1/M_{\mathcal{Q}}}\rangle \leq C\|u\|^2_{\delta,\delta}. \tag{4.20}$$

Using the hypoelliptic estimate (4.20), (4.19), the Prop.B.1 in [38] and the triangle inequality for $z = \operatorname{Re} z + i \operatorname{Im} z$, we get

$$\begin{aligned}\frac{1}{2}|z+1|^{2/M_{\mathcal{Q}}}\|u\|^2 &\leq C\|u\|^2_{\delta,\delta} + \|(\mathcal{QKQ}-z)\|^2 \\ &\leq C_{\mathcal{Q}}([1+\operatorname{Re} z]^2\|u\|^2 + \|(\mathcal{QKQ}-z)u\|^2).\end{aligned}$$

Together with the compactness of the resolvent of $\mathcal{QKQ}$, this implies that every z in the spectrum of $\mathcal{QKQ}$ satisfies inequality

$$\frac{1}{4}|\operatorname{Im} z|^{2/M_{\mathcal{Q}}}\|u\|^2 \leq \frac{1}{4}|z+1|^{2/M_{\mathcal{Q}}}\|u\|^2 \leq C_{\mathcal{Q}}(1+\operatorname{Re} z)^2\|u\|^2,$$

therefore the spectrum of $\mathcal{QKQ}$ is contained in the exact same cusp-shape region $\mathcal{S}_{\mathcal{QKQ}}$. Moreover, for $z \notin \mathcal{S}_{\mathcal{QKQ}}$, we have resolvent estimate

$$\|(\mathcal{QKQ}-z)^{-1}\| \leq \sqrt{8C_{\mathcal{Q}}}|z+1|^{-1/M_{\mathcal{Q}}}. \tag{4.21}$$

□

Remark. The central argument used to prove Theorem 4.1.3 is that when $\mathcal{P}$ is compact, self-adjoint projection operator, the hypoelliptic estimate (4.2) holds for $\mathcal{QKQ}$ since both $\mathcal{KP}$ and $\mathcal{PKQ}$ are bounded operators. We remark that such approach does not apply to non-compact, self-adjoint projection operators simply because they are infinite-rank. A typical example is the conditional expectation, also known as the Zwanzig type projection operator that frequently used in the Mori-Zwanzig framework [103, 17, 18, 110], the analysis of $\mathcal{QKQ}$ corresponding to infinite rank projection operator would be very different, and probably much harder. We leave it as an open problem awaiting further investigations.

With the resolvent estimate (4.21) for $\mathcal{QKQ}$, by following the exact same procedure in the proof of Theorem 4.1.2 and Corollary 4.1.2.1, we shall get the semigroup estimate for $e^{-t\mathcal{QKQ}}$ and the convergence of its power series expansion.

Theorem 4.1.4. *We assume that $\mathcal{K}$ satisfies all the conditions listed in Theorem 4.1.1. If the spectrum of $\mathcal{QKQ}$ in L^2 satisfies*

$$\sigma(\mathcal{QKQ}) \cap i\mathbb{R} = \sigma_{dis}(\mathcal{QKQ}) \cap i\mathbb{R} \subset \{0\}. \tag{4.22}$$

with the possible the eigenvalue 0 has finite algebraic multiplicity. Then for any $0 < \alpha_{\mathcal{Q}} < \min(\mathrm{Re}\,\sigma(\mathcal{QKQ})/\{0\})$, there exits a positive constant $C = C(\alpha_{\mathcal{Q}})$ such that the estimate

$$\|e^{-t\mathcal{QKQ}}u_0 - \sqcap_0^{\mathcal{Q}} u_0\| \leq Ce^{-\alpha_{\mathcal{Q}} t}\|u_0\| \tag{4.23}$$

holds for all $u_0 \in L^2$ and for all $t > 0$, where $\sqcap_0^{\mathcal{Q}}$ is the spectral projection on the kernel of $\mathcal{QKQ}$.

Corollary 4.1.4.1. *Assume that $\mathcal{K}$ satisfies all the conditions listed in Theorem 4.1.2. Then for any initial $u_0 \in L^2(\mathbb{R}^n)$, the power series expansion of the semigroup*

$$e^{-(t+s)\mathcal{QKQ}}u_0 = \sum_{n=0}^{+\infty} \frac{(-s)^n}{n!} e^{-t\mathcal{QKQ}}(\mathcal{QKQ})^n u_0 \tag{4.24}$$

converges in operator norm for any $t > 0$, $s > t$.

The convergence result for power series expansion of the semigroup $e^{-t\mathcal{K}}$ and $e^{-t\mathcal{QKQ}}$ is rarely seen in common semigroup perturbation theory since it holds for all $t > 0$ and does not depend on any small paramters such as the time and the magnititude of nonlinearity. An immediate consequence of this result is that any first-principle calculation method based on the power series expansion of the semigroup is convergent when expansion order $N \to +\infty$. Hence the combinatorial method that would be proposed in Chapter 6 can be proven to be convergent when applied to stochastic system (2.21) as long as $\mathcal{P}$ and $\mathcal{K}$ satisfy the required conditions. From estimate (4.23), immediately we deduce the exponentially decaying of the EMZE/GLE memory kernel as claimed in the abstract.

Corollary 4.1.4.2. *For any scalar observable function $u = u(x(t))$ with the initial condition $u(0) = u_0$, the memory kernel of the Mori-type EMZE/GLE exponentially decays to constant $\langle \mathcal{QK}^* u_0, \sqcap_0^{\mathcal{Q}}(\mathcal{K}u_0)\rangle$ with rate $\alpha_{\mathcal{Q}}$, i.e. there exists positive constants C such that*

$$|K(t) - \langle \mathcal{QK}^* u_0, \sqcap_0^{\mathcal{Q}}(\mathcal{K}u_0)\rangle| \leq Ce^{-\alpha_{\mathcal{Q}} t}. \tag{4.25}$$

The same estimate holds for the entries of the memory matrix (2.35c) *for vector form EMZE/GLE.*

Proof. It is sufficient to consider the one-dimensional GLE with a scalar memory kernel $K(t)$. According to the definition (2.35c), we have

$$\begin{aligned} |K(t) - \langle \mathcal{QK}^* u_0, \sqcap_0^{\mathcal{Q}}(\mathcal{K}u_0)\rangle| &= |\langle u_0, \mathcal{K}e^{t\mathcal{QKQ}}\mathcal{QK}u_0\rangle - \langle \mathcal{QK}^* u_0, \sqcap_0^{\mathcal{Q}}(\mathcal{K}u_0)\rangle| \\ &= |\langle \mathcal{QK}^* u_0, e^{t\mathcal{QKQ}}\mathcal{K}u_0\rangle - \langle \mathcal{QK}^* u_0, \sqcap_0^{\mathcal{Q}}(\mathcal{K}u_0)\rangle| \end{aligned} \tag{4.26}$$

Since $\sqcap_0^{\mathcal{Q}}$ is an projection operator, $\sqcap_0^{\mathcal{Q}}(\mathcal{K}u_0 - \sqcap_0^{\mathcal{Q}}(\mathcal{K}u_0)) = 0$. Using estimate (4.23),(4.26) and Cauchy-Schwartz inequality, we get

$$|K(t) - \langle \mathcal{Q}\mathcal{K}^* u_0, \sqcap_0^{\mathcal{Q}}(\mathcal{K}u_0)\rangle| \leq C\|\mathcal{Q}\mathcal{K}^* u_0\|\|\mathcal{K}u_0 - \sqcap_0^{\mathcal{Q}}(\mathcal{K}u_0)\|e^{-\alpha_{\mathcal{Q}}t}. \tag{4.27}$$

Using the formal expression of the memory matrix (2.35c), one can easily get that similar results hold entry-wise for vector form EMZE/GLE. □

4.2 Applications

All the results we obtained in Section 4.1 are based on pure functional analysis. In this section, we provide examples of operator $\mathcal{K}$ and $\mathcal{Q}\mathcal{K}\mathcal{Q}$ arising from Langevin dynamics, and then get the exact exponentially decaying estimate for corresponding semigroups and the memory kernel. Specifically, we will focus our discussion on how to determine the spectral projector $\sqcap_0^{\mathcal{Q}}$ and the kernel of $\mathcal{Q}\mathcal{K}\mathcal{Q}$, which turns out to be a more tricky question than the common case for operator $\mathcal{K}$.

4.2.1 Application to Langevin dynamics

The Langevin dynamics is a system of SDEs in $\mathbb{R}^{2d}_{p,q}$ with the equation of motion:

$$\begin{cases} dq = \dfrac{1}{m} p dt \\ dp = -\nabla V(q)dt - \dfrac{\gamma}{m} p dt + \sigma d\mathcal{W}_t, \end{cases} \tag{4.28}$$

where $\mathcal{W}_t$ is a d-dimensional Wiener process. The parameters γ and σ represent the magnitude of the fluctuations and of the dissipation respectively, and are linked by the fluctuation-dissipation relation $\sigma = (2\gamma/\beta)^{1/2}$, where $\beta \propto 1/T$ is the thermodynamic parameter. The Langevin dynamics is frequently used in statistical mechanics to model the mesoscopic scale dynamics of liquids and gases. Another coarse-grained model, the overdamped Langevin equation, can be obtained from the Langevin dynamics (4.28) by letting the mass m go to zero and by setting $\gamma = 1$. Hence the analysis we are going to present holds for the overdamped case in a natural way. For SDE (4.28), The position and momentum $\{q_t, p_t\}$ define a Markov process $\mathcal{M}(t, 0)$ with the generator

$$-\mathcal{K} = \frac{p}{m} \cdot \nabla_q - \nabla_q V(q) \cdot \nabla_p + \gamma(-\frac{p}{m}\nabla_p + \beta^{-1}\Delta_p), \tag{4.29}$$

where $\cdot$ denotes the vector product. For potential energy satisfying $V(q) > 0$ at ∞, the Langevin equation (4.28) admits an unique invariant distribution given by the Gibbs form $\rho_{eq} = e^{-\beta\mathcal{H}(p,q)}/Z$ with $\mathcal{H}(p, q) = \frac{1}{2m}\|p\|^2 + V(q)$ and Z being the partition function. To formulate the estimation problem in a suitable framework, we introduce the standard unitary transformation $\mathcal{U} : L^2(\mathbb{R}^{2d}) \to L^2(\mathbb{R}^{2d}, \rho_{eq})$ given by the formula

$$(\mathcal{U}g)(p, q) = \sqrt{Z}e^{\frac{\beta}{2}\mathcal{H}}g(p, q) \tag{4.30}$$

where $L^2(\rho_{eq})$ is weighted Hilbert space endowed with inner product $\langle f, g\rangle_{eq} = \int fg\rho_{eq}dpdq$. The map $\mathcal{U} : L^2(\mathbb{R}^{2d}) \to L^2(\mathbb{R}^{2d}, \rho_{eq})$ introduces an isomorphism between these two Hilbert spaces, and for any $\tilde{u} \in L^2(\mathbb{R}^{2d})$, there is an unique $u \in L^2(\mathbb{R}^{2d}, \rho_{eq})$ such that $\tilde{u} = (e^{-\beta\mathcal{H}/2}/\sqrt{Z})u$ and

$$\|\tilde{u}\|_{L^2} = \|u\|_{L^2_{eq}}. \tag{4.31}$$

By some simple calculation, one can verify that the transformed Kolmogorov-type operator $\tilde{\mathcal{K}} = \mathcal{U}^{-1}\mathcal{K}\mathcal{U}$ is explicitly given by

$$\tilde{\mathcal{K}} = -\frac{p}{m}\cdot\nabla_q + \nabla V(q)\cdot\nabla_p + \frac{\gamma}{\beta}\left(-\partial_{p_i} + \frac{\beta}{2m}p_i\right)\cdot\left(\partial_{p_i} + \frac{\beta}{2m}p_i\right) \tag{4.32}$$

which can be written in the canonical form $\tilde{K} = \sum_{i=1}^{d} X_i^* X_i - X_0$, with

$$\begin{aligned} X_0 &= \frac{p}{m}\cdot\nabla_q - \nabla V(q)\cdot\nabla_p \\ X_i &= \sqrt{\frac{\gamma}{\beta}}(\partial_{p_i} + \frac{\beta}{2m}p_i) \\ X_i^* &= \sqrt{\frac{\gamma}{\beta}}(-\partial_{p_i} + \frac{\beta}{2m}p_i) \end{aligned} \tag{4.33}$$

where X_0 is skew-adjoint (Liouville) operator in $L^2(\mathbb{R}^{2d})$ satisfying $X_0^* = -X_0$. X_i^* and X_i can be interpreted as the creation and annihilation operator of a quantum harmonic oscillator. Naturally the adjoint operator of $\tilde{\mathcal{K}}$ is given by $\tilde{\mathcal{K}}^* = \sum_{i=1}^{d} X_i^* X_i + X_0$. $\tilde{\mathcal{K}}$ and its formal adjoint $\tilde{\mathcal{K}}^*$ are proven [39, 38, 24] to be accretive, closable operators, with maximally accretive closure in $L^2(\mathbb{R}^n)$. We use the same notation for the closure of $\mathcal{K}$ and $\mathcal{K}^*$. The domain of the closed $\mathcal{K}$, $\mathcal{K}^*$ are given by

$$D(\mathcal{K}) = \{u \in L^2(\mathbb{R}^n), \mathcal{K}u \in L^2(\mathbb{R}^n)\},$$
$$D(\mathcal{K}^*) = \{u \in L^2(\mathbb{R}^n), \mathcal{K}^*u \in L^2(\mathbb{R}^n)\}.$$

The projection operator $\mathcal{P}$ and $\mathcal{Q}$ used in the Mori-Zwanzig formulation can also be similarly transformed into the operators on flat Hilbert space $L^2(\mathbb{R}^{2d})$ as $\tilde{\mathcal{P}}$ and $\tilde{\mathcal{Q}}$. The relationship between $L^2(\mathbb{R}^{2d})$, $L^2(\mathbb{R}^{2d}, \rho_{eq})$ and the operators defined therein can be summarized in the following commutative diagram

$$\begin{array}{ccc} L^2(\mathbb{R}^{2d}) & \xrightarrow{\ \mathcal{U}\ } & L^2(\mathbb{R}^{2d}, \rho_{eq}) \\ {\scriptstyle \tilde{\mathcal{P}}, \tilde{\mathcal{K}}, \tilde{\mathcal{Q}}}\big\downarrow & & \big\downarrow{\scriptstyle \mathcal{P}, \mathcal{K}, \mathcal{Q}} \\ L^2(\mathbb{R}^{2d}) & \xleftarrow[\ \mathcal{U}^{-1}\ ]{} & L^2(\mathbb{R}^{2d}, \rho_{eq}) \end{array} \tag{4.34}$$

The properties of operators defined in these two Hilbert spaces are almost the same due to the unitarity of the transformation operator $\mathcal{U}$. For instance, for compact, self-adjoint projection operator $\mathcal{P}, \mathcal{Q}$ in $L^2(\mathbb{R}^{2d}, \rho_{eq})$, their $L^2(\mathbb{R}^{2d})$ correspondence $\tilde{\mathcal{P}}, \tilde{\mathcal{Q}}$ are also compact, self-adjoint projection operators.

We have seen that the transformed Kolmogorov operator $\tilde{\mathcal{K}}$ is now defined in $L^2(\mathbb{R}^{2d})$ and can be written in the canonical form $\tilde{\mathcal{K}}^* = \sum_{i=1}^{d} X_i^* X_i + X_0$. To use the analytical results in Section 4.1.1 and 4.1.2, one only need to verify for SDE (4.28) whether $\{X_i\}_{i=0}^{d}$ satisfies the poly-Hörmander condition and estimate (4.3). This can be done by imposing some additional, normally rather weak assumptions on the potential energy $V(q)$, and then follow the steps in Prop. 3.7 [24] to get (4.3). To focus more on the analyis of the operator $\mathcal{QKQ}$ we shall not follow this routine since it requires the proof (4.3), instead we use the following results obtained by Hérau, Nier and Helffer in [38, 39] which holds for Langevin dynamics specifically. In accordance with [38], we assume $V(q)$ satisfies the following weak ellipticity hypothesis:

Hypothesis 1. *The potential $V(q)$ belongs to $C^\infty(\mathbb{R}^d)$ and satisfies:*

1. $\forall \alpha \in \mathbb{N}^d, |\alpha| = 1, \forall q \in \mathbb{R}^d$, $|\partial_q^\alpha V(q)| \leq C_\alpha \sqrt{1 + \|\nabla V(q)\|^2}$ *for some constant* $C_\alpha > 0$.

2. $\exists M, C \geq 1, \forall q \in \mathbb{R}^d$, $C^{-1}\langle q \rangle^{1/M} \leq \sqrt{1 + \|\nabla V(q)\|^2} \leq C\langle q \rangle^M$

where we used the multi-index notation ∂_q^α and $\langle q\rangle = (1+\|q\|^2)^{1/2}$.

Hypothesis 1 holds for any potential energy grows as $V(q) \simeq \|q\|^M$ when $q \to \infty$. With this hypothesis, Helffer and Nier proved the following theorem in [38]:

Proposition 2 (Helffer and Nier [38])**.** *For Langevin dynamics* (4.28) *with potential energy function $V(q)$ satisfying Hypothesis 1, operator $\tilde{\mathcal{K}}$ has compact resolvent, with the spectrum bounded by $\mathcal{S}_{\mathcal{K}}$. In addition, there exists a positive constant $C = C(\alpha)$ such that the estimate*

$$\|e^{-t\tilde{\mathcal{K}}}u_0 - \tilde{\sqcap}_0\tilde{u}_0\| \le Ce^{-\alpha t}\|\tilde{u}_0\| \tag{4.35}$$

holds for all $\tilde{u}_0 \in L^2(\mathbb{R}^{2d})$ and for all $t > 0$, where $\tilde{\sqcap}_0$ is the orthogonal projection onto the kernel of $\tilde{\mathcal{K}}$ in $L^2(\mathbb{R}^{2d})$, which is given by $Ker(\tilde{\mathcal{K}}) = \mathbb{R}e^{-\beta\mathcal{H}/2}$.

Proposition 2 can be rewritten in $L^2(\mathbb{R}^{2d}, \rho_{eq})$. With the unitary transformation (4.30), naturally we get the $L^2(\mathbb{R}^{2d}, \rho_{eq})$ equivalence of the estimate (4.35):

$$\|e^{-t\mathcal{K}}u_0 - \sqcap_0 u_0\|_{L^2_{eq}} = \|e^{-t\tilde{\mathcal{K}}}\tilde{u_0} - \tilde{\sqcap}_0\tilde{u}_0\|_{L^2} \le Ce^{-\alpha t}\|\tilde{u}_0\|_{L^2} = Ce^{-\alpha t}\|u_0\|_{L^2_{eq}}. \tag{4.36}$$

$\sqcap_0 = \mathcal{U}\tilde{\sqcap}_0\mathcal{U}^{-1}$ is the orthogonal projection onto linear subspace $\mathbb{1}$ with respect to $L^2(\mathbb{R}^{2d}, \rho_{eq})$ norm, which has the simple form $\sqcap_0(\cdot) = \mathbb{E}[(\cdot)]$. For Langevin dynamics, it is possible to get a prior estimate on the convergence rate α by building connections between the Kolmogorov operator and the Witten Laplacian. In [39], Hèrau and Nier specifically discussed the asymptotics of α at the low temperature ($\beta \to \infty$) and the high temperature ($\beta \to 0$) regime. Readers may refer to [39, 38] for more details. Estimate (4.5) was proved for Langevin dynamics in [39] using a different method, from which one can get the convergence of the power series expansion of $e^{-(t+s)\mathcal{K}}$ for any $t > 0$, $s > t$.

Our next task is to get estimate of $\mathcal{QKQ}$. For compact, self-adjoint projection operator $\mathcal{P}$, according to Theorem 4.1.3, the spectrum of $\tilde{\mathcal{Q}}\tilde{\mathcal{K}}\tilde{\mathcal{Q}}$ is bounded by $\mathcal{S}_{\tilde{\mathcal{Q}}\tilde{\mathcal{K}}\tilde{\mathcal{Q}}}$. With Theorem 4.1.4, to get the exponential convergence estimate of $e^{-t\tilde{\mathcal{Q}}\tilde{\mathcal{K}}\tilde{\mathcal{Q}}}$ (or $e^{-t\mathcal{QKQ}}$), we need to verify that operator $\tilde{\mathcal{Q}}\tilde{\mathcal{K}}\tilde{\mathcal{Q}}$ satisfies (4.22). But before that, we need to determine the exact form of the spectral projection $\tilde{\sqcap}_0^{\tilde{\mathcal{Q}}}$. This turns out to be a tricky question for $\tilde{\mathcal{Q}}\tilde{\mathcal{K}}\tilde{\mathcal{Q}}$. As we will see, the kernel of $\tilde{\mathcal{Q}}\tilde{\mathcal{K}}\tilde{\mathcal{Q}}$ is multi-dimensional and depends on the definition of $\mathcal{P}$. We use the following Mori-type projection operator $\mathcal{P}$ and its $L^2(\mathbb{R}^{2d})$ correspondence $\tilde{\mathcal{P}} = \mathcal{U}^{-1}\mathcal{P}\mathcal{U}$ as an example to explicitly calculate the kernel of $\tilde{\mathcal{Q}}\tilde{\mathcal{K}}\tilde{\mathcal{Q}}$

$$\mathcal{P}(\cdot) := \sum_{i=1}^{N}\langle(\cdot), f_i(q)\rangle_{eq} f_i(q) \qquad \tilde{\mathcal{P}}(\cdot) := \sum_{i=1}^{N}\langle(\cdot), f_i(q)\rangle_{eq/2} f_i(q) e^{-\beta\mathcal{H}/2}. \tag{4.37}$$

Here we used the shorthand notation $\langle\cdot\rangle_{eq/2} = \int e^{-\beta\mathcal{H}/2}dpdq/Z$. $f_j(q)$ is an arbitrary scalar function of the position $q \in \mathbb{R}^d$. Moreover, we assume that $\{f_j(q)\}_{j=1}^N$ form an orthonormal basis of *span*$\{f_j(q)\}$ with respect to ρ_{eq} and $\langle f_j(q)\rangle_{eq} = 0$. The assumption holds without loss of generality since when using EMZE/GLE to study the dynamics of observable $\{f_j(q)\}_{i=1}^N$, it is equivalent to study the linearly transformed $\{\hat{f}_j(q)\}_{i=1}^N$ which satisfies aforementioned conditions. For operator $\tilde{\mathcal{Q}}\tilde{\mathcal{K}}\tilde{\mathcal{K}}$, we have the following facts:

Lemma 4.2.1. *For Langevin dynamics* (4.28) *with potential energy function* $V(q)$ *satisfying Hypothesis 1, if we introduce a projection operator* $\tilde{\mathcal{P}}$ *defined as* (4.37), *then*

$$Ker(\tilde{\mathcal{Q}}\tilde{\mathcal{K}}\tilde{\mathcal{Q}}) = Ker(\tilde{\mathcal{K}}) \oplus Id(\tilde{\mathcal{P}}) \tag{4.38}$$

$$\sigma(\tilde{\mathcal{Q}}\tilde{\mathcal{K}}\tilde{\mathcal{Q}}) \cap i\mathbb{R} = \sigma_{dis}(\tilde{\mathcal{Q}}\tilde{\mathcal{K}}\tilde{\mathcal{Q}}) \cap i\mathbb{R} \subset \{0\}. \tag{4.39}$$

where $Id(\tilde{\mathcal{P}})$ *is the finite dimensional linear space defined as* $Id(\tilde{\mathcal{P}}) := \{u \in L^2(\mathbb{R}^{2d}) | \tilde{\mathcal{P}}u = u\}$.

Proof. We first prove (4.38). For $u \in Ker(\tilde{\mathcal{Q}}\tilde{\mathcal{K}}\tilde{\mathcal{Q}})$, $\tilde{\mathcal{Q}}\tilde{\mathcal{K}}\tilde{\mathcal{Q}}u = 0$. u can only be one of the following three cases: i) $\tilde{\mathcal{Q}}u = 0$, ii) $\tilde{\mathcal{Q}}u \neq 0, \tilde{\mathcal{K}}\tilde{\mathcal{Q}}u = 0$ and iii) $\tilde{\mathcal{Q}}u \neq 0, \tilde{\mathcal{K}}\tilde{\mathcal{Q}}u \neq 0, \tilde{\mathcal{Q}}\tilde{\mathcal{K}}\tilde{\mathcal{Q}}u = 0$. Now we discuss these cases separately.

Case 1. When $\tilde{\mathcal{Q}}u = 0$, immediately we have $\tilde{\mathcal{P}}u = u - \tilde{\mathcal{Q}}u = u$, this indicates $u \in Id(\tilde{\mathcal{P}})$.

Case 2. When $\tilde{\mathcal{Q}}u \neq 0$, $\tilde{\mathcal{K}}\tilde{\mathcal{Q}}u = 0$, naturally we get $\tilde{\mathcal{Q}}u \in Ker(\tilde{\mathcal{K}})$. We first prove $\tilde{\mathcal{Q}}u \in Ker(\tilde{\mathcal{K}}) \Rightarrow u \in Id(\tilde{\mathcal{P}}) \oplus Ker(\tilde{\mathcal{K}})$. Since $\tilde{\mathcal{P}}$ is defined as (4.37), which is an orthogonal, finite rank projection in $L^2(\mathbb{R}^n)$, we have $Id(\tilde{\mathcal{P}}) = Ran(\tilde{\mathcal{P}}) = span\{f_i(q)e^{-\beta\mathcal{H}/2}\}_{i=1}^N$. With these two facts, we may rewrite the suitable u as

$$u = \tilde{\mathcal{P}}u + \tilde{\mathcal{Q}}u = \sum_{i=1}^{N} c_i f_i(q)e^{-\beta\mathcal{H}/2} + c'e^{-\beta\mathcal{H}/2} \in Id(\tilde{\mathcal{P}}) \oplus Ker(\tilde{\mathcal{K}}), \qquad c_i, c' \in \mathbb{R}$$

We now verify when $u \in Id(\tilde{\mathcal{P}}) \oplus Ker(\tilde{\mathcal{K}})$, $\tilde{\mathcal{K}}\tilde{\mathcal{Q}}u = 0$, i.e. $\tilde{\mathcal{Q}}u \in Ker(\tilde{\mathcal{K}})$, Since $Id(\tilde{\mathcal{P}}) = Ran(\tilde{\mathcal{P}})$, u admits the general form $u = \sum_{i=1}^N c_i f_i(q)e^{-\beta\mathcal{H}/2} + c'e^{-\beta\mathcal{H}/2}$, $c_i, c' \in \mathbb{R}$, we have

$$\tilde{\mathcal{Q}}u = u - \tilde{\mathcal{P}}u = c'e^{-\beta\mathcal{H}/2} - c'\tilde{\mathcal{P}}e^{-\beta\mathcal{H}/2} = c'e^{-\beta\mathcal{H}/2} \in Ker(\tilde{\mathcal{K}}) \Rightarrow \tilde{\mathcal{K}}\tilde{\mathcal{Q}}u = 0.$$

Case 3. When $\tilde{\mathcal{Q}}u \neq 0$, $\tilde{\mathcal{K}}\tilde{\mathcal{Q}}u \neq 0$ and $\tilde{\mathcal{Q}}\tilde{\mathcal{K}}\tilde{\mathcal{Q}}u = 0$. For u satisfies these conditions, we have

$$\langle \tilde{\mathcal{Q}}\tilde{\mathcal{K}}\tilde{\mathcal{Q}}u, u\rangle = \langle \tilde{\mathcal{K}}\tilde{\mathcal{Q}}u, \tilde{\mathcal{Q}}u\rangle = 0, \qquad \tilde{\mathcal{Q}}u \neq 0, \tilde{\mathcal{K}}\tilde{\mathcal{Q}}u \neq 0. \tag{4.40}$$

Now we set $g = \tilde{\mathcal{Q}}u$. Equivalently we have for g that $\langle \tilde{\mathcal{K}}g, g\rangle = 0, g \neq 0, \tilde{\mathcal{K}}g \neq 0$. This indicates $\mathrm{Re}\langle \tilde{\mathcal{K}}g, g\rangle = \sum_{i=1}^{d}\langle X_i g, X_i g\rangle = 0$, which requires g belonging to the kernel space of all annihilation operators X_i, $1 \leq i \leq d$. Therefore g must be of the form $g = \mu(q)e^{-\frac{\beta}{4m}\|p\|^2}$. On the other hand, since $\tilde{\mathcal{Q}}\tilde{\mathcal{K}}\tilde{\mathcal{Q}}u = \tilde{\mathcal{Q}}\tilde{\mathcal{K}}g = 0$, we shall get that $\tilde{\mathcal{P}}\tilde{\mathcal{K}}g = \tilde{\mathcal{K}}g$. By plugging in the general form of g into this expression, we obtain

$$\begin{aligned}\tilde{\mathcal{K}}g = X_0 g &= \sum_{i=1}^{d} p_i \left[\frac{1}{m}\partial_{q_i}\mu(q) - \frac{\beta}{2m}\mu(q)\partial_{q_i}V(q)\right] e^{-\frac{\beta}{4m}\|p\|^2} \\ &= \sum_{i=1}^{d}[F(q)]_i p_i e^{-\frac{\beta}{4m}\|p\|^2} \qquad (4.41)\\ \tilde{\mathcal{P}}\tilde{\mathcal{K}}g &= \sum_{i=1}^{N}\langle \tilde{\mathcal{K}}g, f_i(q)\rangle_{eq/2} f_j(q) e^{-\frac{\beta}{2}V(q)} e^{-\frac{\beta}{4m}\|p\|^2} \qquad (4.42)\end{aligned}$$

Since $F(q)$ is a d-dimensional vector function of the position q, for projection operator $\tilde{\mathcal{P}}$ defined as (4.37), it is easy to get that equation (4.41)=(4.42) only when $\tilde{\mathcal{K}}g = 0$, which contradicts the assumption $\tilde{\mathcal{K}}\tilde{\mathcal{Q}}u \neq 0$. Hence we conclude that there is no u such that $\tilde{\mathcal{Q}}\tilde{\mathcal{K}}\tilde{\mathcal{Q}}u = 0$ while $\tilde{\mathcal{Q}}u \neq 0$ and $\tilde{\mathcal{K}}\tilde{\mathcal{Q}}u \neq 0$.

Combining these three cases, we can get that the kernel of operator $\tilde{\mathcal{K}}\tilde{\mathcal{Q}}\tilde{\mathcal{K}}$ is given by finite dimensional linear space $Ker(\tilde{\mathcal{K}}) \oplus Id(\tilde{\mathcal{P}})$.

(4.39) claims that the only eigenvalue that operator $\tilde{\mathcal{Q}}\tilde{\mathcal{K}}\tilde{\mathcal{Q}}$ has in $i\mathbb{R}$ is the origin. i.e. for all u in real Hilbert space $L^2(\mathbb{R}^{2d})$ satisfying $\tilde{\mathcal{Q}}\tilde{\mathcal{K}}\tilde{\mathcal{Q}}u = i\lambda u$ for $\lambda \in \mathbb{R}$, we have $\lambda = 0$. To see this, we choose $g = \tilde{\mathcal{Q}}u$, $\tilde{\mathcal{Q}}\tilde{\mathcal{K}}\tilde{\mathcal{Q}}u = i\lambda u$ implies $\mathrm{Re}\langle \tilde{\mathcal{Q}}\tilde{\mathcal{K}}\tilde{\mathcal{Q}}u, u\rangle = \mathrm{Re}\langle \tilde{\mathcal{K}}g, g\rangle = 0$, which again implies $g = \mu(q)e^{-\frac{\beta}{4m}\|p\|^2}$. Since g

is a real function and X_0 a real operator, we have $\text{Im}\langle\tilde{\mathcal{K}}g, g\rangle = \text{Im}\langle X_0 g, g\rangle = 0$. Hence when $\tilde{\mathcal{Q}}\tilde{\mathcal{K}}\tilde{\mathcal{Q}}u = i\lambda u$ for $\lambda \in \mathbb{R}$, we must have $\lambda = 0$. $\square$

We use (4.37) as an example to show the kernel of $\mathcal{QKQ}$ is given by $Ker(\mathcal{K}) \oplus Id(\mathcal{P})$. In fact, for other Mori-type projection operators associated with observables different from $\{f_j(q)\}_{i=1}^N$, it is quite rare that the kernel of $\mathcal{QKQ}$ is *not* $Ker(\mathcal{K}) \oplus Id(\mathcal{P})$ simply because equation (4.41)=(4.42) in Case 3 mostly hold only when $\tilde{\mathcal{K}}g = 0$. Thus we can roughly claims that Lemma 4.2.1 are almost always true for Mori-type projection operator. However, we must emphasize it is possible to have (4.41)=(4.42) when $\tilde{\mathcal{K}}g \neq 0$, such that $dim(Ker(\mathcal{QKQ})) > dim(Ker(\mathcal{K}) \oplus Id(\mathcal{P})$. This happens when $\mathcal{P}$ are projectors to p_j. For instance, if $\mathcal{P}, \tilde{\mathcal{P}}$ are defined as

$$\mathcal{P}(\cdot) = \langle(\cdot), p_j\rangle_{eq} p_j \qquad \tilde{\mathcal{P}}(\cdot) = \langle(\cdot), p_j\rangle_{eq/2} p_j e^{-\beta\mathcal{H}/2} \tag{4.43}$$

Then for Case 3, we have

$$\tilde{\mathcal{K}}g = \sum_{i=1}^{d} [F(q)]_i p_i e^{-\frac{\beta}{4m}\|p\|^2} \tag{4.44}$$

$$\tilde{\mathcal{P}}\tilde{\mathcal{K}}g = \langle\tilde{\mathcal{K}}g, p_j\rangle_{eq/2} p_j e^{-\frac{\beta}{2}V(q)} e^{-\frac{\beta}{4m}\|p\|^2} \tag{4.45}$$

where $g = \tilde{\mathcal{Q}}u$. Equation (4.44) = (4.45) when the following ODE system admits a non-zero solution $\mu(q)$

$$\begin{cases} [F(q)]_i = 0, \quad i \neq j \\ [F(q)]_j = \langle [F(q)]_j, p_j\rangle_{eq/2} e^{-\frac{\beta}{2}V(q)}, \quad i = j \end{cases} \tag{4.46}$$

(4.46) is composed of a system of linear ODEs and a linear integro-differential equation. Hence for some suitable $V(q)$, it admits non-zero solution $\mu(q)$. By solving another linear equation $u = \tilde{\mathcal{P}}u + g$, one may get a non-zero u such that $\tilde{\mathcal{Q}}\tilde{\mathcal{K}}\tilde{\mathcal{Q}}u = 0$ while $\tilde{\mathcal{Q}}u \neq 0$ and $\tilde{\mathcal{K}}\tilde{\mathcal{Q}}u \neq 0$. This special case needs to be taken care of when discussing the behaviour of the memory kernel. Since $dim(Ker(\mathcal{QKQ})) > dim(Ker(\mathcal{K}) \oplus Id(\mathcal{P})$, $K(t)$ will no longer converge to $\mathbb{E}[\mathcal{K}u_0]\mathbb{E}[\mathcal{QK}^*u_0]$ as it usually does (see Corollary 4.2.1.1). With Lemma 4.2.1, we get the following exponential convergence estimate for semigroup $e^{-t\mathcal{QKQ}}$.

Proposition 3. *For Langevin dynamics* (4.28) *with potential energy function $V(q)$ satisfying Hypothesis 1, there exits a positive constant $C = C(\alpha_{\mathcal{Q}})$ such that the estimate*

$$\|e^{-t\mathcal{QKQ}}u_0 - \sqcap_0^{\mathcal{Q}}u_0\|_{L^2_{eq}} \leq Ce^{-\alpha_{\mathcal{Q}}t}\|u_0\|_{L^2_{eq}} \tag{4.47}$$

holds for all $u_0 \in L^2(\mathbb{R}^{2d}, \rho_{eq})$ and for all $t > 0$. In (4.47)*, $\mathcal{P}$ is a finite rank projection operator defined as* (4.37)*, $\sqcap_0^{\mathcal{Q}}$ is the orthogonal projection on linear space $Ker(\mathcal{K}) \oplus Id(\mathcal{P})$, with the specific form*

$$\sqcap_0^{\mathcal{Q}}(\cdot) = \sqcap_0(\cdot) + \mathcal{P}(\cdot) \tag{4.48}$$

unique up to isomorphism.

Proof. We still perform analysis on the flat Hilbert space $L^2(\mathbb{R}^{2d})$. The $L^2(\mathbb{R}^{2d})$ equivalent of the estimate (4.23) is given by

$$\|e^{-t\mathcal{QKQ}}u_0 - \sqcap_0^{\mathcal{Q}}u_0\|_{L^2_{eq}} = \|e^{-t\tilde{\mathcal{Q}}\tilde{\mathcal{K}}\tilde{\mathcal{Q}}}\tilde{u}_0 - \tilde{\sqcap}_0^{\tilde{\mathcal{Q}}}\tilde{u}_0\|_{L^2} \leq Ce^{-\alpha_{\mathcal{Q}}t}\|\tilde{u}_0\|_{L^2} = Ce^{-\alpha t}\|u_0\|_{L^2_{eq}} \tag{4.49}$$

where $\tilde{\sqcap}_0^{\tilde{\mathcal{Q}}} = \mathcal{U}^{-1}\sqcap_0^{\mathcal{Q}}\mathcal{U}$. According to Proposition 2, The transformed Kolmogorov operator $\tilde{\mathcal{K}}$ is of the form (4.1), which has compact resolvent and cusp shape spectrum. Then by Theorem 4.1.3 ($\tilde{\mathcal{P}}$ is a finite rank projection in $L^2(\mathbb{R}^{2d})$), the operator $\tilde{\mathcal{Q}}\tilde{\mathcal{K}}\tilde{\mathcal{Q}}$ has the same properties. Theorem 4.1.4 can be used to get estimate (4.47) after we verify the following facts

1. The spectral projection appeared in Theorem 4.1.4 can be replaced by an projection operator $\tilde{\sqcap}_0^{\tilde{\mathcal{Q}}}$ in $L^2(\mathbb{R}^{2d})$, which projects any $u \in L^2(\mathbb{R}^{2d})$ on linear subspace $Ker(\tilde{\mathcal{Q}}\tilde{\mathcal{K}}\tilde{\mathcal{Q}})$, i.e., $Ran(\tilde{\sqcap}_0^{\tilde{\mathcal{Q}}}) = Ker(\tilde{\mathcal{Q}}\tilde{\mathcal{K}}\tilde{\mathcal{Q}}) = Id(\tilde{\mathcal{P}}) \oplus Ker(\tilde{\mathcal{K}})$.
2. $\sqcap_0^{\mathcal{Q}}$ is an orthogonal operator in $L^2(\mathbb{R}^{2d}, \rho_{eq})$ and given by the form (4.48), which is unique up to isomorphism.

Claim 1. We notice that for self-adjoint operator $\tilde{\mathcal{P}}$, $Id(\tilde{\mathcal{P}}) = Id(\tilde{\mathcal{P}}^*)$. In [38], Nier and Helffer already showed that $Ker(\tilde{\mathcal{K}}) = Ker(\tilde{\mathcal{K}}^*) = \mathbb{R}e^{-\beta\mathcal{H}}$. Hence we shall get

$$Ker(\tilde{\mathcal{Q}}\tilde{\mathcal{K}}\tilde{\mathcal{Q}}) = Ker(\tilde{\mathcal{Q}}\tilde{\mathcal{K}}^*\tilde{\mathcal{Q}}) = Ker(\tilde{\mathcal{K}}) \oplus Id(\tilde{\mathcal{P}}) \tag{4.50}$$

We now consider the orthogonal decomposition of the Hilbert space $L^2(\mathbb{R}^{2d})$

$$L^2(\mathbb{R}^{2d}) = Ker(\tilde{\mathcal{Q}}\tilde{\mathcal{K}}\tilde{\mathcal{Q}}) \overset{\perp}{\oplus} Ker(\tilde{\mathcal{Q}}\tilde{\mathcal{K}}\tilde{\mathcal{Q}})^{\perp},$$

If we define a projection operator $\sqcap_0^{\tilde{\mathcal{Q}}}$ such that $Ran(\sqcap_0^{\tilde{\mathcal{Q}}}) = Ker(\tilde{\mathcal{Q}}\tilde{\mathcal{K}}\tilde{\mathcal{Q}})$, then for any $u_0 \in L^2(\mathbb{R}^{2d})$, we have the orthogonal decomposition

$$\tilde{u_0} = \sqcap_0^{\tilde{\mathcal{Q}}}\tilde{u_0} + (\tilde{u_0} - \sqcap_0^{\tilde{\mathcal{Q}}}\tilde{u}_0), \quad \text{where} \quad \sqcap_0^{\tilde{\mathcal{Q}}}\tilde{u}_0 \in Ker(\tilde{\mathcal{Q}}\tilde{\mathcal{K}}\tilde{\mathcal{Q}}), \tilde{u}_0 - \sqcap_0^{\mathcal{Q}}\tilde{u}_0 \in Ker(\tilde{\mathcal{Q}}\tilde{\mathcal{K}}\tilde{\mathcal{Q}})^{\perp}$$

We now verify that $\tilde{\mathcal{Q}}\tilde{\mathcal{K}}\tilde{\mathcal{Q}}$ maps linear subspace $Ker(\tilde{\mathcal{Q}}\tilde{\mathcal{K}}\tilde{\mathcal{Q}})^{\perp}$ into itself, i.e. for any $u \in Ker(\tilde{\mathcal{Q}}\tilde{\mathcal{K}}\tilde{\mathcal{Q}})^{\perp}$, $\tilde{\mathcal{Q}}\tilde{\mathcal{K}}\tilde{\mathcal{Q}}u \in Ker(\tilde{\mathcal{Q}}\tilde{\mathcal{K}}\mathcal{Q})^{\perp}$. For any $w \in Ker(\tilde{\mathcal{Q}}\tilde{\mathcal{K}}\tilde{\mathcal{Q}})$, since

$Ker(\tilde{\mathcal{Q}}\tilde{\mathcal{K}}\tilde{\mathcal{Q}}) = Ker(\tilde{\mathcal{Q}}\tilde{\mathcal{K}}^*\tilde{\mathcal{Q}})$, we have

$$\langle \tilde{\mathcal{Q}}\tilde{\mathcal{K}}\tilde{\mathcal{Q}}u, w\rangle = \langle u, \tilde{\mathcal{Q}}\tilde{\mathcal{K}}^*\tilde{\mathcal{Q}}w\rangle = 0$$

When $u \in Ker(\tilde{\mathcal{Q}}\tilde{\mathcal{K}}\tilde{\mathcal{Q}})^\perp$, $\tilde{\mathcal{Q}}\tilde{\mathcal{K}}\tilde{\mathcal{Q}}u \neq 0$, hence we must have $\tilde{\mathcal{Q}}\tilde{\mathcal{K}}\tilde{\mathcal{Q}}u \in Ker(\tilde{\mathcal{Q}}\tilde{\mathcal{K}}\tilde{\mathcal{Q}})^\perp$. In general, we can verify that the evolution operator $\tilde{\mathcal{Q}}\tilde{\mathcal{K}}\tilde{\mathcal{Q}}$ and its adjoint $\tilde{\mathcal{Q}}\tilde{\mathcal{K}}^*\mathcal{Q}$ can be decomposed as

$$\begin{aligned}
\tilde{\mathcal{Q}}\tilde{\mathcal{K}}\tilde{\mathcal{Q}} &= \tilde{\mathcal{Q}}\tilde{\mathcal{K}}\tilde{\mathcal{Q}}\Big|_{Ker(\tilde{\mathcal{Q}}\tilde{\mathcal{K}}\tilde{\mathcal{Q}})} \overset{\perp}{\oplus} \tilde{\mathcal{Q}}\tilde{\mathcal{K}}\tilde{\mathcal{Q}}\Big|_{Ker(\tilde{\mathcal{Q}}\tilde{\mathcal{K}}\tilde{\mathcal{Q}})^\perp} \\
\tilde{\mathcal{Q}}\tilde{\mathcal{K}}^*\tilde{\mathcal{Q}} &= \tilde{\mathcal{Q}}\tilde{\mathcal{K}}^*\tilde{\mathcal{Q}}\Big|_{Ker(\tilde{\mathcal{Q}}\tilde{\mathcal{K}}\tilde{\mathcal{Q}})} \overset{\perp}{\oplus} \tilde{\mathcal{Q}}\tilde{\mathcal{K}}^*\tilde{\mathcal{Q}}\Big|_{Ker(\tilde{\mathcal{Q}}\tilde{\mathcal{K}}\tilde{\mathcal{Q}})^\perp}
\end{aligned}$$

Combined with the fact that $\tilde{\sqcap}_0^{\tilde{\mathcal{Q}}}$ is a projection operator therefore $(\tilde{\sqcap}_0^{\tilde{\mathcal{Q}}})^2 = \tilde{\sqcap}_0^{\tilde{\mathcal{Q}}}$, for initial condition $\tilde{u}_0 - \sqcap_0^{\mathcal{Q}}\tilde{u}_0 \in Ker(\tilde{\mathcal{Q}}\tilde{\mathcal{K}}\tilde{\mathcal{Q}})^\perp$ we can deform the boundary of Dunford integral from $[-i\infty, +i\infty]$ to the $\mathcal{S}'_{\tilde{\mathcal{Q}}\tilde{\mathcal{K}}\tilde{\mathcal{Q}}}$ as we did in Theorem 4.1.2. This leads to

$$e^{t\tilde{\mathcal{Q}}\tilde{\mathcal{K}}\tilde{\mathcal{Q}}}\tilde{u}_0 - \tilde{\sqcap}_0^{\tilde{\mathcal{Q}}}\tilde{u}_0 = e^{t\tilde{\mathcal{Q}}\tilde{\mathcal{K}}\tilde{\mathcal{Q}}}(\tilde{u}_0 - \tilde{\sqcap}_0^{\tilde{\mathcal{Q}}}\tilde{u}_0) = \frac{1}{2\pi i}\int_{\partial\mathcal{S}'_{\tilde{\mathcal{Q}}\tilde{\mathcal{K}}\tilde{\mathcal{Q}}}} e^{-tz}(z - \tilde{\mathcal{Q}}\tilde{\mathcal{K}}\tilde{\mathcal{Q}})^{-1}\tilde{u}_0 dz.$$

Following the exact same procedure in Theorem 4.1.2 and 4.1.3, we can get semi-group estimate (4.47).

Claim 2. Equivalently, we prove $\tilde{\sqcap}_0^{\tilde{\mathcal{Q}}}$ is an orthogonal operator in $L^2(\mathbb{R}^{2d})$ and given by the form

$$\tilde{\sqcap}_0^{\tilde{\mathcal{Q}}}(\cdot) = \tilde{\sqcap}_0(\cdot) + \tilde{\mathcal{P}}(\cdot) = \langle\cdot\rangle_{eq/2}e^{-\beta\mathcal{H}/2} + \sum_{i=1}^{N}\langle(\cdot), f_i(q)\rangle_{eq/2} f_i(q)e^{-\beta\mathcal{H}/2}. \tag{4.51}$$

Since $\tilde{\sqcap}_0^{\tilde{\mathcal{Q}}}$ and its L^2 adjoint operator $[\tilde{\sqcap}_0^{\tilde{\mathcal{Q}}}]^*$ are both projection operator, the Hilbert space $L^2(\mathbb{R}^{2d})$ can be decomposed as

$$L^2(\mathbb{R}^{2d}) = Ker(\tilde{\sqcap}_0^{\tilde{\mathcal{Q}}}) \oplus Ran(\tilde{\sqcap}_0^{\tilde{\mathcal{Q}}}), \qquad L^2(\mathbb{R}^{2d}) = Ker([\tilde{\sqcap}_0^{\tilde{\mathcal{Q}}}]^*) \oplus Ran([\tilde{\sqcap}_0^{\tilde{\mathcal{Q}}}]^*)$$

It follows from (4.50) that $Ran(\tilde{\sqcap}_0^{\tilde{\mathcal{Q}}}) = Ran([\tilde{\sqcap}_0^{\tilde{\mathcal{Q}}}]^*) = Ker(\tilde{\mathcal{Q}}\tilde{\mathcal{K}}\tilde{\mathcal{Q}})$. This also implies $Ker(\tilde{\sqcap}_0^{\tilde{\mathcal{Q}}}) = Ker([\tilde{\sqcap}_0^{\tilde{\mathcal{Q}}}]^*)$. Since for any $u, w \in L^2(\mathbb{R}^{2d})$, $w - \tilde{\sqcap}_0^{\tilde{\mathcal{Q}}} w \in Ker(\tilde{\sqcap}_0^{\tilde{\mathcal{Q}}})$, we have

$$\langle \tilde{\sqcap}_0^{\tilde{\mathcal{Q}}} u, w - \tilde{\sqcap}_0^{\tilde{\mathcal{Q}}} w \rangle = \langle u, [\tilde{\sqcap}_0^{\tilde{\mathcal{Q}}}]^* (w - \tilde{\sqcap}_0^{\tilde{\mathcal{Q}}} w) \rangle = 0.$$

Therefore $\tilde{\sqcap}_0^{\tilde{\mathcal{Q}}}$ is an orthogonal operator. On the other hand, since $Id(\tilde{\mathcal{P}}) = Ran(\tilde{\mathcal{P}})$ and $\tilde{\mathcal{P}}$ itself is an orthogonal operator, the projection $\tilde{\sqcap}_0^{\tilde{\mathcal{Q}}}$ is therefore given by the form (4.51) (unique up to isomorphism), which has adjoint given by (4.48). □

Using the same approach, we can prove that the power series expansion of the orthogonal semigroup $e^{-t\mathcal{Q}\mathcal{K}\mathcal{Q}}$ converges in operator norm for any $t > 0$. On the other hand, since the kernel of $\mathcal{Q}\mathcal{K}\mathcal{Q}$ is determined, we can get an exact estimate for the exponentially decaying memory kernel.

Corollary 4.2.1.1. *For Langevin dynamics* (4.28) *with potential energy function $V(q)$ satisfying Hypothesis 1, if we introduce a Mori-type projection operator $\mathcal{P}$ associated with scalar observable $u = u(x(t))$ such that $Ker(\mathcal{Q}\mathcal{K}\mathcal{Q}) = Ker(\mathcal{K}) \oplus Id(\mathcal{P})$, then the memory kernel of* (2.35c) *exponentially converges to $\mathbb{E}[\mathcal{K}u_0]\mathbb{E}[\mathcal{Q}\mathcal{K}^* u_0]$ with rate $\alpha_{\mathcal{Q}}$. i.e.*

$$|K(t) - \mathbb{E}[\mathcal{K}u_0]\mathbb{E}[\mathcal{Q}\mathcal{K}^* u_0]| \leq Ce^{-\alpha_{\mathcal{Q}} t}. \tag{4.52}$$

The same estimate holds for the entries of the memory matrix (2.35c) *for vector form EMZE/GLE.*

Proof. With estimate (4.25) and definition (4.48), using the fact that $\mathcal{P}\mathcal{Q} = 0$ we shall get the result. $\square$

Some open problems The mathematical framework we used to study the EMZE was based on the Hörmander analysis of hypoelliptic operator $\mathcal{K}$. We believe the philosophy and methodology of revising the estimate of $\mathcal{K}$, $e^{-t\mathcal{K}}$ to get similar results for $\mathcal{Q}\mathcal{K}\mathcal{Q}$, $e^{-t\mathcal{Q}\mathcal{K}\mathcal{Q}}$ is applicable to other stochastic systems. Beside the theoretical extension in this direction, here we give some possible developments of this work that are closely related with the Mori-Zwanzig theory itself.

1. Estimate of convergence rate $\alpha_{\mathcal{Q}}$: The real part of the first non-zero eigenvalue of $\mathcal{Q}\mathcal{K}\mathcal{Q}$ corresponds to the maximum convergence rate $\alpha_{\mathcal{Q}}$ of the semigroup $e^{-t\mathcal{Q}\mathcal{K}\mathcal{Q}}$ and MZ memory kernel. For operator $\mathcal{K}$, Hèrau and Nier [39] already show that the convergence rate α is closely related with the lowest positive eigenvalue of the corresponding Witten Laplacian. It would be very meaningful if some connection can be built between α and $\alpha_{\mathcal{Q}}$. For instance, numerical simulations of classical Hamiltonian systems have shown the memory kernel often converges faster than the dynamics which indicates a relationship $\alpha_{\mathcal{Q}} < \alpha$. If such estimate can be verified by analysis, then we would justify the utility of the MZ framework from a theoretic point of view.

2. Analysis for Zwanzig-type projection operator: Zwanzig-type conditional expectation is another frequenty used projection operator within the MZ framework. However, an useful methodology for the systematic treatment of Zwanzig -type MZ equation is lacking for a rather long time in the com-

munity. Most theoretical results, including the well-posedness study in [34] and what we obtained in [103] and this paper, focus on Mori-type MZ equation. As we mentioned before, for an infinite rank projection operator like Zwanzig's projection, one can not get useful estimates for $\mathcal{QKQ}$ simply by revising the analytic results of the kinetic equation. New ideas are needed in this direction.

3. (Effective) Mori-Zwanzig theory for SDE driven by fractional Brownian motion (FBM): In nonequilibrium statistical mechanics, fractional Brownian motion was introduced to model complex dynamics with anomalous long-time behavior [5, 21]. It would be interesting to see if one can use a similar method as in Secion 2.2 and 2.3 to derive the MZ and effective MZ equation for the FBM system.

4.2.2 Comparison with the classical dynamics

Combined with the conclusions we obtained in Chapter 3, we summarized the analysis results of the Mori-Zwanzig equation for the classical and Langevin dynamics of a Hamiltonian system. We have found that the analytical properties of the evolution operator for the full and the orthogonal dynamics are closely related with the spectrum of the semigroup generator $\mathcal{G}$, which exhibits different features for linear and nonlinear dynamics. Generally speaking, the spectral properties of the Liouville operator $\mathcal{L}$ make the classical dynamics harder to study both theoretically and numerically. The generic continuity of the spectrum indicates that the exponential convergence results do not hold for a general classical dynamical system. In fact, even for the linear case where the spectrum is discrete, $\|e^{t\mathcal{L}} - \sqcap_0\|_{eq}$ does not converge to 0 exponentially. A typical example is the harmonic oscillator chain model studied in [104, 33], the velocity correlation function

Table 4.1: Spectrum of the generator $\mathcal{G}$ and the analytical properties of semigroup $e^{-t\mathcal{G}}$, $e^{-t\mathcal{Q}\mathcal{G}\mathcal{Q}}$. $\sim$ means the same as the previous column.

Name	Classical dynamics		Stochastic dymamics	
Generator $\mathcal{G}$	Liouville operator $\mathcal{L}$		Kolmogorov operator $\mathcal{K}$	
System type	Linear	Nonlinear	Linear	Nonlinear
Spectrum of $\mathcal{G}$	$\sigma(\mathcal{L}) \in i\mathbb{R}$ discrete	$\sigma(\mathcal{L}) \in i\mathbb{R}$ continuous[2]	Cone,discrete	Cusp,discrete
$e^{-t\mathcal{G}}$	$\Vert e^{-t\mathcal{L}}\Vert_{eq} \leq 1$	$\sim$	$\Vert e^{-t\mathcal{K}} - \sqcap_0\Vert_{eq} \leq Ce^{-\alpha t}$	$\sim$
$e^{-t\mathcal{Q}\mathcal{G}\mathcal{Q}}$	$\Vert e^{-t\mathcal{Q}\mathcal{L}\mathcal{Q}}\Vert_{eq} \leq 1$	$\sim$	$\Vert e^{-t\mathcal{Q}\mathcal{K}\mathcal{Q}} - \sqcap_0^{\mathcal{Q}}\Vert_{eq} \leq Ce^{-\alpha_{\mathcal{Q}} t}$	$\sim$
Memory Kernel	$\vert K(t)\vert \leq C$	$\sim$	$\vert K(t) - \langle \mathcal{Q}\mathcal{K}^* u_0, \sqcap_0^{\mathcal{Q}}(\mathcal{K}u_0)\rangle\vert \leq Ce^{-\alpha_{\mathcal{Q}} t}$	$\sim$
Semigroup Taylor expansion	Convergent if $u_0 \in \{x_{i0}\}_{i=1}^d$	Convergent if $t \to 0$	Convergent if $t > 0, s > t$	$\sim$

of tagged oscillator is given by a first kind Bessel function $J_0(2t)$, which decays to 0 slower than any exponential function.

For the Langevin dynamics, the spectrum for Kolmogorov operator $\mathcal{K}$ is discrete and bounded in a subset of the right half complex plane, which lead to the ergodicity and exponential convergence of the Markovian semigroup. This makes the Langevin dynamics an easier mathematical object to dealt with from both theoretical and computational point of view. However, this also means that it *cannot* be used to simulate systems which exhibit very clear non-exponential decaying features, such as the supercooled liquid [81].

4.3 Summary

In this chapter, we provide a systematic way to study the orthogonal propagator $e^{t\mathcal{Q}\mathcal{K}\mathcal{Q}}$ in the EMZE based on the established theory of Hörmander analysis for linear hypoelliptic equations. In particular, we proved that for suitable $\mathcal{P}$ and $\mathcal{K}$, the spectrum of operator $\mathcal{QKQ}$ locates within a cusp shape region in the complex plane, and the analytic method used to deduce the exponential convergence to equilibrium for Fokker-Planck equation can applied to get the exponential convergence of the MZ memory kernel. On the other hand, the convergence of the power series expansion we obtained for semigroup $e^{t\mathcal{K}}$, $e^{t\mathcal{QKQ}}$ confirmed the convergence of the combinatorial algorithm (for the first time) we introduced in [105] when it is applied to stochastic system (2.21). The abstract theory is applied to Langevin dynamics to get specific convergence results. Through these discussions, we wish to convey the following messages:

1. The properties of the orthogonal propagator $e^{t\mathcal{QGQ}}$ and its generator $\mathcal{QGQ}$ are closely related with the propagator $e^{t\mathcal{G}}$ for original dynamics and its generator $\mathcal{G}$. We believe other analytical results obtained for the general kinetic equation (1.1) can be extended to get similar results for $e^{t\mathcal{QGQ}}$, at least for Mori's projection operator.

2. Although formally simpler, the dynamical properties of the classical system are more complicated than the stochastic system since the general ergodicity for the first one is lacking. Corresponding, the analysis of the MZ equation for the classical dynamics would be harder for the stochastic case due to the spectrum differences of the generator $\mathcal{G}$ and $\mathcal{QGQ}$.

Chapter 5

Faber approximation of the MZE

In previous chapters, we introduced the Mori-Zwanzig equation for classical and stochastic dynamical systems and used functional analysis to get some prior estimations on the MZ memory integral. From this chaper, we turn to the study of the MZ equation from a computational point of view. Generally speaking, computing the solution to the MZ equation is a very challenging task that relies on approximations and appropriate numerical schemes. One of the main difficulties is the approximation of the memory integral (convolution term), which encodes the effects of the so-called orthogonal dynamics in the observable of interest. The orthogonal dynamics is essentially a high-dimensional flow that satisfies a complex integro-differential equation. In statistical systems far from equilibrium, such flow has the same order of magnitude and dynamical properties as the observable of interest, i.e., there is no scale separation between the observable of interest and the orthgonal dynamics. In these cases, the computation of the MZ memory can be addressed only by problem-class-dependent approximations. In Chapter 3, we have introduced various approximations to the MZ memory integral such as the t-model, H_t-model[15, 18, 85, 13] and th hierarchical perturbation methods [87, 96]. In this chapter, we will propose a new approximation of the MZ equation based on *global* operator series expansions of the orthogonal dynamics propagator. In particular, we study the Faber series, which yields asymptotically optimal approximations converging at least R-superlinearly with the polynomial order. The advantages of expanding the orthogonal dynamics propagator in terms of globally defined operator series are similar to those we obtain when we approximate a smooth function in terms of orthogonal polynomials rather than Taylor series. As we will see, the proposed MZ memory approximation method based on global operator series outperform in terms of accuracy and computational efficiency the hierarchical memory approximation techniques discussed in Chapter 3 which are based on Taylor-type expansions.

5.1 New Approximation of the MZ Memory Integral

In this section, we develop new approximations of the Mori-Zwanzig memory integral

$$\int_0^t \mathcal{P}e^{s\mathcal{L}}\mathcal{P}\mathcal{L}e^{(t-s)\mathcal{Q}\mathcal{L}}\mathcal{Q}\mathcal{L}u_0 ds \tag{5.1}$$

based on series expansions of the orthogonal dynamics propagator $e^{t\mathcal{Q}\mathcal{L}}$ in the form

$$e^{t\mathcal{Q}\mathcal{L}} = \sum_{n=0}^{\infty} a_n(t)\Phi_n\left(\mathcal{Q}\mathcal{L}\right), \tag{5.2}$$

where Φ_n are polynomial basis functions, and $a_n(t)$ are temporal modes. Series expansions in the form (5.2) can be rigorously defined in the context of matrix theory [65, 66], i.e., for operators $\mathcal{Q}\mathcal{L}$ between finite-dimensional vector spaces. The series expansion (5.2) needs to be handled with care if $\mathcal{L}$ is an unbounded operator, e.g., the generator of the Koopman semigroup (2.3) (see [48], p. 481).

In this case, $e^{t\mathcal{L}}$ should be properly defined as

$$e^{t\mathcal{L}} = \lim_{q\to\infty} \left(1 - \frac{t\mathcal{L}}{q}\right)^{-q}.$$

In fact, $(1 - t\mathcal{L}/q)^{-1}$ is the resolvent of $\mathcal{L}$ (modulus a constant factor), which can be rigorously defined for both bounded and unbounded operators. Despite theoretical issues associated with the existence of convergent series expansions of semigroups generated by unbounded operators [28, 48], when it comes to computing we always need to use a finite dimensional Hilbert space H and consider the approximated dynamics $e^{t\mathcal{QL}} : H \to H$ within H. In this setting, $e^{t\mathcal{QL}}$ is truly a matrix exponential, which can be expanded as in (5.2).

5.1.1 MZ-Dyson Expansion

Consider the classical Taylor series expansion of the orthogonal dynamics propagator

$$e^{t\mathcal{QL}} = \sum_{n=0}^{\infty} \frac{t^n}{n!} (\mathcal{QL})^n. \tag{5.3}$$

A substitution of this expansion into the MZ equation (2.6) yields

$$\begin{aligned}
\frac{\partial}{\partial t}\mathcal{P}e^{t\mathcal{L}}u_0 &= \mathcal{P}e^{t\mathcal{L}}\mathcal{PL}u_0 + \int_0^t \mathcal{P}e^{s\mathcal{L}}\mathcal{PL}e^{(t-s)\mathcal{QL}}\mathcal{QL}u_0 ds,\\
&= \mathcal{P}e^{t\mathcal{L}}\mathcal{PL}u_0 + \int_0^t \sum_{n=0}^{\infty} \frac{(t-s)^n}{n!} \underbrace{\mathcal{P}e^{s\mathcal{L}}\mathcal{PL}(\mathcal{QL})^n\mathcal{QL}u_0}_{\mathcal{C}_n(s)u_0} ds,\\
&= \mathcal{P}e^{t\mathcal{L}}\mathcal{PL}u_0 + \int_0^t \underbrace{\left[\sum_{n=0}^{\infty} \mathcal{C}_n(s)\frac{(t-s)^n}{n!}\right]}_{\mathcal{G}(t-s,s)} u_0 ds,\\
&= \mathcal{P}e^{t\mathcal{L}}\mathcal{PL}u_0 + \int_0^t \mathcal{G}(t-s,s)u_0 ds,
\end{aligned} \tag{5.4}$$

where the *memory operator*[1] $\mathcal{G}(t-s,s)$ is defined as

$$\mathcal{G}(t-s,s)=\sum_{n=0}^{\infty}\frac{(t-s)^n}{n!}\mathcal{C}_n(s),\qquad \mathcal{C}_n(s)=\mathcal{P}e^{s\mathcal{L}}\mathcal{P}\mathcal{L}(\mathcal{Q}\mathcal{L})^n\mathcal{Q}\mathcal{L},\quad n\geq 0.\quad (5.5)$$

We shall call this series expansion of the MZ equation as *MZ-Dyson expansion*. The reason for such definition is that (5.4) is equivalent to the H-model discussed in Chapter 3, which in turn is equivalent to a Dyson series expansion in the form

$$\frac{\partial}{\partial t}\mathcal{P}e^{t\mathcal{L}}u_0=\mathcal{P}e^{t\mathcal{L}}\mathcal{P}\mathcal{L}u_0+w_0(t)$$

where

$$\begin{aligned}w_0(t)=&\int_0^t\mathcal{P}e^{s\mathcal{L}}\mathcal{P}\mathcal{L}\mathcal{Q}\mathcal{L}x_0ds+\int_0^t\int_0^{\tau_1}\mathcal{P}e^{s\mathcal{L}}\mathcal{P}\mathcal{L}\mathcal{Q}\mathcal{L}\mathcal{Q}\mathcal{L}x_0dsd\tau_1\\&+...+\int_0^t\int_0^{\tau_{n-1}}...\int_0^{\tau_1}\mathcal{P}e^{s\mathcal{L}}\mathcal{P}\mathcal{L}(\mathcal{Q}\mathcal{L})^nx_0dsd\tau_1...d\tau_{n-1}+....\end{aligned}\quad (5.6)$$

To prove such equivalence, we just need to prove that

$$\int_0^t\int_0^{\tau_{n-1}}...\int_0^{\tau}\mathcal{P}e^{s\mathcal{L}}\mathcal{P}\mathcal{L}(\mathcal{Q}\mathcal{L})^ndsd\tau_1...d\tau_{n-1}=\int_0^t\frac{(t-s)^{n-1}}{(n-1)!}\mathcal{P}e^{s\mathcal{L}}\mathcal{P}\mathcal{L}(\mathcal{Q}\mathcal{L})^n\mathcal{Q}\mathcal{L}ds.\quad (5.7)$$

We proceed by induction. To this end, we first define

$$\begin{aligned}\mathcal{A}_n(t)&=\int_0^t\int_0^{\tau_{n-1}}...\int_0^{\tau}\mathcal{P}e^{s\mathcal{L}}\mathcal{P}\mathcal{L}(\mathcal{Q}\mathcal{L})^ndsd\tau_1...d\tau_{n-1}\\\mathcal{B}_n(t)&=\int_0^t\frac{(t-s)^{n-1}}{(n-1)!}\mathcal{P}e^{s\mathcal{L}}\mathcal{P}\mathcal{L}(\mathcal{Q}\mathcal{L})^n\mathcal{Q}\mathcal{L}ds.\end{aligned}\quad (5.8)$$

For $n=1$ we have $\mathcal{A}_1=\mathcal{B}_1$. For $n\geq 2$ we have $\mathcal{A}_n'(t)=\mathcal{A}_{n-1}(t)$, $\mathcal{B}_n'(t)=\mathcal{B}_{n-1}(t)$ and $\mathcal{A}_n(0)=\mathcal{B}_n(0)$. Hence, by induction we conclude that $\mathcal{A}_n(t)=\mathcal{B}_n(t)$, and

[1] Note that $\mathcal{G}(t-s,s)$ here is not a function but a linear operator.

therefore the memory integral in (5.4), with $\mathcal{G}$ given in (5.5), is equivalent to a Dyson series.

5.1.2 MZ-Faber Expansion

The Faber series of the orthogonal dynamics propagator $e^{t\mathcal{QL}}$ is an operator series in the form (see Appendix 5.A)

$$e^{t\mathcal{QL}} = \sum_{j=0}^{\infty} a_j(t)\mathcal{F}_j(\mathcal{QL}), \tag{5.9}$$

where $\mathcal{F}_j$ is the $j-$th order Faber polynomial, and $a_j(t)$ are suitable temporal modes defined hereafter. The series expansion (5.9) is *asymptotically optimal*, in the sense that its m-th order truncation uniformly approximates the best sequence of operator polynomials converging to $e^{t\mathcal{QL}}$ as $m \to \infty$ [27]. A substitution of (5.9) into (5.1) yields the following expansion of the MZ equation (2.6)

$$\begin{aligned}
\frac{\partial}{\partial t}\mathcal{P}e^{t\mathcal{L}}u_0 &= \mathcal{P}e^{t\mathcal{L}}\mathcal{PL}u_0 + \int_0^t \mathcal{P}e^{s\mathcal{L}}\mathcal{PL}e^{(t-s)\mathcal{QL}}\mathcal{QL}u_0 ds, \\
&= \mathcal{P}e^{t\mathcal{L}}\mathcal{PL}u_0 + \int_0^t \sum_{j=0}^{\infty} a_j(t-s) \underbrace{\mathcal{P}e^{s\mathcal{L}}\mathcal{PL}\mathcal{F}_j(\mathcal{QL})\mathcal{QL}u_0}_{\mathcal{C}_j(s)u_0} ds, \\
&= \mathcal{P}e^{t\mathcal{L}}\mathcal{PL}u_0 + \int_0^t \mathcal{G}(t-s,s)u_0 ds,
\end{aligned} \tag{5.10}$$

where

$$\mathcal{G}(t-s,s) = \sum_{j=0}^{\infty} a_j(t-s)\mathcal{C}_j(s), \tag{5.11}$$

and

$$a_j(t-s) = \frac{1}{2\pi i}\int_{|w|=R} \frac{e^{(t-s)\psi(w)}}{w^{j+1}} dw, \qquad \mathcal{C}_j(s) = \mathcal{P}e^{s\mathcal{L}}\mathcal{PL}\mathcal{F}_j(\mathcal{QL})\mathcal{QL}. \tag{5.12}$$

Here, $\psi(w)$ is the conformal map at the basis of the Faber series (see Appendix 5.A). The coefficients of the Laurent expansion of ψ determine the recurrence relation of the Faber polynomials. High-order Laurent series usually yield higher convergence rates, but complicated recurrence relations (see equation (5.101)). Moreover, the computation of the integrals in (5.12) can be quite cumbersome if high-order Laurent series are employed. To avoid such drawbacks, in this paper we choose the conformal map $\psi(w) = w + c_0 + c_1/w$. This yields the following expression for the coefficients $a_j(t-s)$

$$a_j(t-s) = \frac{e^{(t-s)c_0}}{(\sqrt{-c_1})^j} J_j\left(2(t-s)\sqrt{-c_1}\right), \tag{5.13}$$

where J_j denotes the $j-$th Bessel function of the first kind. In Section 5.2 we prove that the Faber expansion of the MZ memory integral converges for any linear dynamical system and any finite integration time with rate that is at least R-superlinear.

Remark The MZ-Dyson expansion we discussed in Section 5.1.1 is a subcase of the Faber expansion. In fact, Faber polynomials $\mathcal{F}_j(\mathcal{QL})$ corresponding to the conformal mapping $\psi(w) = w$ are simply monomials $(\mathcal{QL})^j$ (see Appendix 5.A). Moreover, the temporal modes (5.13) reduce to $(t-s)^j/j!$ if we set $c_0 = 0$ and take the limit $c_1 \to 0$.

5.1.3 Other Series Expansions of the MZ-Memory Integral

The operator exponential $e^{t\mathcal{QL}}$ (propagator of the orthogonal dynamics) can be expanded relative to basis functions other than Faber polynomials [65, 66]. This yields different approximations of the MZ memory integral and, correspondingly, different expansions of the MZ equation. Hereafter we discuss two relevant cases.

MZ-Lagrange Expansion

The MZ-lagrange expansion is based on the following semigroup expansion

$$e^{t\mathcal{QL}} = \sum_{j=1}^{n} e^{\lambda_j t} \prod_{\substack{k=1 \\ k\neq j}}^{n} \frac{(\mathcal{QL} - \lambda_k \mathcal{I})}{(\lambda_j - \lambda_k)}, \tag{5.14}$$

where $\{\lambda_1, ..., \lambda_n\} = \sigma(\mathcal{QL})$ is the spectrum of the matrix representation of the operator $\mathcal{QL}$ (eigenvalues counted with their multiplicity). Note that (5.14) is in the form (5.2) with

$$a_j(t) = e^{\lambda_j t}, \quad \text{and} \quad \Phi_j(\mathcal{QL}) = \prod_{\substack{k=1 \\ k\neq j}}^{n} \frac{(\mathcal{QL} - \lambda_k \mathcal{I})}{(\lambda_j - \lambda_k)}. \tag{5.15}$$

A substitution of (5.14) into the MZ equation yields the MZ-Lagrange expansion

$$\frac{\partial}{\partial t}\mathcal{P}e^{t\mathcal{L}}u_0 = \mathcal{P}e^{t\mathcal{L}}\mathcal{PL}u_0 + \int_0^t \mathcal{G}(t-s,s)u_0 ds, \tag{5.16}$$

where

$$\mathcal{G}(t-s,s) = \sum_{j=1}^{n} e^{(t-s)\lambda_j}\mathcal{C}_j(s), \quad \text{and} \quad \mathcal{C}_j(s) = \mathcal{P}e^{s\mathcal{L}}\mathcal{PL}\prod_{\substack{k=1 \\ k\neq j}}^{n} \frac{(\mathcal{QL} - \lambda_k \mathcal{I})}{(\lambda_j - \lambda_k)}, \qquad j \geq 1. \tag{5.17}$$

MZ-Newton Expansion

The MZ-Newton expansion is based on the following semigroup expansion

$$e^{t\mathcal{QL}} = f_{1,1}(t)\mathcal{I} + \sum_{j=2}^{n} f_{1,j}(t) \prod_{k=1}^{j-1} (\mathcal{QL} - \lambda_k \mathcal{I}), \tag{5.18}$$

where $f_{1,j}(t)$ is the divided difference defined recursively by

$$f_{1,j}(t) = \begin{cases} e^{\lambda_1 t} & j = 1, \\ \dfrac{e^{t\lambda_1} - e^{t\lambda_2}}{\lambda_1 - \lambda_2} & j = 2, \\ \dfrac{f_{1,j-1}(t) - f_{2,j}(t)}{\lambda_1 - \lambda_j} & j \geq 3. \end{cases} \tag{5.19}$$

A substitution of the Newton expansion (5.18) into the MZ equation yields the following MZ-Newton expansion

$$\frac{\partial}{\partial t}\mathcal{P}e^{t\mathcal{L}}u_0 = \mathcal{P}e^{t\mathcal{L}}\mathcal{P}\mathcal{L}u_0 + \int_0^t \mathcal{G}(t-s,s)u_0 ds, \tag{5.20}$$

where

$$\begin{aligned} \mathcal{G}(t-s,s) &= \mathcal{C}_1(s)e^{(t-s)\lambda_1} + \sum_{j=2}^{n} \mathcal{C}_j(s) f_{1,j}(t) \\ \mathcal{C}_j(s) &= \begin{cases} \mathcal{P}e^{s\mathcal{L}}\mathcal{P}\mathcal{L} & j = 1 \\ \mathcal{P}e^{s\mathcal{L}}\mathcal{P}\mathcal{L}\displaystyle\prod_{k=1}^{j-1}(\mathcal{Q}\mathcal{L} - \lambda_k\mathcal{I})\mathcal{Q}\mathcal{L} & j \geq 2 \end{cases}. \end{aligned} \tag{5.21}$$

In conclusion, for Mori-Zwanzig memory operator

$$\mathcal{G}(t-s,s) = \sum_{j=0}^{\infty} h_j(t-s)\mathcal{C}_j(s)$$

we summarized different series expansions and their corresponding representations of the memory kernel in Table 5.1. All series expansion methods we considered so far aim at representing the memory integral in the Mori-Zwazing equation for the same phase space function. Therefore, such series should be related to each other. Indeed, as shown in Table 5.1, they basically represent the same memory operator $\mathcal{G}(t-s,s)$ relative to different bases. This also means that the series can

Type	Temporal bases $h_j(t)$	Operators $\mathcal{C}_j(s)$
MZ-Dyson	$\dfrac{t^j}{j!}$	$\mathcal{P}e^{s\mathcal{L}}\mathcal{P}\mathcal{L}(\mathcal{Q}\mathcal{L})^j\mathcal{Q}\mathcal{L}$
MZ-Faber	$e^{tc_0}\dfrac{J_j(2t\sqrt{-c_1})}{(\sqrt{-c_1})^j}$	$\mathcal{P}e^{s\mathcal{L}}\mathcal{P}\mathcal{L}\mathcal{F}_j(\mathcal{Q}\mathcal{L})\mathcal{Q}\mathcal{L}$
MZ-Lagrange	$e^{t\lambda_j}$	$\mathcal{P}e^{s\mathcal{L}}\mathcal{P}\mathcal{L}\prod_{\substack{k=1\\k\neq j}}^{n}\dfrac{(\mathcal{Q}\mathcal{L}-\lambda_k\mathcal{I})}{(\lambda_j-\lambda_k)}$
MZ-Newton	$f_{1,j}(t)$	$\begin{cases}\mathcal{P}e^{s\mathcal{L}}\mathcal{P}\mathcal{L} & j=1\\ \mathcal{P}e^{s\mathcal{L}}\mathcal{P}\mathcal{L}\prod_{k=1}^{j-1}(\mathcal{Q}\mathcal{L}-\lambda_k\mathcal{I})\mathcal{Q}\mathcal{L} & j\geq 2\end{cases}$

Table 5.1: Series expansions of the Mori-Zwanzig memory operator. Here J_j is the jth Bessel function of the first kind, c_0 and c_1 are real numbers, $f_{1,j}(t)$ are defined in (5.19), and λ_j are the eigenvalues of any matrix representation of $\mathcal{Q}\mathcal{L}$.

have different convergence rate. For example, as we will demonstrate numerically in Section 6.3 the MZ-Faber expansion converges much faster than the MZ-Dyson series.

5.1.4 Generalized Langevin Equation

We have seen in Section 5.1 that expanding the orthogonal dynamics propagator $e^{t\mathcal{Q}\mathcal{L}}$ in an operator series in the form (5.2) yields the Mori-Zwanzig equation[2]

$$\frac{\partial}{\partial t}\mathcal{P}e^{t\mathcal{L}}u_0 = \mathcal{P}e^{t\mathcal{L}}\mathcal{P}\mathcal{L}u_0 + \sum_{j=0}^{\infty}\int_0^t h_j(t-s)\mathcal{C}_j(s)u_0 ds, \tag{5.22}$$

[2] We emphasize that if we also apply the semigroup expansion (5.9) on the noise term $e^{t\mathcal{Q}\mathcal{L}}\mathcal{Q}\mathcal{L}u_0$ of the full MZ equation, we end up with a stochastic differential equation. However, getting a convergence theorem for this SDE is challenging since almost inevitably, phenomenological approximation to the noise term has to be introduced which contaminates the MZ equation with uncontrollable error.

where $h_j(t-s)$ are temporal modes, and $\mathcal{C}_j(s)$ are operators defined in Table 5.1. For example, if we consider the MZ-Dyson expansion, we have

$$h_j(t-s) = \frac{(t-s)^j}{j!}, \qquad \mathcal{C}_j(s) = \mathcal{P}e^{s\mathcal{L}}\mathcal{P}\mathcal{L}(\mathcal{Q}\mathcal{L})^j\mathcal{Q}\mathcal{L}. \tag{5.23}$$

Equation (5.22) is the exact generalized Langevin equation (GLE) governing the projected dynamics of a quantity of interest. Such equation has different forms depending on the choice of the projection operator $\mathcal{P}$. In particular, if we choose Chorin's projection, then (5.22) is an equation for the conditional expectation of the quantity of interest. On the other hand, if we choose Mori's projection, then (5.22) becomes an equation for the autocorrelation function of the quantity of interest.

Evolution Equation for the Conditional Expectation

If we consider Chorin's projection (2.7), then (5.22) becomes an unclosed evolution equation for the conditional expectation of the quantity of interest (see Section 2.1.1). However, in the special case where the dynamical system (2.1) is *linear* and the quantity of interest is $u(\boldsymbol{x}) = x_1(t)$, it can be shown that the evolution equation for the conditional expectation is closed. To this end, let us first recall that if $\mathcal{P}$ is Chorin's projection and $u(\boldsymbol{x}) = x_1$ then

$$\mathcal{P}e^{t\mathcal{L}}x_1(0) = \langle x_1(t)\rangle_{\rho_0} = \int x_1(t, \boldsymbol{x}_0)\rho_0(\boldsymbol{x}_0)d\boldsymbol{x}_0.$$

In this case, (5.22) reduces to

$$\frac{d}{dt}\langle x_1(t)\rangle_{\rho_0} = a\langle x_1(t)\rangle_{\rho_0} + b + \int_0^t g(t-s)\langle x_1(s)\rangle_{\rho_0}ds + \int_0^t f(t-s)ds, \tag{5.24}$$

where the constants a, b, the *MZ memory kernel* $g(t-s)$, and the function $f(t-s)$ are defined by

$$\mathcal{P}\mathcal{L}x_1(0) = ax_1(0)+b, \qquad g(t-s) = \sum_{j=0}^{\infty} g_j h_j(t-s), \qquad f(t-s) = \sum_{j=0}^{\infty} f_j h_j(t-s). \tag{5.25}$$

The coefficients g_j, f_j and the temporal bases $h_j(t-s)$ appearing in the series expansions above depend on the series expansion of the orthogonal dynamics propagator $e^{t\mathcal{Q}\mathcal{L}}$. Specifically, g_j and f_j are determined by the equation

$$\mathcal{C}_j(s)x_1(0) = g_j\langle x_1(s)\rangle_{\rho_0} + f_j, \tag{5.26}$$

while $h_j(t-s)$ and $\mathcal{C}_j(s)$ are defined in Table 5.1. To derive equation (5.26) we used the identity $\mathcal{P}e^{s\mathcal{L}}f_j = f_j$. In the case of MZ-Dyson and MZ-Faber expansions we explicitly obtain

$$\mathcal{P}\mathcal{L}(\mathcal{Q}\mathcal{L})^j\mathcal{Q}\mathcal{L}x_1(0) = g_j^D x_1(0) + f_j^D, \qquad \mathcal{P}\mathcal{L}\mathcal{F}_j(\mathcal{Q}\mathcal{L})\mathcal{Q}\mathcal{L}x_1(0) = g_j^F x_1(0) + f_j^F, \tag{5.27}$$

where the superscripts D and F stand for "Dyson" and "Faber", respectively.

Evolution Equation for the Autocorrelation Function

If we choose the projection operator $\mathcal{P}$ to be Mori's projection (3.70), then equation (5.22) becomes a closed evolution equation for the autocorrelation function $C_u(t)$ of the quantity of interest. Such equation has the form

$$\frac{dC_u(t)}{dt} = aC_u(t) + \int_0^t g(t-s)C_u(s)ds, \tag{5.28}$$

where a and g are defined as

$$\mathcal{P}\mathcal{L}u_0 = au_0, \qquad g(t-s) = \sum_{j=0}^{\infty} g_j h_j(t-s). \tag{5.29}$$

As before, the temporal modes h_j and the coefficients g_j in the expansion of the MZ-memory kernel $g(t-s)$ depend on the expansion of the orthogonal dynamics propagator $e^{t\mathcal{Q}\mathcal{L}}$. Specifically, in the case of MZ-Dyson and MZ-Faber expansions we obtain, respectively,

$$\mathcal{P}\mathcal{L}(\mathcal{Q}\mathcal{L})^j\mathcal{Q}\mathcal{L}u_0 = g_j^D u_0, \qquad \mathcal{P}\mathcal{L}\mathcal{F}_j(\mathcal{Q}\mathcal{L})\mathcal{Q}\mathcal{L}u_0 = g_j^F u_0. \tag{5.30}$$

Analytical Solution to the Generalized Langevin Equation

The analytical solution to the MZ equations (5.24) and (5.28) can be computed through Laplace transforms. To this end, let us first notice that both equations are in the form of a Volterra equation

$$\frac{dy(t)}{dt} = ay(t) + b + \int_0^t g(t-s)y(s)ds + \int_0^t f(t-s)ds. \tag{5.31}$$

Applying the Laplace transform

$$\mathscr{L}[\cdot](s) = \int_0^{\infty} (\cdot)e^{-st}dt \tag{5.32}$$

to both sides of (5.31) yields

$$sY(s) - y(0) = aY(s) + \frac{b}{s} + Y(s)G(s) + \frac{F(s)}{s}, \tag{5.33}$$

i.e.,

$$Y(s) = \frac{(F(s)+b)/s + y(0)}{s - G(s) - a}, \tag{5.34}$$

where

$$Y(s) = \mathscr{L}[y(t)], \qquad F(s) = \mathscr{L}[f(t)], \qquad G(s) = \mathscr{L}[g(t)]. \tag{5.35}$$

Thus, the exact solution to the Volterra equation (5.31) can be written as

$$y(t) = \mathscr{L}^{-1}\left[\frac{(F(s)+b)/s + y(0)}{s - G(s) - a}\right]. \tag{5.36}$$

The Laplace transform of the memory kernel $g(t)$, i.e., $G(s)$, can be computed analytically in many cases. For example, in the case of MZ-Dyson and MZ-Faber expansions we obtain, respectively

$$G(s) = \sum_{j=0}^{\infty} \frac{g_j^D}{s^{j+1}} \qquad \text{(MZ-Dyson)}, \tag{5.37}$$

$$G(s) = \sum_{j=0}^{\infty} \frac{g_j^F}{2^j(\sqrt{-c_1})^{2j}} \frac{(\sqrt{s^2-4c_1}-s)^j}{\sqrt{s^2-4c_1}} \qquad \text{(MZ-Faber)}. \tag{5.38}$$

The coefficients g_j^F and g_j^D are explicitly defined in (5.27), or (5.30), depending on whether we are interested in the mean or the correlation function of the quantity of interest.

Remark The recurrence relation at the basis of the Faber polynomials (see equation (5.101)) induces a recurrence relation in the Laplace transform $G(s)$ of the MZ memory kernel. Therefore, a connection between the MZ-Faber approximation method we propose here and the method of recurrence relations of Lee [67] can be established.

Generalized Langevin Equation for nonlinear system

In the previous sections, we used Chorin's projection and Mori's projection to derive the GLE for linear systems. In fact, similar series expansion of the GLE can be derived for general nonlinear dynamical systems by using Mori's projection. It is not hard to verify that under Mori's projection, $\mathcal{P}\mathcal{L}(\mathcal{Q}\mathcal{L})^i u_0 \in \{A_j\}_{j=1}^N$, where $\{A_j\}_{j=1}^N$ is a finite dimensional linear subspace and A_j are phase variables. As a consequence, for arbitrary $j-$th order operator polynomial $\Phi_j(\mathcal{Q}\mathcal{L})$, $\mathcal{P}\mathcal{L}\Phi_j(\mathcal{Q}\mathcal{L})u_0$ also lies in this linear space. Under this setting, if we choose $u_0 = A_k \in \{A_j\}_{j=1}^N$, $\mathcal{P}e^{s\mathcal{L}}\mathcal{P}\mathcal{L}\Phi_i(\mathcal{Q}\mathcal{L})A_k$ admits a matrix representation $\boldsymbol{M}(\Phi_i)\boldsymbol{a}(s)$, where $\boldsymbol{a}(s) = [A_1(s), ..., A_N(s)]^T$, $A_k(s) = \mathcal{P}e^{s\mathcal{L}}A_k$. $\boldsymbol{M}(\Phi_i)$ are $N \times N$ coefficient matrices for different Φ_i. Furthermore, we have

$$\mathcal{P}e^{s\mathcal{L}}\mathcal{P}\mathcal{L}e^{(t-s)\mathcal{Q}\mathcal{L}}\mathcal{Q}\mathcal{L}A_k = \sum_{i=1}^{n} g_i(t-s)\mathcal{P}e^{s\mathcal{L}}\mathcal{P}\mathcal{L}\Phi_i(\mathcal{Q}\mathcal{L})A_k = \sum_{j=1}^{N}\sum_{i=1}^{n} g_i(t-s)M_{kj}^{(i)}A_j(s)$$

where $g_j(t-s)$ can be any one of the temporal bases listed in Table 5.1, depending on the types of different MZ-expansion. This leads to the following matrix form GLE

$$\frac{d}{dt}A_k(t) = \sum_{j=1}^{N}\Omega_{jk}A_j(s) + \sum_{j=1}^{N}\int_0^t K_{kj}(t-s)A_j(s)ds \tag{5.39}$$

where $\mathcal{P}\mathcal{L}A_k = \sum_{j=1}^N \Omega_{jk}A_j$ yields the streaming matrix $\boldsymbol{\Omega}$ and the memory kernel matrix $\boldsymbol{K}(t-s)$ is given by $K_{kj}(t-s) = \sum_{i=1}^n g_i(t-s)M_{kj}^{(i)}$. To be noticed that, for nonlinear dynamical systems calculating the coefficient matrix $M_{kj}^{(i)}$ analytically is hard, especially for large i (the difficulty can be seen in Stinis's series work [86, 85, 87]). They are different cumulants (moments) of the random initial condition $x_0 \sim \rho_0$ [83], which can be computed in a data-driven setting [57, 6]. As we pointed out, different series expansion of the orthogonal dynamics propagator have different convergence rates, and can yield different errors in the MZ memory approximation for the same polynomial order n. From both theoretical and computational point of view, the Faber series expansion is superior than the commonly used Taylor series expansion. To see this, without loss of generality we consider the one-dimensional GLE where the MZ memory kernel is only a scalar function. The truncated Faber series

$$K(t) \simeq \sum_{q=1}^{n} M_q e^{tc_0} \frac{J_q(2t\sqrt{-c_1})}{(\sqrt{-c_1})^q} \tag{5.40}$$

has the following merits in representing the MZ memory kernel:

1. According to the properties of the Bessel function, Faber series (5.40) is uniformly bounded and converges to 0 as $t \to +\infty$ if modeling coefficient $c_0 \leq 0$. This guarantees the numerical stability of the approximation scheme and its long-time accuracy in describing the behavior of dynamical observables (such as time autocorrelation function) in ergodic systems. The Taylor (Dyson) series expansion obviously has no such properties.

2. As we will see immediately, the Faber series expansion converges faster than the Taylor series expansion.

3. Although formally more complicated, as we will show immediately once the expansion coefficients of the Taylor series are determined form data driven or first principle method, the expansion coefficient M_q of the Faber series can be obtained exactly through some recurrence formula.

4. When compared with other formally exact memory approximation schemes such as the Padé approximation used in [57] and the continued fraction introduced by Mori and Lee [67, 56], Faber series approximation (5.40) is

conducted in the temporal domain t instead of the Laplace transformed frequency domain s. This avoids the compuation of inverse Laplace transform which is numerically unstable.

Generally speaking, the orthogonal semigroup expansion theory provides a systematic way to derive different memory kernel approximation schemes. A lot of phenomenlogical models of the memory kernel can be incorporated into the framework. For instance, the trigonometric series used in [6] exactly is temporal modes corresponding to MZ-Lagrange expansion.

5.2 Convergence Analysis

In this section, we develop a thorough convergence analysis of the MZ-Faber expansion[3] of the Mori-Zwanzig equation (2.6). The key theoretical results at the basis of our analysis can be found in Chapter 3. Here we focus, in particular, on high-dimensional linear systems in the form

$$\dot{\boldsymbol{x}}(t) = \boldsymbol{A}\boldsymbol{x}(t), \qquad \boldsymbol{x}(0) = \boldsymbol{x}_0(\omega), \tag{5.41}$$

where $\boldsymbol{x}_0(\omega)$ is a random initial state. Our goal is to prove that the norm of the approximation error

$$\begin{aligned} E_n(t) &= \int_0^t \mathcal{P}e^{s\mathcal{L}}\mathcal{P}\mathcal{L}e^{(t-s)\mathcal{Q}\mathcal{L}}\mathcal{Q}\mathcal{L}u_0 ds - \underbrace{\sum_{j=0}^{n}\int_0^t a_j(t-s)\mathcal{P}e^{s\mathcal{L}}\mathcal{P}\mathcal{L}\mathcal{F}_j(\mathcal{Q}\mathcal{L})\mathcal{Q}\mathcal{L}u_0 ds}_{\text{MZ-Faber series}} \\ &= \int_0^t \sum_{j=n+1}^{\infty} a_j(t-s)\mathcal{P}e^{s\mathcal{L}}\mathcal{P}\mathcal{L}\mathcal{F}_j(\mathcal{Q}\mathcal{L})\mathcal{Q}\mathcal{L}u_0 ds \end{aligned} \tag{5.42}$$

[3] We recall that the MZ-Dyson series expansion is a subcase of the MZ-Faber expansion. Therefore convergence of MZ-Faber implies convergence of MZ-Dyson.

decays as we increase the polynomial order n, for any fixed integration time $t > 0$, i.e.,

$$\lim_{n\to\infty} \|E_n(t)\| = 0.$$

Throughout this Section $\|\cdot\|$ denotes either an operator norm, a norm in a function space or a standard norm in $\mathbb{C}^N$, depending on the context. The convergence proof of MZ-Faber series clearly depends on the choice of the projection operator and the phase space function $u(\boldsymbol{x})$ (quantity of interest). In this Section, we consider

$$u(\boldsymbol{x}(t)) = x_1(t), \tag{5.43}$$

and Chorin's projection (2.7). Similar results can be obtained for Mori's projection. We begin with the following

Lemma 5.2.1. *Consider the linear dynamical system* (5.41) *and the phase space function* (5.43). *If we set $\mathcal{P}$ to be Chorin's projection* (2.7) *with arbitrary initial distribution ρ_0 (not necessarily i.i.d), $\mathcal{Q} = \mathcal{I} - \mathcal{P}$, $\mathcal{L} = \boldsymbol{A}\boldsymbol{x}\cdot\nabla$ and p_k an arbitrary polynomial of degree k, then we have the following operator polynomial equality*

$$\mathcal{P}\mathcal{L}p_k(\mathcal{Q}\mathcal{L})\mathcal{Q}\mathcal{L}x_1(0) = \left[\boldsymbol{b}\cdot p_k(\mathbf{M}_{11}^T)\boldsymbol{a}\right]x_1(0) + \left[p_k(\mathbf{M}_{11}^T)\mathbf{M}_{11}^T\boldsymbol{a}\right]\cdot\langle\boldsymbol{x}_{-1}(0)\rangle_{\rho_0},$$

where

$$\boldsymbol{x}_{-1}(0) = [x_2(0), x_3(0), \dots, x_N(0)]^T \qquad \boldsymbol{a} = [A_{12}, \dots, A_{1N}]^T, \qquad \boldsymbol{b} = [A_{21}, \dots, A_{N1}]^T,$$

and $\mathbf{M}_{11}$ is the matrix obtained from $\boldsymbol{A}$ by removing the first row and the first column.

Proof. By a direct calculation, it can be verified that

$$
\begin{aligned}
(\mathcal{Q}\mathcal{L})^n x_1(0) &= \left[\left(\boldsymbol{M}_{11}^T\right)^{n-1}\boldsymbol{a}\right]\cdot\left[\boldsymbol{x}_{-1}(0) - \langle\boldsymbol{x}_{-1}(0)\rangle_{\rho_0}\right],\\
\mathcal{L}(\mathcal{Q}\mathcal{L})^n x_1(0) &= \left[\boldsymbol{b}^T\left(\boldsymbol{M}_{11}^T\right)^{n-1}\boldsymbol{a}\right] x_1(0) + \left[\left(\boldsymbol{M}_{11}^T\right)^{n}\boldsymbol{a}\right]\cdot\boldsymbol{x}_{-1}(0),\\
\mathcal{P}\mathcal{L}(\mathcal{Q}\mathcal{L})^n\mathcal{Q}\mathcal{L}x_1(0) &= \left[\boldsymbol{b}^T\left(\mathbf{M}_{11}^T\right)^{n}\boldsymbol{a}\right] x_1(0) + \left[\left(\mathbf{M}_{11}^T\right)^{n}\mathbf{M}_{11}^T\boldsymbol{a}\right]\cdot\langle\boldsymbol{x}_{-1}(0)\rangle_{\rho_0}. \quad (5.44)
\end{aligned}
$$

Note that each entry of the vector $\langle\boldsymbol{x}_{-1}(0)\rangle_{\rho_0} = [\langle x_2(0)\rangle_{\rho_0}, ..., \langle x_N(0)\rangle_{\rho_0}]^T$ is $\langle x_i(0)\rangle_{\rho_0} = \mathcal{P}x_i(0)$ $(i = 2, ..., N)$. Thus, for any polynomial function in the form

$$
p_k(\mathcal{Q}\mathcal{L}) = \sum_{j=0}^{k}\beta_k(\mathcal{Q}\mathcal{L})^j, \tag{5.45}
$$

we have

$$
\begin{aligned}
\mathcal{P}\mathcal{L}p_k(\mathcal{Q}\mathcal{L})\mathcal{Q}\mathcal{L}x_1(0) =& \sum_{j=0}^{k}\beta_j\mathcal{P}\mathcal{L}(\mathcal{Q}\mathcal{L})^j\mathcal{Q}\mathcal{L}x_1(0),\\
=& \sum_{j=0}^{k}\beta_j\left(\left[\boldsymbol{b}^T\left(\mathbf{M}_{11}^T\right)^n\boldsymbol{a}\right]x_1(0) + \left[\left(\mathbf{M}_{11}^T\right)^n\mathbf{M}_{11}^T\boldsymbol{a}\right]\cdot\langle\boldsymbol{x}_{-1}(0)\rangle_{\rho_0}\right),\\
=& \left[\boldsymbol{b}\cdot p_k\left(\mathbf{M}_{11}^T\right)\boldsymbol{a}\right]x_1(0) + \left[p_k\left(\mathbf{M}_{11}^T\right)\mathbf{M}_{11}^T\boldsymbol{a}\right]\cdot\langle\boldsymbol{x}_{-1}(0)\rangle_{\rho_0}.
\end{aligned}
$$

This completes the proof of the Lemma. □

To prove convergence of MZ-Faber series we need two more Lemmas involving Faber polynomials in the complex plane (see Appendix 5.A).

Lemma 5.2.2. *Let γ be the capacity of $\Omega \subseteq \mathbb{C}$. If Ω is symmetric with respect to the real axis, then for any $R > \gamma$ the conformal map* (5.96) *satisfies*

$$
\psi(R) \leq \psi(\gamma) + R - \frac{\gamma^2}{R}.
$$

Proof. We first notice that

$$\psi(R) = \psi(\gamma) + \int_{\gamma}^{R} \psi'(t)dt.$$

By using Lemma 4.2 in [70], i.e.,

$$|\psi'(t)| \leq 1 + \left(\frac{\gamma}{|t|}\right)^2, \quad |t| > \gamma$$

we have

$$\psi(R) - \psi(\gamma) \leq |\psi(R) - \psi(\gamma)| = \left|\int_{\gamma}^{R} \psi'(t)dt\right| \leq \int_{\gamma}^{R} |\psi'(t)|dt = R - \frac{\gamma^2}{R},$$

which completes the proof. □

Next, consider an arbitrary matrix $\boldsymbol{A}$ and define the *field value* of $\boldsymbol{A}$ as

$$FV(\mathbf{A}) = \left\{\boldsymbol{z}^H \mathbf{A} \boldsymbol{z} : \boldsymbol{z} \in \mathbb{C}^N,\, \boldsymbol{z}^H \boldsymbol{z} = 1\right\}.$$

The field value of $\boldsymbol{A}$ is a subset of the complex plane. Also, denote the truncated Faber series of the exponential matrix $e^{t\boldsymbol{A}}$ as

$$\boldsymbol{P}_m(t) = \sum_{j=0}^{m} a_j(t)\mathcal{F}_j(\mathbf{A}). \tag{5.46}$$

With this notation, we have the following

Lemma 5.2.3. *Let $\Omega \subset \mathbb{C}$ be symmetric with respect to the real axis, convex and with capacity γ. Consider an $N \times N$ matrix $\mathbf{A}$ with spectrum $\sigma(\boldsymbol{A})$, and an $N \times 1$ vector $\boldsymbol{v}$. If $\sigma(\boldsymbol{A}) \subseteq \Omega$ and the field value $FV(\mathbf{A}) \subseteq \Omega(q)$ for some $q \geq \gamma$, then* *the approximation error*

$$\boldsymbol{e}_m(t-s)\boldsymbol{v} = e^{(t-s)\mathbf{A}}\boldsymbol{v} - \boldsymbol{P}_{(m-1)}(t-s)\boldsymbol{v} \qquad t \geq s$$

satisfies

$$\|\boldsymbol{e}_m(t-s)\boldsymbol{v}\| \leq C_3 e^{(t-s)E}\left(\frac{qe^{t-s}}{m}\right)^{m-1} \qquad m \geq 4q,$$

where

$$C_3 = C_3(v) = 8e\|\boldsymbol{v}\|q\left(1+\frac{1}{8q}\right) \quad \text{and} \quad E = 1+\psi(\gamma).$$

Proof. If $q \geq \gamma$ then we have, thanks to the convexity of Ω and the analyticity of the exponential function,

$$\|\boldsymbol{e}_m(t-s)\boldsymbol{v}\| \leq 8\|\boldsymbol{v}\|e\left(1+\frac{1}{8q}\right)m\left(\frac{q}{m}\right)^m \max_{|z|\in\Gamma(m)}\left|e^{(t-s)z}\right| \qquad m \geq 4q \tag{5.47}$$

(see Theorem 4.2 in [70]). On the other hand,

$$\max_{|z|\in\Gamma(m)}\left|e^{(t-s)z}\right| = e^{(t-s)\psi(m)} \qquad m \geq 4q. \tag{5.48}$$

By using Lemma 5.2.2 we have

$$\psi(m) \leq \psi(\gamma) + m - \frac{\gamma^2}{m} \leq \psi(\gamma) + m, \tag{5.49}$$

and therefore

$$e^{(t-s)\psi(m)} \leq e^{(t-s)(m-1)}e^{(t-s)(1+\psi(\gamma))} \quad m \geq 4q \geq \gamma. \tag{5.50}$$

Combining (5.47) , (5.48) and (5.50), we obtain

$$\|\boldsymbol{e}_m(t-s)\boldsymbol{v}\| \leq C_3 \exp((t-s)E) \left(\frac{qe^{t-s}}{m}\right)^{m-1}, \tag{5.51}$$

where

$$C_3 = 8e\|\boldsymbol{v}\|q\left(1+\frac{1}{8q}\right) \quad \text{and} \quad E = 1+\psi(\gamma).$$

□

At this point, we we have all elements to prove the following

Theorem 5.2.4. (Convergence of the MZ-Faber Expansion) *Consider the linear dynamical system* (5.41)*, the phase space function* (5.43) *and the projection operator* (2.7)*. The norm of the approximation error* (5.42) *satisfies*[4]

$$\|E_n(t)\| \leq K\left(\frac{q}{n+1}\right)^n \frac{e^{t\beta}-e^{t(E+n)}}{\beta-E-n} \qquad t\geq 0, \quad n \geq 4q, \tag{5.52}$$

where n is the Faber polynomial order, while q, K, β and E are suitable constants defined in the proof of the theorem.

Proof. We aim at finding an upper bound for

$$\|E_n(t)\| = \left\|\int_0^t \mathcal{P}e^{s\mathcal{L}} \sum_{j=n+1}^{\infty} a_j(t-s)\mathcal{P}\mathcal{L}\mathcal{F}_j(\mathcal{Q}\mathcal{L})\mathcal{Q}\mathcal{L}x_1(0)ds\right\|. \tag{5.53}$$

To this end, we fist notice that quantity $\mathcal{F}_j(\mathcal{Q}\mathcal{L})\mathcal{Q}\mathcal{L}$ is a $(j+1)$-th order operator

[4]It can be shown that the upper bound in (5.52) is always positive.

polynomial in $\mathcal{QL}$. Thus, we can apply Lemma 5.2.1 to obtain

$$\mathcal{PLF}_j(\mathcal{QL})\mathcal{QL}x_1(0) = \left[\boldsymbol{b}\cdot\mathcal{F}_j(\mathbf{M}_{11}^T)\boldsymbol{a}\right]x_1(0) + \left[\mathcal{F}_j(\mathbf{M}_{11}^T)\mathbf{M}_{11}^T\boldsymbol{a}\right]\cdot\langle\boldsymbol{x}_{-1}(0)\rangle_{\rho_0}. \tag{5.54}$$

Let us now set

$$\eta_n(t-s) = \left\|\sum_{j=n+1}^{\infty} a_j(t-s)\mathcal{PLF}_j(\mathcal{QL})\mathcal{QL}x_1(0)\right\|. \tag{5.55}$$

By using (5.54) and the Cauchy-Schwartz inequality we have

$$\eta_n(t-s) \leq C_4\left\|\sum_{j=n+1}^{\infty} a_j(t-s)\mathcal{F}_j(\boldsymbol{M}_{11}^T)\boldsymbol{a}\right\| + C_5\left\|\sum_{j=n+1}^{\infty} a_j(t-s)\mathcal{F}_j(\boldsymbol{M}_{11}^T)\mathbf{M}_{11}^T\boldsymbol{a}\right\|, \tag{5.56}$$

where $C_4 = \|\boldsymbol{b}^T\||x_1(0)|$, $C_5 = \|\langle\boldsymbol{x}_{-1}(0)\rangle_{\rho_0}\|$. The two sums in (5.56) represent the error in the Faber approximation of the matrix exponential $e^{(t-s)\boldsymbol{M}_{11}^T}$. In fact,

$$\boldsymbol{e}_{(n+1)}(t-s) = e^{(t-s)\boldsymbol{M}_{11}^T} - \sum_{j=1}^{n} a_j(t-s)\mathcal{F}_j(\boldsymbol{M}_{11}^T) = \sum_{j=n+1}^{\infty} a_j(t-s)\mathcal{F}_j(\boldsymbol{M}_{11}^T). \tag{5.57}$$

Combining (5.53), (5.55), (5.56) and (5.51) yields

$$\begin{aligned}
\|E_n(t)\| &\leq \int_0^t \eta_n(t-s)\left\|\mathcal{P}e^{s\mathcal{L}}\right\| ds,\\
&\leq \int_0^t \left(C_4\left\|\boldsymbol{e}_{(n+1)}(t-s)\boldsymbol{a}\right\| + C_5\left\|\boldsymbol{e}_{(n+1)}(t-s)\boldsymbol{M}_{11}^T\boldsymbol{a}\right\|\right)\|\mathcal{P}e^{s\mathcal{L}}\|ds,\\
&\leq \int_0^t Ke^{s\beta}e^{(t-s)(E+n)}\left(\frac{q}{n+1}\right)^n ds,\\
&\leq K\left(\frac{q}{n+1}\right)^n \frac{e^{t\beta}-e^{t(E+n)}}{\beta-E-n} \quad n\geq 4q.
\end{aligned} \tag{5.58}$$

Here we used the semigroup estimation $\|e^{s\mathcal{L}}\| \le We^{s\beta}$. The constants in (5.58) are

$$K = \|\mathcal{P}\|C_6W, \qquad C_6 = 2\max\{C_4C_3, C_5C_3^*\}, \qquad E = 1 + \psi(\gamma), \tag{5.59}$$

where

$$C_3 = 8e\|\boldsymbol{a}\|q\left(1 + \frac{1}{8q}\right), \qquad C_3^* = 8e\|\boldsymbol{M}_{11}^T\boldsymbol{a}\|q\left(1 + \frac{1}{8q}\right). \tag{5.60}$$

It can be shown that the upper bound (5.58) is always positive, and goes to zero as we send the Faber polynomial order n to infinity. This implies that

$$\lim_{n\to\infty}\|E_n(t)\| = 0, \tag{5.61}$$

i.e., the MZ-Faber expansion converges for any finite time $t \ge 0$. This completes the proof. □

Next, we estimate the convergence rate of the MZ-Faber expansion. To this end, let us define

$$R(t,n) = K\left(\frac{q}{n+1}\right)^n \frac{e^{t\beta} - e^{t(E+n)}}{\beta - E - n}, \quad n \ge 4q \tag{5.62}$$

the be the upper bound (5.58). We have the following

Corollary 5.2.4.1. (Convergence Rate of the MZ-Faber Expansion) *With the same the notation of Theorem 5.2.4, the MZ-Faber expansion converges at least R-superlinearly with the polynomial order, i.e.*

$$\lim_{n\to\infty}\frac{R(t,n+1)}{R(t,n)} = 0 \tag{5.63}$$

for any finite time $t \ge 0$.

Proof. By a direct calculation it is easy to verify that (5.63) holds true. In fact,

$$\frac{R(t,n+1)}{R(t,n)} = \frac{q}{n+2}\left(\frac{n+1}{n+2}\right)^n \frac{e^{t\beta}-e^{t(E+n+1)}}{e^{t\beta}-e^{t(n+E)}}\frac{\beta-E-n}{\beta-E-(n+1)} \tag{5.64}$$

Therefore[5],

$$\lim_{n\to+\infty}\frac{R(t,n+1)}{R(t,n)} = \lim_{n\to+\infty}\frac{qe^t}{n+2}\left(\frac{n+1}{n+2}\right)^n = 0, \qquad t<\infty. \tag{5.66}$$

□

By using asymptotic analysis we can show theoretically that also the MZ-Dyson expansion converge R-superlinearly. To this end, let us define the MZ-Dyson approximation error

$$E_n(t) = \int_0^t \mathcal{P}e^{s\mathcal{L}}\mathcal{P}\mathcal{L}e^{(t-s)\mathcal{Q}\mathcal{L}}\mathcal{Q}\mathcal{L}u_0 ds - \underbrace{\sum_{j=0}^{n}\int_0^t a_j(t-s)\mathcal{P}e^{s\mathcal{L}}\mathcal{P}\mathcal{L}(\mathcal{Q}\mathcal{L})^j\mathcal{Q}\mathcal{L}u_0 ds}_{\text{MZ-Dyson series}}. \tag{5.67}$$

By following the same steps we used in the proof of Theorem 5.2.4, we can bound the norm of (5.67) as

$$\|E_n(t)\| \leq F(t,n). \tag{5.68}$$

where

$$F(t,n) = C\frac{(At)^n}{(n+1)!} \qquad A, C \geq 0. \tag{5.69}$$

Such upper bound plays the same role as $R(t,n)$ in the MZ-Faber expansion of $E_n(t)$ (see Eqs. (5.58) and (5.62)). Taking the ratio between $F(t,n+1)$ and

[5] We recall that

$$\lim_{n\to+\infty}\left(\frac{n+1}{n+2}\right)^n = \frac{1}{e}. \tag{5.65}$$

$F(t,n)$ we obtain

$$\lim_{n\to\infty}\frac{F(t,n+1)}{F(t,n)}=\lim_{n\to\infty}\frac{At}{n+2}=0. \tag{5.70}$$

5.3 Numerical Examples

In this section, we demonstrate the accuracy and effectiveness of the MZ-Dyson and MZ-Faber expansion methods we developed in this paper in applications to prototype problems involving random wave propagation and harmonic chains of oscillators interacting on a Bethe lattice.

5.3.1 Random Wave Propagation

Consider the following initial/boundary value problem for the wave equation in an annulus with radii $r_1 = 1$ and $r_2 = 11$

$$\frac{\partial^2 w}{\partial t^2}=\frac{\partial^2 w}{\partial r^2}+\frac{1}{r}\frac{\partial w}{\partial r}+\frac{1}{r^2}\frac{\partial^2 w}{\partial \theta^2}, \tag{5.71}$$

where

$$w(t,r_1,\theta)=0, \qquad w(t,r_2,\theta)=0 \qquad w(0,r,\theta)=w_0(r,\theta;\omega), \qquad \frac{\partial w(0,r,\theta)}{\partial t}=0. \tag{5.72}$$

The field $w(t,r,\theta)$ represents the wave amplitude at time t, while $w_0(r,\theta;\omega)$ is the wave field at initial time, which is set to be random. We seek the for an approximation of the solution $w(t,r,\theta)$ in the form

$$w_N(t,r,\theta)=\sum_{n=1}^{N}\widehat{w}_n(t)\psi_n(r,\theta), \tag{5.73}$$

where $\psi_n(r,\theta)$ are standard trigonometric functions. The random wave field at initial time is represented as

$$w_0(r,\theta;\omega) = \sum_{n=1}^{M} \widehat{w}_n(0)\psi_n(r,\theta), \qquad M \leq N, \tag{5.74}$$

where $\widehat{w}_n(0)$ are i.i.d Gaussian random variables. We substitute (5.73) into (5.71) and impose that the residual is orthogonal to the space spanned by the basis $\{\psi_1, ..., \psi_N\}$ [40]. This yields the linear system

$$\frac{d^2}{dt^2}\widehat{\boldsymbol{w}}(t) = \boldsymbol{A}\widehat{\boldsymbol{w}}(t), \tag{5.75}$$

where $\boldsymbol{A}$ is an $N \times N$ matrix with entries

$$A_{mn} = \frac{\displaystyle\int_{r_1}^{r_2}\int_0^{2\pi}\left(\frac{\partial^2\psi_n}{\partial r^2} + \frac{1}{r}\frac{\partial \psi_n}{\partial r} + \frac{1}{r^2}\frac{\partial^2\psi_n}{\partial \theta^2}\right)\psi_m dr d\theta}{\displaystyle\int_{r_1}^{r_2}\int_0^{2\pi}\psi_m^2 dr d\theta}. \tag{5.76}$$

We are interested in building a *convergent* reduced-order model for the wave amplitude at a specific point within the annulus, e.g., where we placed a sensor. To this end, we transform the system (5.75) form the modal space to the nodal space defined by an interpolant of at N collocation points. Such transformation can be easily defined by evaluating (5.73) at a set of distinct collocation nodes $\boldsymbol{x}_n = (r_{i(n)}, \theta_{j(n)})$ $(n = 1, ..., N)$ within the annulus. This yields

$$\boldsymbol{w}(t) = \boldsymbol{\Psi}\widehat{\boldsymbol{w}}(t), \tag{5.77}$$

where $\boldsymbol{w}(t) = [w(t, \boldsymbol{x}_1), ..., w(t, \boldsymbol{x}_N)]^T$, while $\boldsymbol{\Psi}$ is the $N \times N$ transformation matrix defined as

$$\boldsymbol{\Psi} = \begin{bmatrix} \psi_1(\boldsymbol{x}_1) & \dots & \psi_N(\boldsymbol{x}_1) \\ \vdots & & \vdots \\ \psi_1(\boldsymbol{x}_N) & \dots & \psi_N(\boldsymbol{x}_N) \end{bmatrix}.$$

Differentiating (5.77) with respect to time we obtain

$$\frac{d}{dt}\begin{bmatrix} \boldsymbol{w} \\ \dot{\boldsymbol{w}} \end{bmatrix} = \begin{bmatrix} \mathbf{0} & \boldsymbol{I} \\ \boldsymbol{\Psi}\mathbf{A}\boldsymbol{\Psi}^{-1} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \boldsymbol{w} \\ \dot{\boldsymbol{w}} \end{bmatrix} = \boldsymbol{B} \begin{bmatrix} \boldsymbol{w} \\ \dot{\boldsymbol{w}} \end{bmatrix} \tag{5.78}$$

This system evolves from the random initial state

$$\boldsymbol{w}(0) = \boldsymbol{\Psi}\widehat{\boldsymbol{w}}(0), \qquad \frac{d\boldsymbol{w}(0)}{dt} = 0. \tag{5.79}$$

For convenience, we still use $\boldsymbol{w}$ to represent vector $[\boldsymbol{w}, \dot{\boldsymbol{w}}]^T$. In Figure 5.1 we plot the mean solution of the random wave equation for initial conditions in the form (5.74) with different number of modes.

Generalized Langevin Equation for the Mean Wave Amplitude We are interested in building a convergent reduced-order model for the mean wave amplitude at a specific point within the annulus, e.g., where we would like to place a sensor. Such a dynamical system can be constructed by using the Mori-Zwanzig formulation and Chorin's projection operator (2.7). In particular, let us define the quantity of interest as $u(\boldsymbol{w}) = w_1(t)$, i.e., the wave amplitude at the spatial point

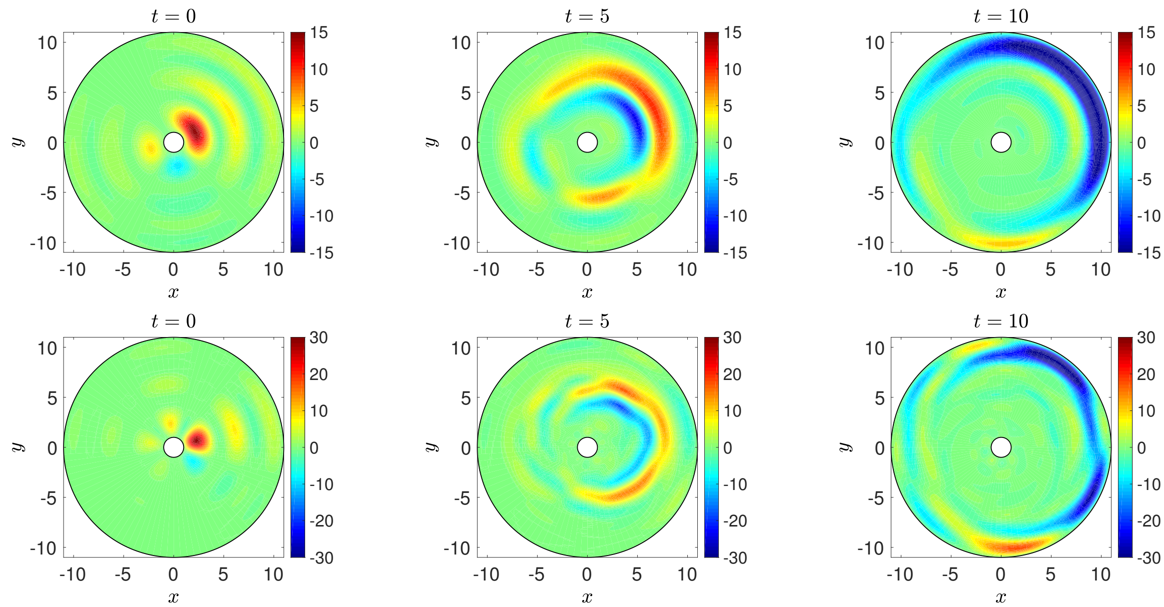

Figure 5.1: Mean solution of the random wave equation in the annulus. We consider two random initial conditions in the form (5.74), with different number of modes: $M = 25$ (first row), $M = 50$ (second row).

$(r, \theta) = (1.1, 0.1)$. The exact evolution equation for mean of $w_1(t)$ was derived in Section 5.1.4, and it is rewritten hereafter for convenience

$$\frac{d}{dt}\langle w_1(t)\rangle_{\rho_0} = a\langle w_1(t)\rangle_{\rho_0} + b + \int_0^t g(t-s)\langle w_1(s)\rangle_{\rho_0} ds + \int_0^t f(t-s)ds. \quad (5.80)$$

We recall that

$$\begin{aligned} \mathcal{PL}w_1(0) =& B_{11}w_1(0) + \boldsymbol{a}\cdot\langle \boldsymbol{w}_{-1}(0)\rangle_{\rho_0} \\ =& a w_1(0) + b, \end{aligned}$$

and $\mathcal{L} = \boldsymbol{Bw}\cdot\nabla$. The memory kernel $g(t-s)$ and the function $f(t-s)$ can be expanded by using in any of the operator series summarized in Table 5.1. For

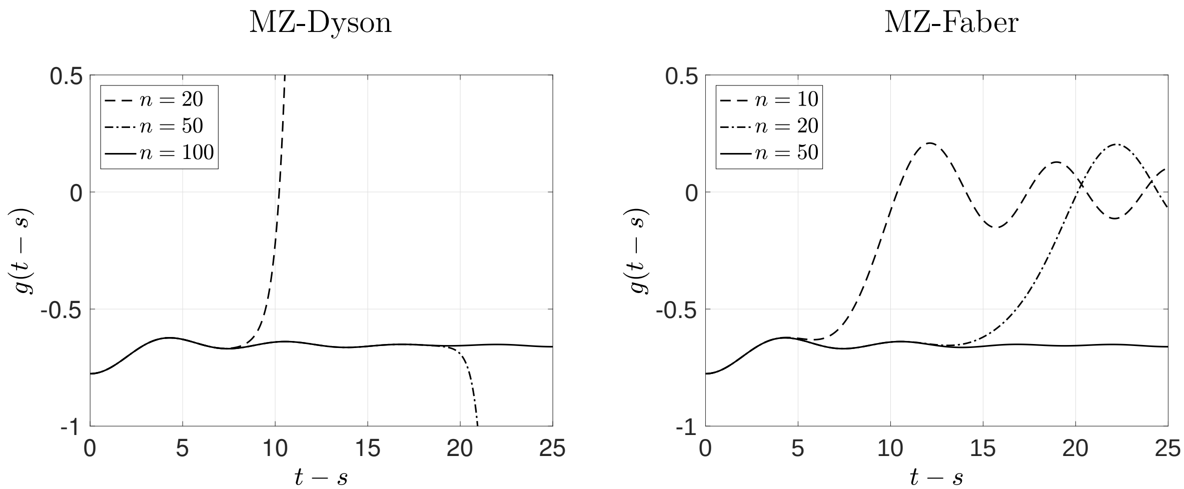

Figure 5.2: Dyson and Faber expansions of the Mori-Zwanzig memory kernel $g(t-s)$ in equation (5.81). Shown are results for different polynomial orders n. It is seen that the MZ-Faber series converges faster that the MZ-Dyson series.

instance, if we employ MZ-Faber series we obtain

$$g(t-s) = \sum_{j=0}^{n} g_j^F e^{tc_0} \frac{J_j(2t\sqrt{-c_1})}{(\sqrt{-c_1})^j}, \qquad f(t-s) = \sum_{j=0}^{n} f_j^F e^{tc_0} \frac{J_j(2t\sqrt{-c_1})}{(\sqrt{-c_1})^j}. \quad (5.81)$$

The coefficients g_j^F and f_j^F are explicitly obtained as

$$g_j^F = \boldsymbol{b}^T \mathcal{F}_j \left(\boldsymbol{M}_{11}^T \right) \boldsymbol{a}, \qquad f_j^F = \left[\mathcal{F}_j \left(\boldsymbol{M}_{11}^T \right) \mathbf{M}_{11}^T \boldsymbol{a} \right] \cdot \langle \boldsymbol{w}_{-1}(0) \rangle_{\rho_0}, \quad (5.82)$$

where

$$\boldsymbol{w}_{-1}(0) = [w_2(0), w_3(0), \dots, w_N(0)]^T, \boldsymbol{a} = [B_{12}, \dots, B_{1N}]^T, \boldsymbol{b} = [B_{21}, \dots, B_{N1}]^T,$$

$\mathbf{M}_{11}$ is the matrix obtained from $\boldsymbol{B}$ by removing the first row and the first column. In Figure 5.2 we study convergence of MZ-Dyson and MZ-Faber series expansions of the memory kernel. In Figure 5.3 we study the accuracy of the MZ-Dyson and the MZ-Faber expansions in representing the mean wave solution as a function of

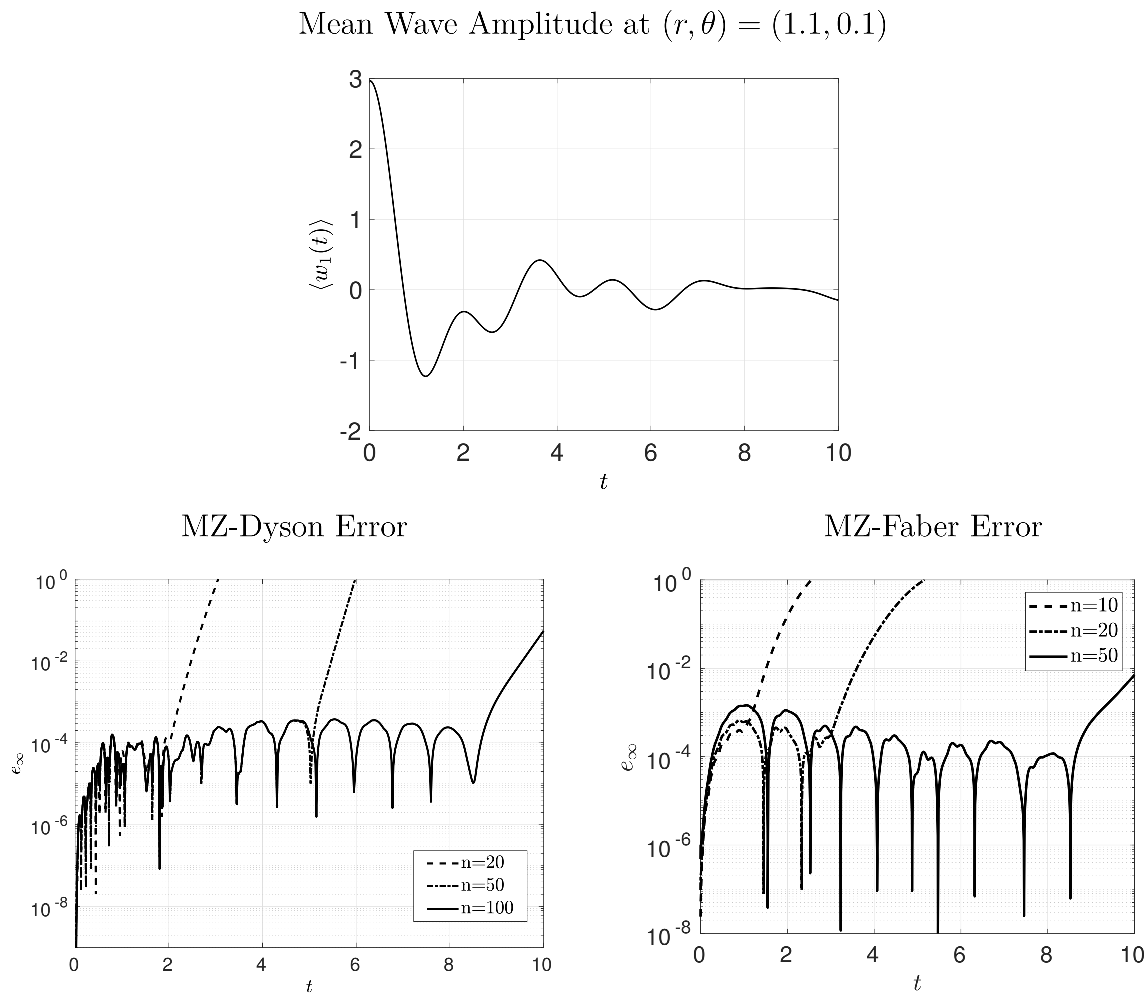

Figure 5.3: MZ-Dyson and MZ-Faber approximation errors of the mean wave amplitude at $(r, \theta) = (1.1, 0.1)$ as a function of the polynomial order n. It is seen that the MZ-Faber expansion converges faster than the MZ-Dyson series.

the polynomial order n. To this end, we solve (5.80) numerically a linear multi-step (explicit) time integration scheme (3rd-order Adams-Bashforth) combined with a trapezoidal rule to discretize the memory integral. As easily seen, that the MZ-Faber expansion converges faster than the MZ-Dyson expansion.

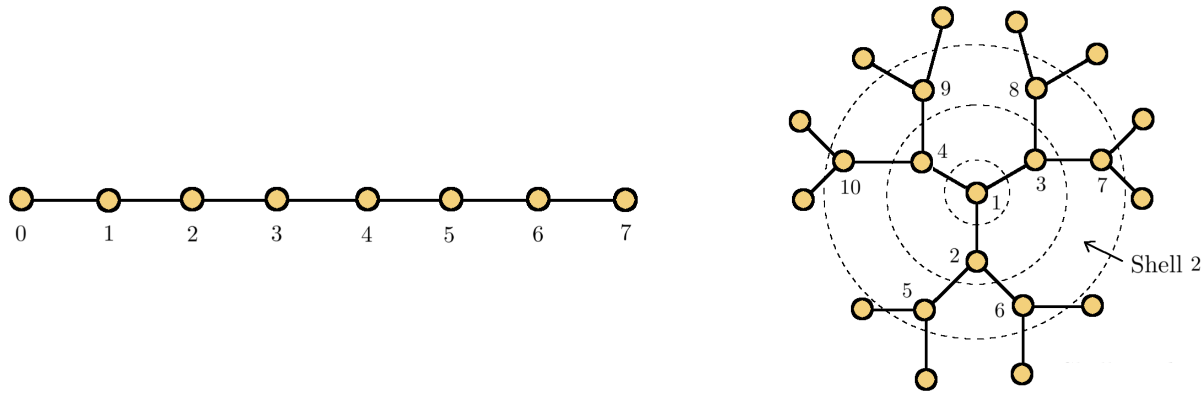

Figure 5.4: Bethe lattices with coordination numbers 2 (left), and 3 (right).

5.3.2 Harmonic Chains on the Bethe Lattice

Dynamics of harmonic chains on Bethe lattices is a simple but illustrative Hamiltonian dynamical system that has been widely studied in statistical mechanics, mostly in relation to Brownian motion [4, 33]. A Bethe lattice is a connected cycle-free graph in which each node interacts only with its neighbors. The number of such neighbors, is a constant of the graph called *coordination number*. This means that each node in the graph (with the exception of the leaf nodes) has the same number of edges connecting it to its neighbors. In Figure 5.4 we show two Bethe lattices with coordination numbers $l = 2$ and $l = 3$, respectively. The Bethe graph is hierarchical and therefore it can be organized into shells, emanating from an arbitrary node. The number of nodes in the k-th shell is given by $N_k = l(l-1)^{k-1}$,while the total number of nodes within S shells is

$$N = 1 + \sum_{k=1}^{S} N_k. \tag{5.83}$$

Next, we consider a coupled system of N harmonic oscillators[6] whose mutual interactions are defined by the adjacency matrix $\boldsymbol{B}^{(l)}$ of a Bethe graph with coordination number l [8]. The Hamiltonian of such system can be written as

$$H(\boldsymbol{p}, \boldsymbol{q}) = \frac{1}{2m} \sum_{i=1}^{N} p_i^2 + \frac{k}{2l} \sum_{i,j=1}^{N} B_{ij}^{(l)} (q_i - q_j)^2, \tag{5.84}$$

where q_i and p_i are, respectively, the displacement and momentum of the $i-$th particle, m is the mass of the particles (assumed constant throughout the network), and k is the elasticity constant that modulates the intensity of the quadratic interactions. We emphasize that the harmonic chain we consider here is one-dimensional. The Bethe graph basically just sets the interaction among the different oscillators. The dynamics of the harmonic chain on the Bethe lattice is governed by the Hamilton's equations

$$\frac{dq_i}{dt} = \frac{\partial H}{\partial p_i}, \qquad \frac{dp_i}{dt} = -\frac{\partial H}{\partial q_i}. \tag{5.85}$$

These equations can be written in a matrix-vector form as

$$\begin{bmatrix} \dot{\boldsymbol{p}} \\ \dot{\boldsymbol{q}} \end{bmatrix} = \begin{bmatrix} \boldsymbol{0} & k\boldsymbol{B}^{(l)} - k\boldsymbol{D}^{(l)} \\ \boldsymbol{I}/m & \boldsymbol{0} \end{bmatrix} \begin{bmatrix} \boldsymbol{p} \\ \boldsymbol{q} \end{bmatrix} = \boldsymbol{C} \begin{bmatrix} \boldsymbol{p} \\ \boldsymbol{q} \end{bmatrix}, \tag{5.86}$$

where $\boldsymbol{B}^{(l)}$ is the adjacency matrix of the graph and $\boldsymbol{D}^{(l)}$ is the degree matrix. Note that (5.86) is a linear dynamical system. The time evolution of any phase

[6] The number of oscillators cannot be set arbitrarily as it must satisfy the topological graph constraints prescribed by (5.83).

space function $u(\boldsymbol{q},\boldsymbol{p})$ (quantity of interest) satisfies

$$\frac{du}{dt} = \{u, H\},$$

where

$$\{u, H\} = \sum_{i=1}^{N} \left(\frac{\partial u}{\partial q_i} \frac{\partial H}{\partial p_i} - \frac{\partial H}{\partial q_i} \frac{\partial u}{\partial p_i} \right) \tag{5.87}$$

denotes the Poisson Bracket. A particular phase space function we consider hereafter is the velocity auto-correlation function of a tagged oscillator, say the one at location $j = 1$ (see Figure 5.4). Such correlation function is defined as

$$C_{p_1}(t) = \frac{\langle p_1(t) p_1(0) \rangle_{eq}}{\langle p_1(0) p_1(0) \rangle_{eq}}, \tag{5.88}$$

where the average is an integral over the Gibbs canonical distribution $\rho_{eq} \propto e^{-\beta H}$.

Analytical Expressions for the Velocity Autocorrelation Function

The simple structure of harmonic chains on the Bethe lattice allows us to determine analytical expressions for the velocity autocorrelation function (5.88), e.g., [4, 50, 33].

Bethe Lattice with Coordination Number 2 Let us set $l = 2$. In this case, the Bethe lattice is a a path graph, i.e., a one-dimensional chain of harmonic oscillators where each oscillator interacts only with the one at the left and at the right. We set fixed boundary conditions at the endpoint of the chain, i.e., $q_0(t) = q_{N+1}(t) = 0$ and $p_0(t) = p_{N+1}(t) = 0$ (particles are numbered from left to right). In this setting, the velocity auto-correlation function of the particle labeled with $j = 1$ can be obtained analytically by employing Lee's continued fraction method [33]. This yields the well-known $J_0 - J_4$ solution

$$C_{p_1}(t) = J_0(2\omega t) - J_4(2\omega t), \tag{5.89}$$

where $J_i(t)$ is the i-th Bessel function of the first kind, and $\omega = k/m$. Here we choose $k = m = 1$. The Hamilton's equations (5.86) for the inner oscillators (Here we exclude the two oscillators at the endpoints of the harmonic chain, since their dynamics is trivial) take the form

$$\begin{bmatrix} \dot{\boldsymbol{p}} \\ \dot{\boldsymbol{q}} \end{bmatrix} = \begin{bmatrix} \boldsymbol{0} & \boldsymbol{B}^{(2)} - \boldsymbol{D}^{(2)} \\ \boldsymbol{I} & \boldsymbol{0} \end{bmatrix} \begin{bmatrix} \boldsymbol{p} \\ \boldsymbol{q} \end{bmatrix}, \tag{5.90}$$

where $\boldsymbol{B}^{(2)}$ and $\boldsymbol{D}^{(2)}$ are the adjacency matrix and the degree matrix of the Bethe lattice with $l = 2$ (see Figure 5.4). As an example, if we consider $N = 5$ oscillators then $\boldsymbol{B}^{(2)}$ and $\boldsymbol{D}^{(2)}$ are given by

$$\boldsymbol{B}^{(2)} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}, \qquad \boldsymbol{D}^{(2)} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}. \tag{5.91}$$

Bethe Lattice with Coordination Number 3 Bethe graphs with $l = 3$ can be represented as planar graphs (see Figure 5.4). The velocity auto-correlation function at the center node can be expressed analytically [50], in the limit of an

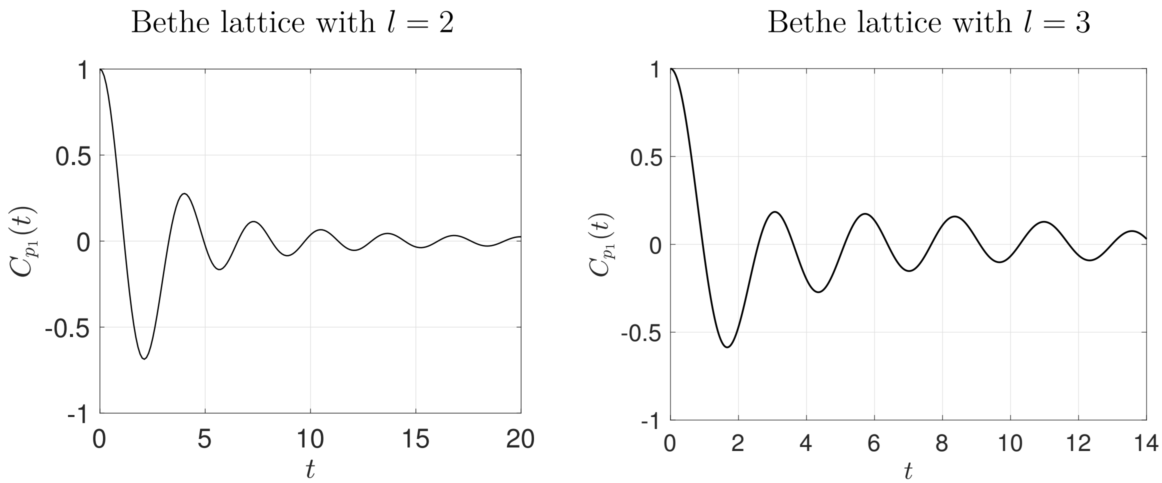

Figure 5.5: Velocity auto-correlation functions (5.89) (left) and (5.92) (right) of a tagged oscillator in an harmonic chain interacting on a Bethe lattice with coordination number $l = 2$ and $l = 3$, respectively.

infinite number of oscillators ($N \to \infty$) as

$$C_{p_1}(t) = \sum_{n=-\infty}^{+\infty} [G_n(t) + H_n(t)] J_{2n}(bt) \tag{5.92}$$

where

$$G_n(l) = \sum_{k=0}^{\infty} \frac{g_k(l)}{b^{2k-2}} \frac{1}{2\pi} \int_a^{\pi/2} d\theta \frac{\cos^2(\theta)}{\sin^{2k}(\theta)} \cos(2n\theta)$$
$$H_n(l) = \sum_{k=0}^{\infty} \frac{h_k(l)}{b^{-2k-2}} \frac{1}{2\pi} \int_a^{\pi/2} d\theta \frac{\cos^2(\theta)}{\sin^{-2k}(\theta)} \cos(2n\theta)$$

with $g_k(l)$ and $h_k(l)$ defined as

$$g_k(l) = -\sum_{j=k}^{\infty} \frac{(2j-1)!!}{[2^j(2j-1)j!]} a^{2j} c^{2(k-j)},$$
$$h_k(l) = -\sum_{j=k}^{\infty} \frac{(2j-1)!!}{[2^j(2j-1)j!]} a^{2(j-k)} c^{-2j}$$

and $a = \sqrt{2} - 1$, $b = \sqrt{2} + 1$ and $c = \sqrt{6}$. The Hamilton's equations of motion in this case are[7] ($k = m = 1$)

$$\begin{bmatrix} \dot{\boldsymbol{p}} \\ \\ \dot{\boldsymbol{q}} \end{bmatrix} = \begin{bmatrix} \mathbf{0} & \boldsymbol{B}^{(3)} - \boldsymbol{D}^{(3)} \\ \\ \boldsymbol{I} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \boldsymbol{p} \\ \\ \boldsymbol{q} \end{bmatrix}. \tag{5.93}$$

where $\boldsymbol{B}^{(3)}, \boldsymbol{D}^{(3)}$ are the adjacency matrix and the degree matrix of the Bethe lattice with $l = 3$ (see Figure 5.4). For example, if we label the oscillators as in Figure 5.4, and assume that the Bethe lattice has only three shells, i.e., 10 oscillators (4 inner nodes, and 6 leaf nodes) then the adjacency matrix and the

[7]Here we implemented a free boundary condition at the outer shell of the chain.

degree matrix are

$$\boldsymbol{B}^{(3)} = \begin{bmatrix} 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \boldsymbol{D}^{(3)} = \begin{bmatrix} 3 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 3 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 3 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

Generalied Langevin Equation for the Velocity Autocorrelation Function

The evolution equation for the velocity autocorrelation function (5.88) was obtained in Section 5.1.4 and it is hereafter rewritten for convenience

$$\frac{dC_{p_1}(t)}{dt} = aC_{p_1}(t) + \int_0^t g(t-s)C_{p_1}(s)ds. \tag{5.94}$$

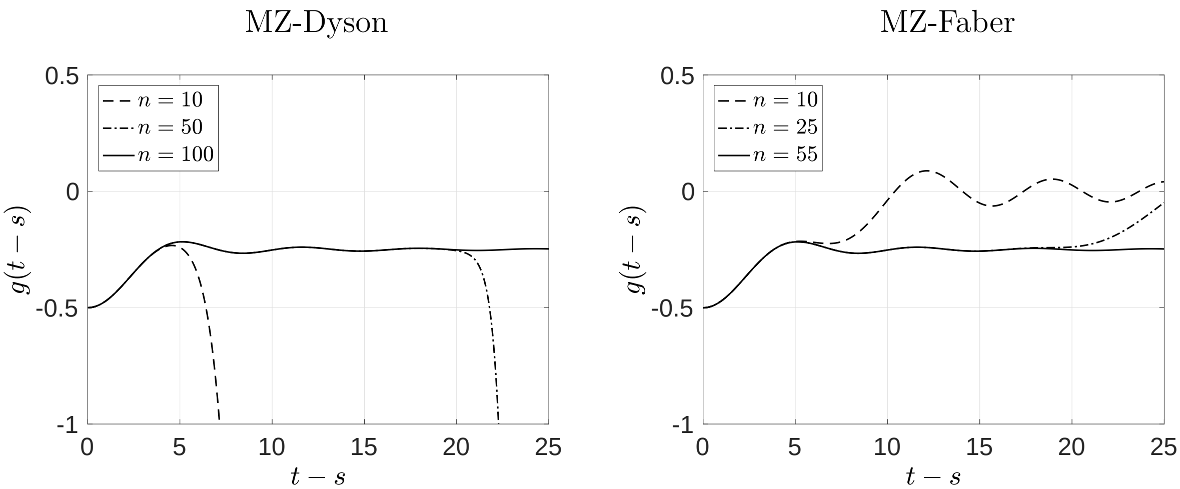

Figure 5.6: Harmonic chains of oscillators. Dyson and Faber expansions of the Mori-Zwanzig memory kernel $g(t-s)$. Shown are results for different polynomial orders n. It is seen that the MZ-Faber series converges faster that the MZ-Dyson series.

The initial condition is $C_{p_i}(0) = 1$. The MZ-Dyson and MZ-Faber series expansions of the the memory kernel $g(t-s)$ are given by

$$g(t-s) = \sum_{j=0}^{n} \frac{g_j^D}{j!}(t-s)^j, \qquad g(t-s) = \sum_{j=0}^{n} g_j^F e^{tc_0} \frac{J_j(2t\sqrt{-c_1})}{(\sqrt{-c_1})^j}$$

where

$$g_j^D = \boldsymbol{b}^T (\mathbf{M}_{11}^T)^j \boldsymbol{a}, \quad g_j^F = \boldsymbol{b}^T \mathcal{F}_j \left(\boldsymbol{M}_{11}^T \right) \boldsymbol{a}.$$

The definition of the matrix $\boldsymbol{M}_{11}^T$ and the vectors $\boldsymbol{a}$, $\boldsymbol{b}$ is the same as before. Here we used the fact that for any quadratic Hamiltonian we have $\langle p_i(0), q_i(0) \rangle_{eq} = 0$ and $\langle p_i(0), p_j(0) \rangle_{eq} = 0$ In Figure 5.6 we study convergence of the MZ-Dyson and the MZ-Faber series expansion of the memory kernel in equation (5.28). As before, the MZ-Faber series converges faster that the MZ-Dyson series. In Figure 5.7 and Figure 5.8, we study the accuracy of the MZ-Dyson and the MZ-Faber expansions

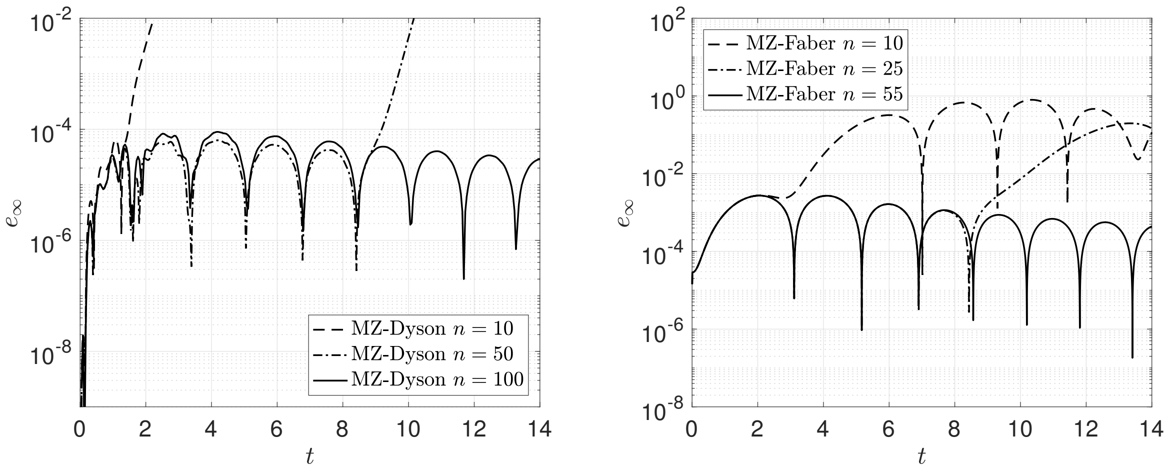

Figure 5.7: Accuracy of the MZ-Dyson and MZ-Faber expansions in representing the velocity auto-correlation function of the tagged oscillator $j = 2$ in an harmonic chain interacting on the Bethe lattice with coordination number 2. It is seen that the MZ-Dyson and the MZ-Faber expansions yield accurate predictions as we increase the polynomial order n. Moreover, the MZ-Faber expansion converges faster than the MZ-Dyson expansion.

in representing the velocity auto-correlation functions (5.89) and (5.92) (see Figure 5.5). Specifically, in these simulations we considered a chain of $N = 100$ oscillators for the case $l = 2$, and 8 shells of oscillators for the case $l = 3$, i.e., a total number of $N = 766$ oscillators. The results in Figure 5.7 and Figure 5.8 show that both the MZ-Dyson and the MZ-Faber expansions of the memory integral yield accurate approximations of the velocity autocorrelation function, and that convergence is uniform with the polynomial order. We emphasize that the new expansion of the MZ memory integral we developed can be employed to calculate phase space functions of harmonic oscillators on graphs with arbitrary topological structure. The following example shows the effectiveness of the proposed technique in calculating the velocity auto-correlation function of a tagged oscillator in a network sampled from the ErdösâĂŞRényi random graph.

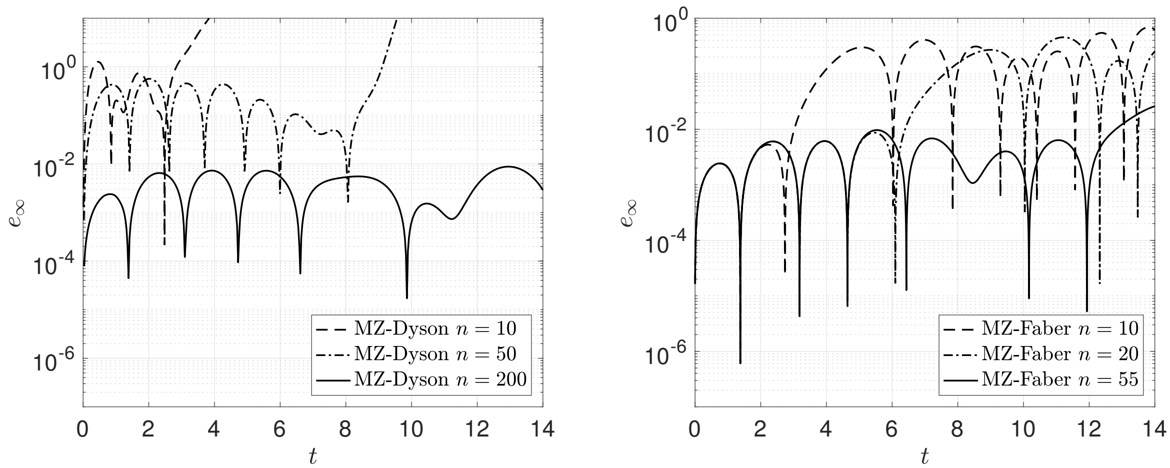

Figure 5.8: Accuracy of the MZ-Dyson and MZ-Faber expansions in representing the velocity auto-correlation function of the oscillator at the center of a Bethe lattice with coordination number 3, 8 shells and $N = 766$ oscillators. It is seen that the MZ-Dyson and the MZ-Faber expansion yield accurate predictions as we increase the polynomial order n. Moreover, the MZ-Faber expansion converges faster than the MZ-Dyson series.

Harmonic Chains on Graphs with Arbitrary Topology

In this section we consider an harmonic chain on a graph with arbitrary topology. The Hamiltonian function is

$$H = \frac{1}{2m}\sum_{i=1}^{N} p_i^2 + \frac{k}{2}\sum_{\substack{i,j=1\\ i<j}}^{N} J_{ij}(q_i - q_j)^2, \tag{5.95}$$

where J_{ij} is here is assumed is to be a sample from the Erdös-Rényi random adjacency matrix [9, 69]. In Figure 5.9 we study the accuracy of the MZ-Dyson and MZ-Faber expansions in approximating the velocity auto-correlation function of a tagged oscillator. In this case, no analytical solution is available and therefore we compared our solution to an accurate Monte Carlo benchmark. The lack of symmetry in each realization of the random network makes the velocity auto-correlation function dependent on the particular oscillator we consider.

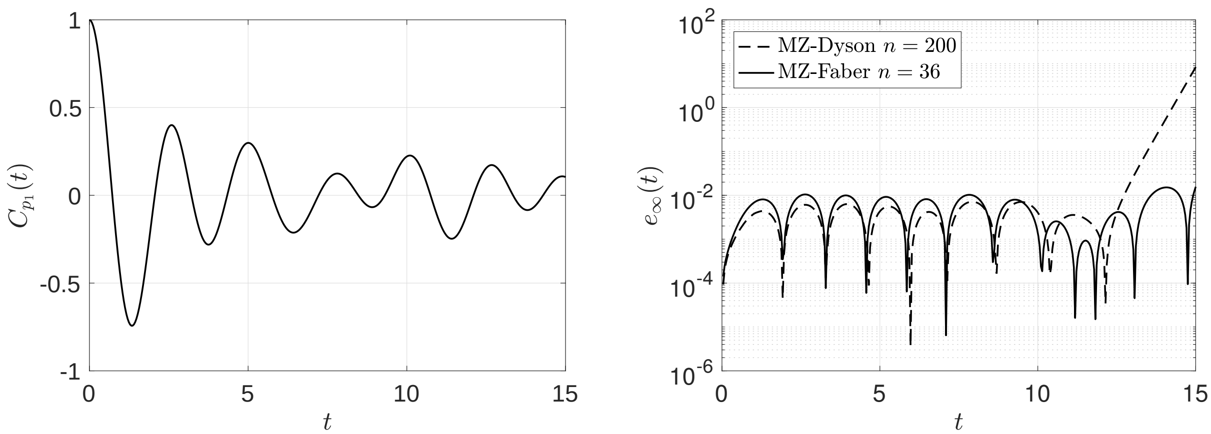

Figure 5.9: Aaccuracy of the MZ-Dyson and MZ-Faber expansions in approximating the velocity auto-correlation function of one tagged oscillator on a network obtained by sampling the Erdös-Rényi graph $G(100, 0.1)$. The benchmark solution is computed by Monte Carlo simulation.

5.4 Summary

In this chapter, we extended the series expansion method for matrix exponentials to the evolution operator for the orthogonal dynamics $e^{t\mathcal{QL}}$, and derived a new family of expansions for the Mori-Zwanzig equation. For Mori type projection operator, this yields series representation of the MZ memory kernel. Under the framework, not only some new approximations, such as the Faber series representation, can be obtained, some known phenomenological models for the memory kernel can also be derived from the first principle through the orthogonal dynamics expansion. For linear system, we provide an accurate convergence estimate for the Faber series expansion and proved that it converges to the exact result R-superlinearly fast. The whole theory is then tested on the well-known random wave propagation problem. We consider the linear wave on different geometries, including the annulus, Bethe lattice and the Erdös-Rényi random graph. For all these cases, the MZ-Faber expansion yields convergence results with a rate larger than the commonly used MZ-Dyson (or Taylor) expansion.

Appendix 5.A Faber Polynomials

In this appendix we briefly review the theory of Faber polynomials in the complex plane. Such polynomial were introduced by Faber in [31] (see [88] for a thorough review), and they play an important role in theory of univalent functions and in the approximation of matrix functions [70]. To introduce Faber polynomials, let

$$M = \{\Omega \subset \mathbb{C} :\, \Omega \neq \{\emptyset\} \text{ is compact and } \mathbb{C} \setminus \Omega \text{ is simply connected}\},$$

Given any set $\Omega \subset M$, by the Riemann mapping theorem there exits a conformal surjection

$$\psi : \hat{\mathbb{C}} \setminus \{w : |w| \leq \gamma\} \to \hat{\mathbb{C}} \setminus \Omega, \quad \psi(\infty) = \infty, \quad \psi'(\infty) = 1, \tag{5.96}$$

where $\hat{\mathbb{C}}$ is the Riemann sphere. The constant γ is called *capacity* of Ω. The $j-$th order Faber polynomial $\mathcal{F}_j(z)$ is defined to be the regular part of the Laurent expansion of $[\psi^{-1}(z)]^j$ at infinity, i.e.,

$$\mathcal{F}_j(z) := z^j + \sum_{k=0}^{j-1} \beta_{j,k} z^k, \quad j \geq 0. \tag{5.97}$$

Let Γ be the boundary of Ω. For $R \geq \gamma$ we define the equipotential curve $\Gamma(R)$ as

$$\Gamma(R) := \{z : \psi^{-1}(z) = R\}. \tag{5.98}$$

We also denote as $\Omega(R)$ the closure of the interior of $\Gamma(R)$. Obviously, if $R = \gamma$ then we have $\Omega(R) = \Omega$ and $\Gamma(R) = R$. Any analytic function $f(z)$ on Ω can be uniquely expanded in terms of Faber polynomials as

$$f(z) = \lim_{m\to\infty} f_m(z) \qquad f_m(z) = \sum_{j=0}^{m} a_j(f)\mathcal{F}_j(z), \tag{5.99}$$

where the coefficients $a_j(f)$ are given be the complex integral

$$a_j(f) = \frac{1}{2\pi i}\int_{|w|=R} \frac{f(\psi(w))}{w^{j+1}}dw. \tag{5.100}$$

It can be shown that $\mathcal{F}_j(z)$ satisfy the following recurrence relation

$$\begin{aligned}
\mathcal{F}_0(z) &= 1,\\
\mathcal{F}_1(z) &= z - c_0,\\
&\vdots\\
\mathcal{F}_j(z) &= (z-c_0)\mathcal{F}_{j-1}(z) - (c_1\mathcal{F}_{j-2}(z) + ... + c_{j-1}\mathcal{F}_0(z)) - (j-1)c_{j-1}, \quad j \geq 2,
\end{aligned} \tag{5.101}$$

where $c_0, c_1, ...$ are the coefficients of the Laurent series expansion of the mapping ψ, i.e.,

$$\psi(w) = w + c_0 + \frac{c_1}{w} + \frac{c_2}{w^2} + \cdots, \qquad |w| > \gamma \tag{5.102}$$

From a computational viewpoint, it is convenient to limit the number of terms in the expansion (5.102). In this way, we can simplify the recurrence relation (5.101), the calculation of (5.100) and therefore significantly speed up computations. In this paper we consider the map

$$\psi(w) = w + c_0 + \frac{c_1}{w}, \tag{5.103}$$

which transforms circles into ellipses. In this case, the coefficients (5.100) can be obtained analytically by computing the integral

$$
\begin{aligned}
a_j(t) &= \frac{1}{2\pi i}\int_{|w|=R} \frac{\exp\{t(w+c_0+c_1/w)\}}{w^{j+1}}dw, \\
&= \frac{1}{(\sqrt{-c_1})^j} e^{tc_0} J_j(2t\sqrt{-c_1}), \qquad (5.104)
\end{aligned}
$$

where $J_j(x)$ is the Bessel function of the first kind. The number of terms in the Laurent series expansion (5.102) should be selected so that the spectrum of the operator $\mathcal{QL}$ lies entirely within the equipotential curve (5.98). In the numerical examples we discuss in Section 6.3 such spectrum turns out to be relatively concentrated around the imaginary axis. Hence, the second-order truncation (5.103), which defines an elliptical equipotential curve, guarantees fast convergence of the Faber series expansion of the orthogonal dynamics propagator.

Appendix 5.B Faber Expansion of the Orthogonal Dynamics Propagator

Given any matrix representation of the operator $\mathcal{QL}$ (generator of the orthogonal dynamics) and a vector v, it is known that the sequence $f_m(\mathcal{QL})v$ (see equation (5.99)) converges to $f(\mathcal{QL})v$ for any analytic function $f(z)$ defined on Ω, provided the spectrum of $\mathcal{QL}$ is in Ω (see [**?**]). Moreover, by the properties of Faber polynomials, it is known that the sequence $f_m(\mathcal{QL})$ approximates asymptotically $f(\mathcal{QL})$ on Ω, as well as the sequence of best uniform approximation polynomials. In this sense, $f_m(\mathcal{QL})$ is said to be *asymptotically optimal* [27]. In particular, if we consider the exponential function $f(z) = e^{tz}$ and the conformal map (5.103), this yields the following m-th order Faber approximation of the orthogonal dynamics semigroup

$$e^{t\mathcal{QL}} \simeq \sum_{j=0}^{m} \frac{1}{(-c_1)^{j/2}} e^{tc_0} J_j(2t\sqrt{-c_1})\mathcal{F}_j(\mathcal{QL}). \tag{5.105}$$

Chapter 6

MZE for system with local interactions

The series expansion method of the orthogonal semigroup developed in Chapter 5 has shown to be a practical computational tool which can be used in dimension reduction. However, due to the complexity of the memory integral, the approximation is only valid for linear systems, which limits the applicability of the method subsetentially. For nonlinear systems with no scale separation, most effective methods are data-driven and the MZ equation is only used as an *ansatz* in this situation. In this chapter, build upon the Faber operator series expansion introduced in Chapter 5, we develop a new combinatorial algorithm that allows us to compute the MZ memory kernel from the first principle, i.e. by using only the structure of the microscopic equations of motion. In addition, we address the full MZ equation (2.5) and develop a new data-driven stochastic process representation method based on Karhunen-Loève (KL) series expansions, which allows us to build simple models of the MZ fluctuation term in systems with invariant measures, e.g., Hamiltonian systems or more general systems [10, 32]. The whole chapter is organized as follows. In Section 6.1 we develop series expansion of the

MZ memory kernel based on the Faber operator series. In particular, the subsection 6.1.1 is devoted to the description of an exact combinatorial algorithm to compute the recurrence coefficients of the MZ memory kernel expansion. In Section 6.2, we develop a new stochastic process representation method to build stochastic reduced-order models for quantities of interest (macroscopic observables) in systems in statistical equilibrium. In Section 6.3 we demonstrate the accuracy of the MZ memory calculation and the reduced-order stochastic modeling technique in applications to nonlinear random wave propagation described by Hamiltonian partial differential equations. We also include a brief Appendix where prove convergence of KL expansions in representing auto-correlation functions of polynomial observables.

6.1 Calculation of the MZ memory kernel from first principles

In this section we develop a new algorithm to calculate the MZ memory kernel (2.16c) based on first-principles, i.e., based on the microscopic equations of motion of the system (2.1). The algorithm we propose is built upon the combinatorial approach originally proposed by Amati, Meyer and Schilling in [1]. To illustrate the main idea in a simple way, hereafter we study the case where the observable $\boldsymbol{u}(t)$ is one-dimensional, i.e., we have only one phase space function $u(t) = u(\boldsymbol{x}(t, \boldsymbol{x}_0))$. In this setting, the series expansion of the memory kernel in eqn (5.39) reduces to

$$K(t) \simeq \sum_{q=0}^{n} g_q(t) M_q, \qquad \text{where} \qquad M_q = \frac{\langle u(0), \mathcal{L}\Phi_q(\mathcal{QL})\mathcal{QL}u(0)\rangle_\rho}{\langle u(0), u(0)\rangle_\rho}. \tag{6.1}$$

Note that $K(t)$ depends only on the set of parameters $\{M_0, \ldots, M_n\}$, since the temporal modes $g_q(t)$ are explicitly available given the polynomial set $\{\Phi_0, \ldots, \Phi_n\}$ (see Table 5.1). We aim at determining $\{M_0, \ldots, M_n\}$ from first principles. For one-dimensional phase space functions, Mori's projection (2.32) reduces to

$$\mathcal{P}f = \frac{\langle f, u(0)\rangle_\rho}{\langle u(0), u(0)\rangle_\rho} u(0). \tag{6.2}$$

At this point, it is convenient to introduce the following notation

$$\mu_i = \frac{\langle \mathcal{L}(\mathcal{QL})^{i-1} u(0), u(0)\rangle_\rho}{\langle u(0), u(0)\rangle_\rho}, \qquad \gamma_i = \frac{\langle \mathcal{L}^i u(0), u(0)\rangle_\rho}{\langle u(0), u(0)\rangle_\rho}. \tag{6.3}$$

Each coefficient μ_i represents the rescaling of $u(0)$ under the action of the operator $\mathcal{PL}(\mathcal{QL})^{i-1}$, i.e. we have

$$\mu_i u(0) = \mathcal{PL}(\mathcal{QL})^{i-1} u(0). \tag{6.4}$$

Clearly, if we are given $\{\mu_1, \ldots, \mu_{n+2}\}$ then we can easily compute M_q in (6.1), and therefore the memory kernel $K(t)$ for any given polynomial function Φ_q. For example, if $\Phi_q(\mathcal{QL}) = (\mathcal{QL})^q$ then $M_q = \mu_{q+2}$ $(q = 0, \ldots, n)$. On the other hand, if $\{\Phi_0, \ldots, \Phi_n\}$ are Faber polynomials [104], then we can write each Φ_q as a linear combination of monomials $(\mathcal{QL})^j$ $(j = 0, \ldots, q)$ and therefore represent M_q as a linear combination of $\{\mu_1, \ldots, \mu_{q+2}\}$. Computing μ_i using the definition (6.3) involves taking operator powers and averaging, which may be computationally expensive. An alternative effective algorithm relies on the following recursive formula [19, 83, 7]

$$\mu_1 = \gamma_1, \qquad \mu_2 = \gamma_2 - \mu_1\gamma_1, \qquad \cdots, \qquad \mu_n = \gamma_n - \sum_{j=1}^{n-1} \mu_{n-j}\gamma_j. \tag{6.5}$$

In practice, (6.5) shifts the problem of computing $\{\mu_1, \dots, \mu_n\}$ to the problem of evaluating the coefficients $\{\gamma_1, \dots, \gamma_n\}$ defined in (6.3). This will be discussed extensively in the subsequent Section 6.1.1. If the Liouville operator $\mathcal{L}$ is skew-adjoint relative to the inner product $\int \rho_0 dx$, then all μ_j and γ_j corresponding to odd indices are identically zero. This allows us to simplify the recursion (6.5) as

$$\mu_{2j} = \gamma_{2j} - \sum_{k=1}^{j-1} \mu_{2j-2k}\gamma_{2k} \qquad j = 1, 2, \dots. \tag{6.6}$$

As a consequence, the streaming term (2.16b) in the MZ equation vanishes identically since $\Omega = \mu_1 = \gamma_1 = 0$. We recall that skew-adjoint Liouville operators arise naturally, e.g., in Hamiltonian dynamical systems at statistical equilibrium.

6.1.1 Systems with polynomial nonlinearities

In this section, we address the problem of calculating the coefficients $\{\gamma_1, \dots, \gamma_n\}$ defined in (6.3) and appearing in the recursion relation (6.5). With such coefficients available, we can compute $\{\mu_1, \dots, \mu_n\}$ and therefore the MZ memory kernel (6.1). The calculation we propose is based on first principles, meaning that we do not rely on any assumption or model to evaluate the averages $\gamma_i = \langle \mathcal{L}^i u(0), u(0)\rangle_\rho / \langle u(0), u(0)\rangle_\rho$. Instead, we develop a combinatorial algorithm that allows us to track all terms in $\mathcal{L}^i u(0)$, hence representing γ_i exactly as a superimposition of a finite, although possibly large, number of terms. The algorithm we develop is built upon the combinatorial algorithm recently proposed by Amati, Meyer and Schiling in [1]. To describe the algorithm, consider the nonlinear dynamical system (2.1) and assume that $\boldsymbol{F}(\boldsymbol{x})$ is a multivariate polynomial in the phase variables $\boldsymbol{x}$. A simple example of such system is the Kraichnan-Orszag three-mode problem [71, 99, 11]

$$\dot{x_1} = x_1 x_3, \qquad \dot{x_2} = -x_2 x_3, \qquad \dot{x_3} = x_2^2 - x_1^2. \tag{6.7}$$

Other examples are the semi-discrete form of the Navier-Stokes equations, or the semi-discrete form of the nonlinear wave equation discussed in Section 6.3. The key observation to compute γ_j for systems with polynomial nonlinearities is that the action of the operator power $\mathcal{L}^i$ on a polynomial observable $u(\boldsymbol{x})$ yields a polynomial function. For instance, consider $u(\boldsymbol{x}) = x_1^3$, and the Liouville operator associated with the system (6.7)

$$\mathcal{L} = x_1 x_3 \frac{\partial}{\partial x_1} - x_2 x_3 \frac{\partial}{\partial x_2} + (x_2^2 - x_1^2)\frac{\partial}{\partial x_3}. \tag{6.8}$$

We have

$$\mathcal{L}x_1^3 = 3x_1^3 x_3, \tag{6.9}$$

$$\mathcal{L}^2 x_1^3 = 9x_1^3 x_3^2 + 3x_1^3 x_2^2 - 3x_1^5, \tag{6.10}$$

$$\mathcal{L}^3 x_1^3 = 27x_1^3 x_3^3 + 18x_1^3 x_2^2 x_3 - 18x_1^5 x_3 + 27x_1^3 x_2^2 x_3 - 6x_1^3 x_2^2 x_3 - 15x_1^5 x_3. \tag{6.11}$$

Clearly, the number of terms in $\mathcal{L}^i x_1^3$ can rapidly increase, if high-order powers of $\mathcal{L}$ are considered. For higher-dimensional systems with non-local interactions, i.e., for systems where each $F_i(\boldsymbol{x})$ $(i = 1, \ldots, N)$ depends on all components of $\boldsymbol{x}$, this problem is serious, and requires multi-core computer-based combinatorics to systematically track all terms in the expansion of $\mathcal{L}^i x_j^q$. Let us introduce the following notation

$$\mathcal{L}^n x_j^q = \sum_{\boldsymbol{b}_i \in B^{(n)}} a_{\boldsymbol{b}_i}^{(n)} x_{k_1}^{m_{k_1}^{(i)}} \cdots x_{k_r}^{m_{k_r}^{(i)}}, \tag{6.12}$$

where $\{a_{\boldsymbol{b}_i}^{(n)}\}$ are polynomial coefficients, and $\{m_{k_j}^{(i)}\}$ are polynomial exponents. The set of indexes representing the relevant phase phase variables appearing in $\mathcal{L}^n x_j^q$, i.e., $\{k_1, \ldots, k_r\}$, is collected in the index set $K(n, j) = \{k_1, \ldots, k_r\}$, which depends on n and j. For example, in (6.9)-(6.11) we have

$$K(1,1) = \{1,3\}, \qquad K(2,1) = \{1,2,3\}, \qquad K(3,1) = \{1,2,3\}. \tag{6.13}$$

Of course, for low-dimensional dynamical systems, the simplest choice for the relevant variables would be the complete set of variables $\{x_1, \cdots, x_N\}$. However, for high-dimensional systems with local interactions this choice could lead to unnecessary computations. In fact, it can be shown that the variables appearing in the polynomial $\mathcal{L}^n x_j^q$ are usually a (possibly small) subset of the phase variables if the system has local interactions. The vector $\boldsymbol{b}_i = [m_{k_1}^{(i)}, \cdots, m_{k_r}^{(i)}]$, collects the exponents of the i-th monomial appearing in the expansion (6.12). Similarly, $a_{\boldsymbol{b}_i}^{(n)}$ is the coefficient multiplying i-th monomial in (6.12). For example, in (6.9) and (6.10) we have, respectively,

$$\boldsymbol{b}_1 = [3,1], \qquad a_{\boldsymbol{b}_1}^{(1)} = 3,$$

$$\boldsymbol{b}_1 = [3,0,2], \quad \boldsymbol{b}_2 = [3,2,0], \quad \boldsymbol{b}_3 = [5,0,0], \quad a_{\boldsymbol{b}_1}^{(2)} = 9, \quad a_{\boldsymbol{b}_2}^{(2)} = 3, \quad a_{\boldsymbol{b}_3}^{(2)} = -3.$$

At this point, it is convenient to define the set of polynomial exponents $B^{(n)} = \{\boldsymbol{b}_1, \boldsymbol{b}_2, \cdots\}$, the set polynomial coefficients $A^{(n)} = \{a_{\boldsymbol{b}_1}^{(2)}, a_{\boldsymbol{b}_2}^{(2)}, \cdots\}$, and the combined index set

$$\mathcal{I}^{(n)} = \{A^{(n)}, B^{(n)}\}. \tag{6.14}$$

Clearly, $\mathcal{I}^{(n)}$ identifies uniquely the polynomial (6.12), i.e., there is a one-to-one correspondence between $\mathcal{I}^{(n)}$ and $\mathcal{L}^n x_j^q$. For example, in the case of (6.9)-(6.11) we have

$$\mathcal{I}^{(1)} =\{\underbrace{\{3\}}_{A^{(1)}}, \underbrace{\{[3,0,1]\}}_{B^{(1)}}\}, \tag{6.15}$$

$$\mathcal{I}^{(2)} =\{\underbrace{\{9,3,-3\}}_{A^{(2)}}, \underbrace{\{[3,0,2],[3,2,1],[5,0,0]\}}_{B^{(2)}}\}, \tag{6.16}$$

$$\mathcal{I}^{(3)} =\{\underbrace{\{27,18,-18,27,-6,-15\}}_{A^{(3)}}, \underbrace{\{[3,0,3],[3,2,1],[5,0,1],[3,2,1],[3,2,1],[5,0,1]\}}_{B^{(3)}}\}. \tag{6.17}$$

If we apply $\mathcal{L}$ to (6.12) we obtain

$$\begin{aligned} \mathcal{L}^{n+1} x_j^q =& \mathcal{L}\mathcal{L}^n x_j^q, \\ =& \mathcal{L} \sum_{\boldsymbol{b}_i \in B^{(n)}} a_{\boldsymbol{b}_i}^{(n)} x_{k_1}^{m_{k_1}^{(i)}} \cdots x_{k_r}^{m_{k_r}^{(i)}}, \\ =& \sum_{\boldsymbol{b}_i \in B^{(n+1)}} a_{\boldsymbol{b}_i}^{(n+1)} x_{k_1}^{m_{k_1}^{(i)}} \cdots x_{k_r}^{m_{k_r}^{(i)}}. \end{aligned} \tag{6.18}$$

Clearly, if we can compute the mapping $\mathcal{I}^{(n)} \xrightarrow{\mathcal{L}} \mathcal{I}^{(n+1)}$, induced by the action of the Liouville operator $\mathcal{L}$ to the polynomial (6.12) (represented by $\mathcal{I}^{(n)}$), then we can compute the *exact* series expansion of $\mathcal{L}^n x_j^q$ for arbitrary n. With such expansion available, we can immediately determine the coefficients γ_j in (6.3) by averaging over the probability density ρ as

$$\gamma_n = \frac{\langle \mathcal{L}^n x_j^q, x_j^q \rangle_\rho}{\langle x_j^q, x_j^q \rangle_\rho} = \sum_{\boldsymbol{b}_i \in B^{(n)}} a_{\boldsymbol{b}_i}^{(n)} \frac{\langle x_{k_1}^{m_{k_1}^{(i)}} \cdots x_{k_r}^{m_{k_r}^{(i)}} x_j^q \rangle_\rho}{\langle x_j^q, x_j^q \rangle_\rho}. \tag{6.19}$$

If the operator $\mathcal{L}$ is skew-adjoint in $L^2(\mathcal{M}, \rho)$, i.e., if $\langle \mathcal{L}f, g\rangle_\rho = -\langle f, \mathcal{L}g\rangle_\rho$, then we have

$$\gamma_{2n} = \frac{\langle \mathcal{L}^{2n} x_j^q, x_j^q\rangle_\rho}{\langle x_j^q, x_j^q\rangle_\rho} = (-1)^n \sum_{\boldsymbol{b}_i, \boldsymbol{b}_j \in B^{(n)}} a_{\boldsymbol{b}_j}^{(n)} a_{\boldsymbol{b}_i}^{(n)} \frac{\langle x_{k_1}^{m_{k_1}^{(i)}+m_{k_1}^{(j)}} \cdots x_{k_r}^{m_{k_r}^{(i)}+m_{k_r}^{(j)}}\rangle_\rho}{\langle x_j^q, x_j^q\rangle_\rho}. \tag{6.20}$$

Remark To guarantee numerical stability of the MZ-Faber expansion we often need to scale the Liouville operator $\mathcal{L}$ by a parameter $\delta < 1$ (see [104, 45]), i.e., we need to compute the Faber operator polynomial series relative to $\delta\mathcal{L}$. Correspondingly, the generalized Langevin equation (2.17) is solved on a time scale t/δ. In this setting, the coefficients (6.19) are also calculated relative to the rescaled Liouville operator $\delta\mathcal{L}$.

Remark Computing γ_j for linear systems reduces to a classical numerical linear algebra problem, i.e., computing the Rayleigh quotient of a matrix power. To show this, consider the N-dimensional linear system $\dot{\boldsymbol{x}} = \boldsymbol{A}\boldsymbol{x}$, evolving from the random initial state $\boldsymbol{x}_0 \sim \rho_0$ ($\boldsymbol{x}_0$ column vector). Suppose we are interested in the first component of the system, i.e., set the observable as $u(t) = x_1(t, \boldsymbol{x}_0)$. Define the linear subspace $V = \text{span}\{x_{01}, x_{02}, \cdots, x_{0N}\} \subset L^2(\mathcal{M}, \rho_0)$. Clearly $u(t) \in V$ for all $t \geq 0$ [103, 104]. This allows us to calculate γ_j as

$$\gamma_j = \langle \left[\boldsymbol{A}^T\right]^j \boldsymbol{x}_0 \cdot \boldsymbol{e}_1\rangle_{\rho_0}, \qquad j = 1, \ldots, n \tag{6.21}$$

where $\boldsymbol{e}_1 = [1, 0, \ldots, 0]^T$.

Mapping the index set $\mathcal{I}^{(n)}$

In this section we describe the algorithm that allows us to compute the polynomial $\mathcal{L}^{n+1}x_k^q$ given the polynomial $\mathcal{L}^n x_k^q$ and the Liouville operator $\mathcal{L}$, i.e., the mapping defined by equation (6.18). This is equivalent to develop a set of algebraic rules to transform the combined index set $\mathcal{I}^{(n)}$ defined in (6.14) into $\mathcal{I}^{(n+1)}$, for arbitrary n. Once such rules are known, we can apply them recursively to compute the polynomial sequence

$$x_j^q \to \mathcal{L}x_j^q \to \mathcal{L}^2 x_j^q \to \mathcal{L}^3 x_j^q \to \cdots \to \mathcal{L}^n x_j^q$$

to any desired order. In this way, we can determine γ_n through (6.19) (or (6.20)), μ_n through (6.5) (or (6.6)), and therefore the MZ memory kernel (6.1). Before formulating the algorithm in full generality, it is useful to examine how it operates in a concrete example. To this end, consider again the Kraichnan-Orszag system (6.7), and the transformation between the polynomials (6.10) and (6.11) defined by the action of the Liouville operator (6.8). We are interested in formulating such transformation in terms of a set of algebraic operations mapping the index set $\mathcal{I}^{(2)}$ into $\mathcal{I}^{(3)}$ (Eqs. (6.16)-(6.17)). We begin by decomposing the three-dimensional Liouville operator (6.8) as

$$\mathcal{L} = \mathcal{L}_1 + \mathcal{L}_2 + \mathcal{L}_3, \quad \text{where} \quad \mathcal{L}_1 = x_1 x_3 \frac{\partial}{\partial x_1}, \quad \mathcal{L}_2 = -x_2 x_3 \frac{\partial}{\partial x_2}, \quad \mathcal{L}_3 = (x_2^2 - x_1^2)\frac{\partial}{\partial x_3}. \tag{6.22}$$

The action of $\mathcal{L}_i$ on any monomial generates a polynomial with S_i terms. In the present example, we have $S_1 = S_2 = 1$ and $S_3 = 2$. Let us now consider the first monomial in (6.10), i.e., $9x_1^3 x_3^2$. Such monomial is represented by the first element of $A^{(2)}$ and $B^{(2)}$ in (6.17). The corresponding combined set is $\{9, [3, 0, 2]\}$. At this point, we apply the operators $\mathcal{L}_1$, $\mathcal{L}_2$ and $\mathcal{L}_3$ to the polynomial $\{9, [3, 0, 2]\}$. This yields

$$\underbrace{\{9,[3,0,2]\}}_{9x_1^3x_3^2} \xrightarrow{\mathcal{L}_1} \underbrace{\{27,[3,0,3]\}}_{27x_1^3x_3^3}, \tag{6.23}$$

$$\underbrace{\{9,[3,0,2]\}}_{9x_1^3x_3^2} \xrightarrow{\mathcal{L}_2} \underbrace{\{0,[3,-1,2]\}}_{0}, \tag{6.24}$$

$$\underbrace{\{9,[3,0,2]\}}_{9x_1^3x_3^2} \xrightarrow{\mathcal{L}_3} \underbrace{\{18,[3,2,1]\} \uplus \{-18,[5,0,1]\}}_{18x_1^3x_2^2x_3 - 18x_1^5x_3} = \{\{18,-18\},\{[3,2,1],[5,0,1]\}\}. \tag{6.25}$$

The transformation associated with $\mathcal{L}_3$ generates the sum of two monomials, namely $18x_1^3x_2^2x_3 - 18x_1^5x_3$, which we represent as a union between two index sets. Such union, here denoted as $\uplus$, is an ordered union that pushes to the left polynomial coefficients and to the right polynomial exponents. Proceeding in a similar manner for all other monomials in (6.10) and taking ordered unions of all sets, yields the desired mapping $\mathcal{I}^{(2)} \to \mathcal{I}^{(3)}$. Let us now examine the action of a more general Liouville operator

$$\mathcal{L}_j = z x_1^{c_1} \cdots x_N^{c_N} \frac{\partial}{\partial x_j} \tag{6.26}$$

on the monomial $a x_1^{m_1} \cdots x_N^{m_N}$ represented by the index set $\{a,[m_1,\ldots,m_N]\}$. We have

$$\{a,[m_1,\ldots,m_N]\} \xrightarrow{\mathcal{L}_j} \{z m_j a, [m_1+c_1,\ldots,m_j+c_j-1,\ldots,m_N+c_N]\}. \tag{6.27}$$

This defines two linear transformations: a scaling transformation in the space of coefficients, and an addition in the space of exponents

$$a \to (zm_1)a, \qquad [m_1, \ldots, m_N] \to [m_1, \ldots, m_N] + [c_1, \ldots, c_j - 1, \ldots, c_N]. \quad (6.28)$$

In a vector notation, upon definition of $\boldsymbol{b} = [m_1, \ldots, m_N]$, $\boldsymbol{\theta}_j = [c_1, \ldots, c_j - 1, \ldots, c_N]$ and $\alpha_j = zm_j$, we can write (6.28) in compact form as

$$a \to \alpha_j a, \qquad \boldsymbol{b} \to \boldsymbol{b} + \boldsymbol{\theta}_j. \quad (6.29)$$

Let us know consider the general case where the Liouville operator is defined as

$$\mathcal{L}(\boldsymbol{x}) = \sum_{k=1}^{N} \mathcal{L}_k(\boldsymbol{x}) \qquad \mathcal{L}_k(\boldsymbol{x}) = F_k(\boldsymbol{x}) \frac{\partial}{\partial x_k} \quad (6.30)$$

and $F_k(\boldsymbol{x})$ is a polynomial involving S_k monomials in either all variables $\{x_1, \ldots, x_N\}$ or a subset of them. The action of $\mathcal{L}$ on each monomial in (6.18) can be written as

$$\mathcal{L} x_{k_1}^{m_{k_1}^{(i)}} \ldots x_{k_r}^{m_{k_r}^{(i)}} = \sum_{q \in K(n,j)} \mathcal{L}_q x_{k_1}^{m_{k_1}^{(i)}} \ldots x_{k_r}^{m_{k_r}^{(i)}}, \quad (6.31)$$

where $K(n, j) = \{k_1, \ldots, k_r\}$ is the set of relevant variables at iteration n. The polynomial (6.31) involves $S_{k_1} + \cdots + S_{k_r}$ terms, each one of which can be explicitly constructed by applying the linear transformation rules (6.29). In summary, we have

$$\mathcal{I}^{(n+1)} = \biguplus_{q \in K(n,j)} \biguplus_{i=1}^{\#B^{(n)}} \biguplus_{s=1}^{S_q} \{\alpha_s^q a_{\boldsymbol{b}_i}^{(n)}, \boldsymbol{b}_i + \boldsymbol{\theta}_s^q\}, \quad (6.32)$$

where $\#B^{(n)}$ denotes the number of elements in $B^{(n)}$. Note that both α_s^s and $\boldsymbol{\theta}_s^q$ depend on $q \in K(n, j)$ (index set of relevant variables).

Remark The recursive algorithm summarized by formula (6.32) is a modified version of the algorithm originally proposed by Amati, Meyer and Schiling in [1]. The key idea is the same, i.e., to compute the expansion coefficients γ_n in (6.19) using polynomial differentiation. However, there are a few differences between our algorithm and the algorithm proposed in [1] which we emphasize hereafter. In [1], the index set $B^{(n)}$ is pre-computed using the so-called spreading operators. Essentially, for each n, the iterative scheme generates a new set of polynomial coefficients $A^{(n)}$, which is subsequently matched with the corresponding indexes in $B^{(n)}$. In our algorithm, the sets $B^{(n)}$ and $A^{(n)}$ are computed on-the-fly at each step of the recursion. By doing so, we avoid calculating the spreading operators. This, in turn, allows us to avoid using numerical tensors to store index sets, since in our formulation there is no matching procedure between the polynomial exponents and the polynomial coefficients. Another difference between the two algorithms is that we utilized a rescaled Liouville operator $\delta\mathcal{L}$ ($\delta \in \mathbb{R}$) to enhance numerical stability when computing the operator polynomials $\Phi_q(\mathcal{QL})$. The algorithm in [1], on the other hand, is based on a Taylor series expansion of the operator exponential $e^{t\mathcal{L}}$, with unscaled Liouville operator[1]

An example: the Fermi-Pasta-Ulam model

Consider a one-dimensional chain of N anharmonic oscillators with Hamiltonian

$$\mathcal{H}(\boldsymbol{p}, \boldsymbol{q}) = \sum_{j=0}^{N-1} \frac{p_j^2}{2m} + \sum_{j=1}^{N-1} V(q_j - q_{j-1}). \tag{6.33}$$

[1] In our recent work [104] (Section 3.1) we proved that a Taylor series of the orthogonal dynamical propagation $e^{t\mathcal{QL}}$ yields an expansion of the MZ memory integral that resembles the classical Dyson series in scattering theory.

In (6.33) $\{q_j, p_j\}$ are, respectively, the generalized coordinate and momentum of the j-th oscillator, while $V(q_i - q_{i-1})$ is the potential energy between two adjacent oscillators. Suppose that the oscillator chain is closed (periodic), i.e., that $q_0 = q_N$ and $p_0 = p_N$. Define the distance between two oscillators as $r_j = q_j - q_{j-1}$. This allows us to write the Hamilton's equations of motion as

$$\begin{cases} \dfrac{dr_j}{dt} &= \dfrac{1}{m}(p_i - p_{i-1}), \\ \dfrac{dp_j}{dt} &= \dfrac{\partial V(r_{j+1})}{\partial r_{j+1}} - \dfrac{\partial V(r_j)}{\partial r_j}. \end{cases}$$

The Liouville operator corresponding to this system is

$$\mathcal{L}(\boldsymbol{p}, \boldsymbol{r}) = \sum_{i=1}^{N-1} \left[\left(\frac{\partial V(r_{i+1})}{\partial r_{i+1}} - \frac{\partial V(r_i)}{\partial r_i} \right) \frac{\partial}{\partial p_i} + \frac{1}{m}(p_i - p_{i-1}) \frac{\partial}{\partial r_i} \right].$$

Setting $V(x) = \alpha x^2/2 + \beta x^4/4$ yields the well-known Fermi-Pasta-Ulam β-model [63], which we study hereafter. To this end, suppose we are interested in the distance between the oscillators j and $j-1$, i.e., in the polynomial observable $u(\boldsymbol{p}, \boldsymbol{r}) = r_j$. The action of $\mathcal{L}^n$ on r_j can be explicitly written as

$$\mathcal{L}^n r_j = \sum_{\boldsymbol{b}_i \in B^{(n)}} a^{(n)}_{\boldsymbol{b}_i} r_{k_1}^{m^{(i)}_{k_1}} \cdots r_{k_u}^{m^{(i)}_{k_u}} p_{l_1}^{s^{(i)}_{l_1}} \cdots p_{l_v}^{s^{(i)}_{l_v}}, \tag{6.34}$$

where $\{k_1, \dots, k_u\}$ and $\{l_1, \dots, l_v\}$ are the relevant degrees of freedom for the polynomials of $\boldsymbol{r}$ and $\boldsymbol{p}$, respectively, at iteration n. We can explicitly compute the sets of such relevant degrees of freedom as

$$K_r(n, j) = \left\{ j - \left\lfloor \frac{n}{2} \right\rfloor, \dots, j + \left\lfloor \frac{n}{2} \right\rfloor \right\} \quad L_p(n, j) = \left\{ j - \left\lfloor \frac{n+1}{2} \right\rfloor, \dots, j + \left\lfloor \frac{n-1}{2} \right\rfloor \right\}, \tag{6.35}$$

The action of the Liouville operator on each monomial appearing in (6.34) can be written as

$$\mathcal{L} r_{k_1}^{m_{k_u}^{(i)}} r_{k_u}^{m_{k_u}^{(i)}} p_{l_1}^{s_{l_1}^{(i)}} \cdots p_{l_v}^{s_{l_v}^{(i)}} = \sum_{v \in K_r(n,j)} \sum_{h \in L_p(n,j)} (\mathcal{L}_{r_v} + \mathcal{L}_{p_h}) r_{k_1}^{m_{k_1}^{(i)}} \cdots r_{k_u}^{m_{k_u}^{(i)}} p_{l_1}^{s_{l_1}^{(i)}} \cdots p_{l_v}^{s_{l_v}^{(i)}}, \tag{6.36}$$

where

$$\mathcal{L}_{r_v} = \frac{1}{m}(p_v - p_{v-1})\frac{\partial}{\partial r_v}, \qquad \text{and} \qquad \mathcal{L}_{p_h} = \left[\alpha(r_{h+1} - r_h) + \beta\left(r_{h+1}^3 - r_h^3\right)\right]\frac{\partial}{\partial p_h}. \tag{6.37}$$

By computing the action of $\mathcal{L}_{r_v}$ and $\mathcal{L}_{p_h}$ on the monomial $r_{k_1}^{m_{k_1}^{(i)}} \cdots r_{k_u}^{m_{k_u}^{(i)}} p_{l_1}^{s_{l_1}^{(i)}} \cdots {}_{l_v}^{s_{l_v}^{(i)}}$ we obtain explicit linear maps of the form (6.29), involving the polynomial exponents

$$\boldsymbol{b}_i = [\boldsymbol{m}^{(i)}, \boldsymbol{s}^{(i)}], \qquad \boldsymbol{m}^{(i)} = [m_{k_1}^{(i)}, \ldots, m_{k_u}^{(i)}], \qquad \boldsymbol{s}^{(i)} = [s_{l_1}^{(i)}, \ldots, s_{l_v}^{(i)}], \tag{6.38}$$

and the polynomial coefficients $a_{\boldsymbol{b}_i}^{(n)}$. With such maps available, we can transform the combined index set $\mathcal{I}^{(n)}$ (representing $\mathcal{L}^n r_j$) to $\mathcal{I}^{(n+1)}$ (representing $\mathcal{L}^{n+1} r_j$) using (6.32). Specifically, we obtain

$$\mathcal{I}^{(n+1)} = \mathcal{I}_r^{(n+1)} \biguplus \mathcal{I}_p^{(n+1)},$$

where

$$\mathcal{I}_r^{(n+1)} = \biguplus_{v\in K_r(n,j)} \biguplus_{i=1}^{\#B^{(n)}} \biguplus_{k=0}^{1} \left\{ m_v^{(i)}(-1)^k a_{\boldsymbol{b}_i}^{(n)}, [\boldsymbol{m}^{(i)} - \boldsymbol{e}_v, \boldsymbol{s}^{(i)} + \boldsymbol{e}_{v-k}] \right\},$$

$$\mathcal{I}_p^{(n+1)} = \biguplus_{h\in L_p(n,j)} \biguplus_{i=1}^{\#B^{(n)}} \biguplus_{k=0}^{1} \Big\{ \{ s_h^{(i)}(-1)^{k+1}\alpha a_{\boldsymbol{b}_i}^{(n)}, s_h^{(i)}(-1)^{k+1}\beta a_{\boldsymbol{b}_i}^{(n)} \},$$
$$\{ [\boldsymbol{m}^{(i)} + \boldsymbol{e}_{h+k}, \boldsymbol{s}^{(i)} - \boldsymbol{e}_h], [\boldsymbol{m}^{(i)} + 3\boldsymbol{e}_{h+k}, \boldsymbol{s}^{(i)} - \boldsymbol{e}_h] \} \Big\}.$$

6.2 Stochastic low-dimensional modeling

In previous Sections we discussed an algorithm to approximate the memory kernel in the MZ equation (2.33) or (2.17) based on the microscopic equations of motion (first-principle calculation). In this Section we construct a statistical model for fluctuation term $\boldsymbol{f}(t)$ appearing in (2.33), which will allow us to compute statistical properties of the quantity of interest $\boldsymbol{u}(t)$ beyond two-point correlations. A possible way to build such model is to expand (2.35d) in a finite-dimensional series[2] (see Eq. (5.2)) as

$$\boldsymbol{f}(t) \simeq \sum_{q=0}^{n} g_q(t)\Phi_q(\mathcal{QL})\mathcal{QL}\boldsymbol{u}(0), \tag{6.39}$$

and evaluate the coefficients $\Phi_q(\mathcal{QL})\mathcal{QL}\boldsymbol{u}(0)$ using the combinatorial approach discussed in Section 6.1.1. However, this may not be straightforward since the operator $\Phi_q(\mathcal{QL})\mathcal{QL}\boldsymbol{u}(0)$ determines a high-dimensional random field. An alternative approach is to ignore the mathematical structure of $\boldsymbol{f}(t)$, i.e., equation (2.16d) or the series expansions (6.39), and simply model $\boldsymbol{f}(t)$ as a stochastic process. In doing so, we need to make sure that the statistical properties of the

[2] Note that $\boldsymbol{f}(t)$ is a random process obtained by mapping the random initial state $\boldsymbol{u}(0) = \boldsymbol{u}(\boldsymbol{x}_0)$ forward in time using the orthogonal dynamics propagator $e^{t\mathcal{QL}(\boldsymbol{x}_0)}$.

reduced-order model, e.g., the equilibrium distribution, are consistent with the full model. Such consistency conditions carry over a certain number of constraints on the process $\boldsymbol{f}(t)$, which allow for its partial identification. As an example, consider the following MZ model recently proposed by Lei *et al.* in [57] to study the dynamics of a tagged particle in a large particle system

$$\begin{cases} \dot{\boldsymbol{q}} &= \dfrac{\boldsymbol{p}}{m} \\ \dot{\boldsymbol{p}} &= \boldsymbol{F}(\boldsymbol{q}) + \boldsymbol{d} \\ \dot{\boldsymbol{d}} &= \boldsymbol{B_0 d} - \boldsymbol{A}_0 \dfrac{\boldsymbol{p}}{m} + \boldsymbol{f}(t) \end{cases} \tag{6.40}$$

It was shown in [57] that if $\boldsymbol{f}(t)$ is Gaussian white noise with auto-correlation function

$$\langle \boldsymbol{f}(t)\boldsymbol{f}(t')\rangle = -\beta \left(\boldsymbol{B}_0\boldsymbol{A}_0 + \boldsymbol{A}_0\boldsymbol{B}_0^T\right)\delta(t-t'), \tag{6.41}$$

then the equilibrium distribution of the particle system has the Boltzmann-Gibbs form

$$\rho(\boldsymbol{p},\boldsymbol{q},\boldsymbol{d}) \propto \exp\left\{-\beta\left(\frac{1}{2m}|\boldsymbol{p}|^2 + \frac{1}{2}\boldsymbol{d}^T\boldsymbol{A}_0^{-1}\boldsymbol{d} + V(\boldsymbol{q})\right)\right\}, \tag{6.42}$$

$V(\boldsymbol{q})$ being the inter-particle potential. However, modeling $\boldsymbol{f}(t)$ as a Gaussian process does not provide satisfactory statistics in MZ equations built upon Mori's projection. In fact, equation (2.15) is linear and therefore the equilibrium distribution of $\boldsymbol{u}(t)$ (assuming it exists) under Gaussian noise $\boldsymbol{f}(t)$ will be necessarily Gaussian. In most applications, however, the equilibrium distribution of $\boldsymbol{u}(t)$ is strongly non-Gaussian. To overcome this difficulty Chu and Li [19] recently developed a multiplicative Gaussian noise model that generalizes (2.15) in the sense that it is not based on additive noise, and it allows for non-Gaussian responses.

In this section we propose a different stochastic modeling approach for $\boldsymbol{f}(t)$

based on bi-orthogonal representations random processes [93, 97, 92, 3, 2]. To illustrate the method, we study the case where the observable $\boldsymbol{u}(t)$ is real valued (one-dimensional) and square integrable. This allows us to develop the theory in a clear and concise way. We also assume that the system is in statistical equilibrium, i.e., that there exists an equilibrium distribution $\rho_{eq}(\boldsymbol{x})$ (or more generally an invariant measure) for the phase variables $\boldsymbol{x}(t, \boldsymbol{x}_0)$, and that $\boldsymbol{x}_0$ is sampled from such distribution. The MZ equation (2.15) for one-dimensional phase space functions $u(t) = u(\boldsymbol{x}(t, \boldsymbol{x}_0))$ reduces to

$$\frac{du(t)}{dt} = \Omega u(t) + \int_0^t K(t-s)u(s)ds + f(t), \tag{6.43}$$

where

$$\Omega = \frac{\langle u(0), \mathcal{L}u(0)\rangle_{eq}}{\langle u(0), u(0)\rangle_{eq}}, \qquad K(t) = \frac{\langle u(0), \mathcal{L}f(t)\rangle_{eq}}{\langle u(0), u(0)\rangle_{eq}}, \qquad f(t) = e^{t\mathcal{QL}}\mathcal{QL}u(0). \tag{6.44}$$

Since $u(t)$ is assumed to be a second-order random process in the time interval $[0, T]$, we can expand it in a truncated Karhunen-Loéve series

$$u(t) \simeq \overline{u} + \sum_{k=1}^{K} \sqrt{\lambda_k}\xi_k e_k(t), \qquad t \in [0, T] \tag{6.45}$$

where $\overline{u}$ denotes the mean of $u(t)$ relative to the equilibrium distribution[3], $\{\xi_1, \ldots, \xi_K\}$ are uncorrelated random variables ($\langle \xi_i \xi_j \rangle_{eq} = \delta_{ij}$), and $\{\lambda_k, e_k(t)\}$ ($k = 1, \ldots, K$) are, respectively, eigenvalues and eigenfunctions of the homogeneous Fredholm

[3]The mean of $u(t) = u(\boldsymbol{x}(t, \boldsymbol{x}_0))$ is necessarily time-independent at statistical equilibrium. In fact, at equilibrium we have that $\boldsymbol{x}_0 \sim \rho_{eq}$ implies that $\boldsymbol{x}(t) \sim \rho_{eq}$ for all $t \geq 0$. A statistically stationary process however, may not be stationary in phase space. Indeed, $\boldsymbol{x}(t)$ evolves in time, eventually in a chaotic way as it happens for systems with strange attractors and invariant measures.

integral equation of the second kind

$$\int_0^T \langle u(t)u(s)\rangle_{eq} e_k(s)ds = \lambda_k e_k(t), \qquad t \in [0,T]. \tag{6.46}$$

We recall that for ergodic systems in statistical equilibrium the auto-correlation function $\langle u(t)u(s)\rangle_{eq}$ decays to zero as $|t-s| \to \infty$. Also, the integral operator at the left hand side of (6.46) is positive-definite and compact [2]. The orthogonal random variables ξ_k and the temporal modes $e_k(t)$ are related to each other by the following dispersion relations [3, 93]

$$\xi_k = \frac{1}{\sqrt{\lambda_k}} \int_0^T u(t)e_k(t)dt, \qquad e_k(t) = \frac{\langle u(t)\xi_k\rangle_{eq}}{\sqrt{\lambda_k}} \qquad k = 1, 2, \ldots. \tag{6.47}$$

Equation (6.47) suggests that if $u(t)$ is a Gaussian random process (e.g., a Wiener process) then $\{\xi_1, \ldots, \xi_K\}$ are necessarily independent Gaussian random variables. On the other hand, if $u(t)$ is non-Gaussian then the joint distribution of $\{\xi_1, \ldots, \xi_K\}$ is unknown, although it can be in principle computed by using the transformation $u(t) \to \xi_k$ $(k = 1, .., K)$ defined in (6.47), given λ_k and $e_k(t)$.

An alternative approach to identify the PDF of $\{\xi_1, \ldots, \xi_K\}$ relies on sampling. In particular, as recently shown by Phoon *et al.* [78, 79], it is possible to develop effective sampling algorithms for the KL expansion (6.45). Such algorithms allow to sample the uncorrelated variables $\{\xi_1, \ldots, \xi_K\}$ in a way that makes the PDF of $u(t)$ consistent with the equilibrium distribution, which can be calculated by mapping $\boldsymbol{x}_0 \sim \rho_{eq}(\boldsymbol{x}_0)$ to $u(\boldsymbol{x}_0)$. At this point, we have available a consistent bi-orthogonal representation of the random process $u(t)$ defined by the series expansion (6.45). It is straightforward to see that such representation yields a corresponding series expansion of the fluctuation term $f(t)$ in (6.43). In fact we have the following

Proposition 4. *For any bi-orthogonal series expansion* (6.45) *of the solution to the MZ-equation* (6.43)*, there exists a unique series expansion of the fluctuation term $f(t)$ of the form*

$$f(t) = \overline{f} + \sum_{k=1}^{K} \sqrt{\lambda_k}\xi_k h_k(t). \tag{6.48}$$

Proof. It is sufficient to prove the theorem for zero-mean processes. To this end, we set $\overline{u} = 0$ and $\overline{f} = 0$ in (6.45) and (6.48). A substitution of (6.45) into (6.43) yields, for all $t \in [0, T]$

$$f(t) = \sum_{k=1}^{K} \sqrt{\lambda_k}\xi_k \left(\frac{de_k(t)}{dt} - \Omega e_k(t) - \int_0^t K(t-s)e_k(s)ds \right). \tag{6.49}$$

Define,

$$h_k(t) = \frac{de_k(t)}{dt} - \Omega e_k(t) - \int_0^t K(t-s)e_k(s)ds. \tag{6.50}$$

This equation does not allow us to compute h_k explicitly quite yet. In fact, the MZ memory kernel $K(t-s)$ depends on $f(t)$ (see Eq. (6.44)). However, a substitution of (6.48) (with $\overline{f} = 0$) into the analytical expression of $K(t)$ yields

$$K(t) = \sum_{i,j=1}^{K} \sqrt{\lambda_i \lambda_j} v_{ij} e_i(0) h_j(t), \qquad \text{where} \qquad v_{ij} = \frac{\langle \xi_i, \mathcal{L}\xi_j \rangle_{eq}}{\langle u(0), u(0) \rangle_{eq}}. \tag{6.51}$$

To evaluate $\langle \xi_i, \mathcal{L}\xi_j \rangle_{eq}$ we need to express $\{\xi_1, \ldots, \xi_K\}$ as a function of $\boldsymbol{x}_0$ (recall that $\mathcal{L}$ operates on functions of $\boldsymbol{x}_0$, see Eq. (2.3)), and then integrate over $\rho_{eq}(\boldsymbol{x}_0)$. This is easily achieved by using the dispersion relation (6.47). Specifically, we have

$$\xi_k(\boldsymbol{x}_0) = \frac{1}{\sqrt{\lambda_k}} \int_0^T u(\boldsymbol{x}(t, \boldsymbol{x}_0)) e_k(t) dt. \tag{6.52}$$

At this point, we substitute (6.51) into (6.50) to obtain

$$h_k(t) = \frac{de_k(t)}{dt} - \Omega e_k(t) - \sum_{i,j=1}^{K} \sqrt{\lambda_i \lambda_j} v_{ij} e_i(0) \int_0^t h_j(t-s) e_k(s) ds. \tag{6.53}$$

Given $\{e_1(t), \ldots, e_K(t)\}$, Ω and v_{ij}, this equation can be solved uniquely for $\{h_1(t), \ldots, h_K(t)\}$ by using Laplace transforms. Note that $\{h_1(t), \ldots, h_K(t)\}$ are not necessarily orthogonal in $L^2([0,T])$. □

Remark Proposition 4 establishes a one-to-one correspondence between the noise process in the MZ equation (6.43) and the biorthogonal series expansion of the solution. If the dynamical system (2.1) is Hamiltonian then the MZ steaming term vanishes, and the MZ memory kernel can be written in terms the fluctuation term as (see Eq. (2.20))

$$K(t) = \frac{\langle f(0), f(t) \rangle_{eq}}{\langle u(0), u(0) \rangle_{eq}}. \tag{6.54}$$

A substitution of this expression into (6.43) yields, after projection onto ξ_k

$$\frac{de_k(t)}{dt} = \int_0^t \sum_{j=1}^{K} \lambda_j \left[h_j(0) h_{k'}(t-s) \right] e_k(s) ds + h_k(t). \tag{6.55}$$

This equation establishes a one-to-one correspondence between the temporal modes of the KL expansion (6.45) and the temporal modes of the fluctuation term (6.49). In particular, given $\{e_1(t), \ldots, e_K(t)\}$, we can determine $\{h_1(t), \ldots, h_K(t)\}$ directly by using Laplace transforms, without building the MZ memory kernel (6.51).

The bi-orthogonal expansion of $u(t)$ already gives an effective model for $u(t)$ which is computed from its time autocorrection function. The ratione behind the introduction of the MZ-KL model (6.43) relies on an empirical fact that the fluctuation force $f(t)$ has faster time scales than $u(t)$, therefore it is possible to use the short time autocorrelation function of $u(t)$ to get effective models (6.49), (6.55) for $f(t)$ and $K(t)$, and then get the longe time dynamics of $u(t)$ through GLE equation (6.43) [57]. The methodology we used is an application of the reverse thinking of traditional approaches such as the ones used in [19, 57]. Instead of finding suitable fluctuation force $f(t)$ such that $u(t)$ has the correct statistics, we directly construct a model of $u(t)$ which has the desired statistics and recover $f(t)$ and $K(t)$ *a posteriori*. Hence within the reconstruction process, the second fluctuation dissipation theorem holds automatically. The resulting GLE (6.43) has the same explanatory power as (6.40).

Stochastic modeling from first principles By combining the MZ-KL model with the combinatorial algorithm we proposed in Section 6.1, we propose the following paradigm to build stochastic models for the observable $u(t)$ at statistical equilibrium from first principles. To this end,

1. Compute the solution to the MZ equation for the temporal correlation function of $u(t)$ (see Eq. (3.62))

$$\frac{dC(t)}{dt} = \Omega C(t) + \int_0^t K(t-s)C(s)ds. \tag{6.56}$$

   The memory kernel $K(t-s)$ can be expanded as in (6.1), and computed from first-principles using the combinatorial approach we discussed in Section 6.1.1. Note that $C(t)$ and $K(t)$ in Eq (6.56) normally are computed within some short-time scale.

2. Build the Karhunen-Loève expansion (6.45) by spectrally decomposing the correlation function $C(t) = \langle u(0)u(t)\rangle_{eq}$ obtained at point 1. Recall that at statistical equilibrium we have $C(t-s) = \langle u(0)u(t-s)\rangle_{eq} = \langle u(s)u(t)\rangle_{eq}$. This yields eigenvalues $\{\lambda_j\}$ and the eigenfunctions $e_j(t)$. The uncorrelated

random variables ξ_k appearing in (6.45) can be sampled consistently with the equilibrium distribution ρ_{eq} by using, e.g., the iterative algorithm recently proposed by Phoon *et al.* [79, 78].

3. With $\{\xi_1, \ldots, \xi_K\}$, $\{e_1(t), \ldots, e_K(t)\}$ and $\{\lambda_1, \ldots, \lambda_K\}$ available, we can uniquely identify the noise process $f(t)$ in the MZ equation (6.43). To this end, we simply use Proposition 4, with the temporal modes $h_k(t)$ obtained by solving equation (6.53) or (6.55) with the Laplace transform.

4. With $K(t)$ and $f(t)$ available, we may plug them in Eq (6.43) to get (long time) dynamics of $u(t)$.

The above paradiam can also be used in a *pure* data-driven setting, in the sense that $C(t)$ appeared in the first step can be replaced by simulated result obtained from direct numerical simulation (for stochastic system Monte-Carlo simulation may be used). Moreover, even when the analytic form of the equilibrium measure is unknown, the reconstruction algorithm proposed by Phoon *et al.* [79, 78] still works for a data-driven, empirical probability density ρ_u which is consistent with ρ_{eq}. This feature makes it a good modeling method for stochastic system in unknown nonequilibrium steady state (NESS) such as the one appeared in heat conduction model [26] and turbulence modeling [32]. After the stochastic model for $u(t)$ is built, one can use it to get *any* statistical quantities related to $u(t)$. In Section 6.3, we calculate time autocorrelation function of higher order moments of $u(t)$ to verify the accuracy of the stochastic model. Other statistical quantities such as the one body intermediate scattering function (ISF) introduced in [1] can be computed in a similar way. The results of this section can be generalized to vector-valued phase space functions $\boldsymbol{u}(t)$ at statistical equilibrium. The starting point is the KL expansion for multi-correlated stochastic processes we recently proposed in [14]. Such expansion is constructed based on cross-correlation information[4], and can be made consistent with the equilibrium distribution of $\boldsymbol{u}(t)$, e.g., by using the sampling strategy proposed in [79, 78]. The correspondence between the KL expansions of $\boldsymbol{u}(t)$ and the vector-valued fluctuation term $\boldsymbol{f}(t)$ can be established by following the same arguments we used in the proof of Proposition 4.

6.3 Applications to nonlinear systems with local interactions

In this Section, we demonstrate the accuracy of the MZ memory calculation method and the reduced-order stochastic modeling technique we discussed in Section 6.1 and Section 6.2, respectively. To this end, we study nonlinear random wave propagation described by Hamiltonian partial differential equations (PDEs). To derive such PDEs consider the nonlinear functional

$$\mathcal{H}([p],[u]) = \int_0^{2\pi} \left[\frac{p^2}{2} + \frac{\alpha}{2}u_x^2 + G(p, u_x, u)\right] dx, \tag{6.57}$$

where $u = u(x,t)$ represents the wave displacement, $p = p(x,t)$ is the canonical momentum (field variable conjugate to $u(x,t)$), $u_x = \partial u/\partial x$, and $G(p, u_x, u)$ is the nonlinear interaction term. By taking functional derivatives of (6.57) with

[4]At statistical equilibrium the cross correlation functions are invariant under temporal shifts. This means that $\langle u_i(s), u_j(t)\rangle_{eq} = \langle u_i(0), u_j(t-s)\rangle_{eq}$ for all $t \geq s$. Hence, the solution to the projected MZ equation (3.62) is sufficient to compute the KL expansion of the multi-correlated process $\boldsymbol{u}(t)$, e.g., using the series expansion method proposed in [14].

respect to p and u (see, e.g., [94]) we obtain the Hamilton's equations of motion

$$\begin{cases} \partial_t u &= \dfrac{\delta \mathcal{H}(p,u)}{\delta p(x,t)} = p + \partial_p G(p, u_x, u), \\ \partial_t p &= -\dfrac{\delta \mathcal{H}(p,u)}{\delta u(x,t)} = \alpha u_{xx} + \partial_x \partial_{u_x} G(p, u_x, u) - \partial_u G(p, u_x, u). \end{cases} \tag{6.58}$$

The corresponding nonlinear wave equation is

$$u_{tt} = \alpha u_{xx} + \partial_t \partial_p G(p, u_x, u) + \partial_x \partial_{u_x} G(p, u_x, u) - \partial_u G(p, u_x, u). \tag{6.59}$$

This equation has been studied extensively in mathematical physics [51, 23, 49, 75], in particular in general relativity, statistical mechanics, and in the theory of viscoelastic fluids. In Figure 6.1 and Figure 6.2, we plot a few sample numerical solutions to (6.59) corresponding to different initial conditions and different nonlinear interaction term $G(p, u_x, u)$. These solutions are computed by an accurate Fourier spectral method with $N = 512$ modes (periodic boundary conditions in $x \in [0, 2\pi]$). Throughout this section, we assume that the initial state $\{u(x,0), p(x,0)\}$ is random and distributed according to the functional Boltzmann-Gibbs equilibrium distribution[5]

$$\rho_{eq}([p],[u]) = \frac{1}{Z(\alpha,\gamma)} e^{-\gamma \mathcal{H}([p],[u])}, \quad \text{where} \qquad Z(\alpha,\gamma) = \int e^{-\gamma \mathcal{H}(p,u)} \mathcal{D}[p(x)] \mathcal{D}[u(x)]. \tag{6.60}$$

We emphasize that $\rho_{eq}([p],[u])$ is invariant under the infinite-dimensional flow generated by (6.59) with periodic boundary conditions, since the Hamiltonian (6.57) is a constant of motion (conserved quantity) in this case.

[5]The partition function $Z(\alpha,\gamma)$ is defined as a functional integral over $u(x)$ and $p(x)$ (see, e.g., [94]).

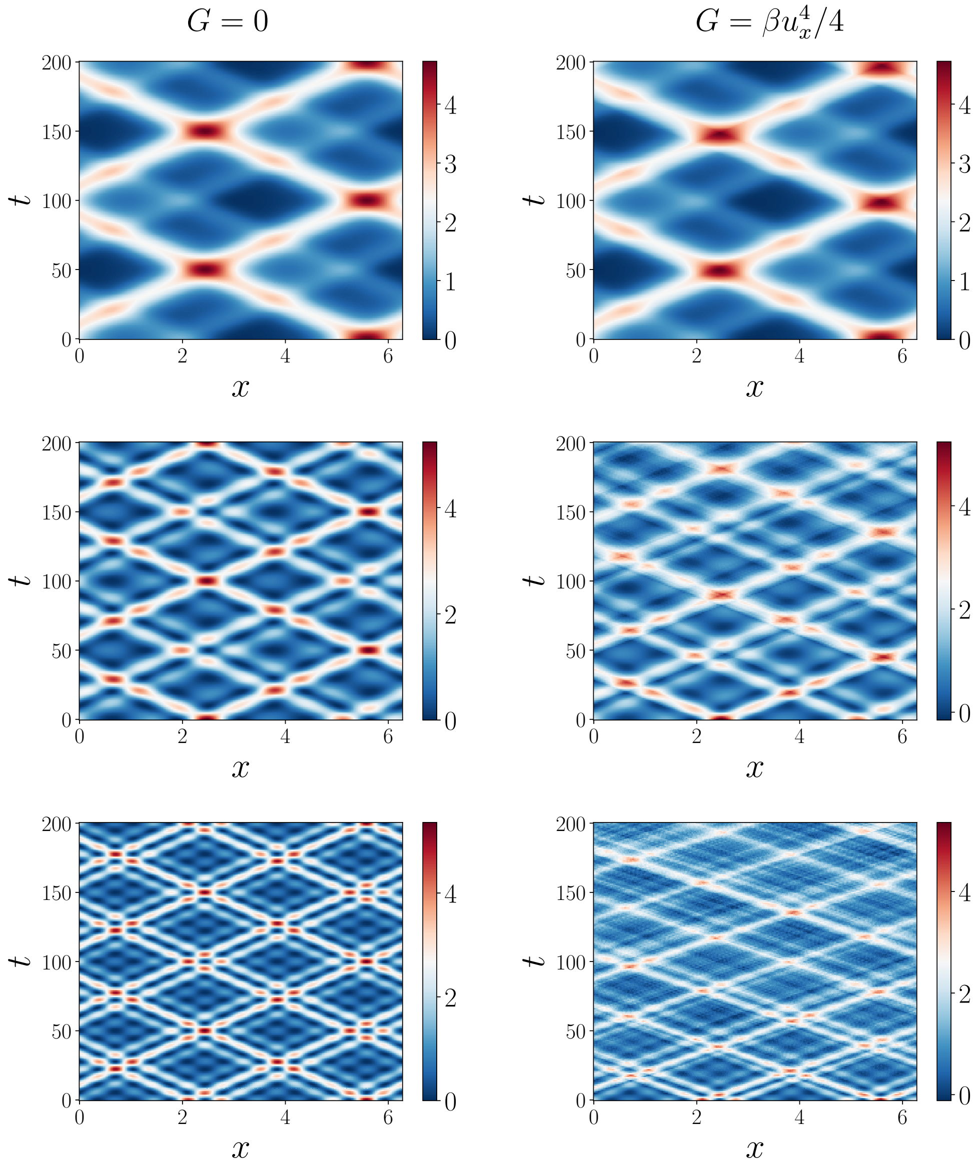

Figure 6.1: Sample solutions of the nonlinear wave equation (6.59) with initial conditions $u(x, 0) = e^{-\sin(2x)}(1 + \cos(x))$ (first row), $u(x, 0) = e^{-\sin(2x)}(1 + \cos(5x))$(second row), and $u(x, 0) = e^{-\sin(2x)}(1 + \cos(9x))$ (third row). We set the group velocity α to $(2\pi/100)^2$ and consider different nonlinear interaction terms: $G = 0$ (first column – linear waves), $G = \beta u_x^4/4$ with $\beta = (2\pi/100)^4$ (second column – nonlinear waves). It is seen that as the initial condition becomes rougher, the nonlinear effects become more important.

6.3.1 Case I: $G(p, \nabla u, u) = 0$

Setting the interaction term $G(p, u_x, u)$ in (6.57) and (6.59) equal to zero yields the well-known linear wave equation

$$u_{tt} = \alpha u_{xx}. \tag{6.61}$$

We discretize (6.61) in space using second-order finite differences on the (periodic) grid $x_j = 2\pi j/N$ ($j = 0, \ldots, N$). This yields the following linear dynamical system

$$\frac{du_j}{dt} = p_j, \qquad \frac{dp_j}{dt} = \frac{\alpha}{h^2}(u_{j+1} - 2u_j + u_{j-1}), \tag{6.62}$$

where $u_j(t) = u(x_j, t)$, $p_j(t) = p(x_j, t)$, and $h = 2\pi/N$ is the mesh size. The Hamilton's function corresponding to the finite-difference scheme (6.62) is obtained by discretizing the integral (6.57), e.g., with the rectangle rule. This yields

$$\mathcal{H}_1(\boldsymbol{p}, \boldsymbol{u}) = \sum_{j=0}^{N-1} \frac{h}{2} p_j^2 + \frac{\alpha_1 h}{2} \sum_{j=0}^{N-1} (u_{j+1} - u_j)^2, \tag{6.63}$$

where we defined $\alpha_1 = \alpha/h^2$. The corresponding finite-dimensional Gibbs distribution can be written as

$$\rho_{eq}(\boldsymbol{p}, \boldsymbol{u}) \propto \exp\left\{-\gamma \left(\sum_{j=0}^{N-1} \frac{1}{2} p_j^2 + \frac{\alpha_1}{2} \sum_{j=0}^{N-1} (u_{j+1} - u_j)^2\right)\right\}, \tag{6.64}$$

$Z_1(\alpha_1, \gamma)$ being the partition function (normalization constant). Note that we absorbed the scaling factor h in the parameter $\gamma > 0$. It is straightforward to verify that the lattice Hamiltonian (6.63) is preserved if $u_0 = u_N$ and $p_0 = p_N$ (periodic boundary conditions). This implies that the PDF (6.64) is invariant under the flow generated by the linear ODE (6.62). Note that the lattice Hamiltonian (6.63)

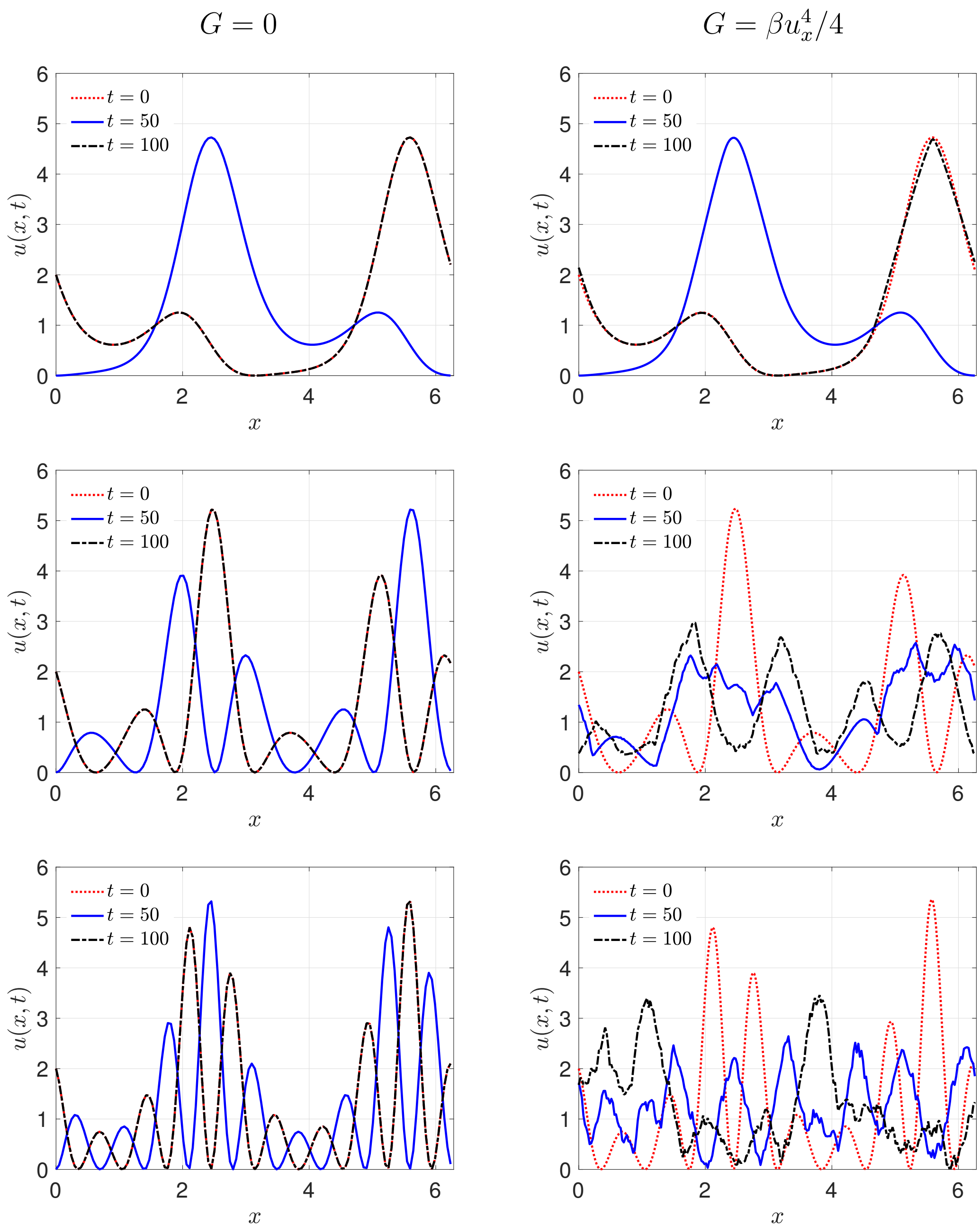

Figure 6.2: Snapshots of the solution shown in Figure 6.1.

coincides with the Hamiltonian of a one-dimensional chain of harmonic oscillators with uniform mass $m = 1$ and spring constants $k = \alpha_1$. We set $N = 100$ and $\alpha = (2\pi/100)^2$ in equation (6.62). In this way, the system (6.62) is 200-dimensional and the modeling parameter α_1 in (6.63)-(6.64) is equal to 1.

MZ memory kernel and auto-correlation functions The Hamiltonian system (6.62) with periodic boundary conditions has many symmetries. In particular, the statistical properties of wave displacement $u(x,t)$ at any point x_j are the same, if the initial state is distributed according to (6.64). In addition, the PDF of the wave momentum[6] $p(x_j,t)$ and the wave displacement $r(x_j,t) = u(x_{j+1},t) - u(x_j,t)$ are both Gaussian (see Eq. (6.64)). Suppose we are interested in the temporal auto-correlation function of the wave momentum $p(x_j,t) = p_j$, at an arbitrary location x_j, i.e.,

$$C_{p_j}(t) = \langle p_j(t), p_j(0) \rangle_{eq}, \tag{6.65}$$

where $\langle , \rangle_{eq}$ is an integral over the equilibrium distribution (6.64). Such correlation function admits the analytical expression (see [33])

$$C_{p_j}(t) = J_0(2t), \qquad \forall \gamma > 0, \tag{6.66}$$

where J_0 is the zero-order Bessel function of the first kind. With $C_{p_j}(t)$ available, we can solve the MZ equation

$$\frac{d}{dt} C_{p_j}(t) = \int_0^t K(t-s) C_{p_j}(s) ds \tag{6.67}$$

for the memory kernel $K(t)$ by using Laplace transforms. This yields the exact MZ kernel

$$K(t) = \frac{J_1(2t)}{t}, \qquad \forall \gamma > 0, \tag{6.68}$$

where J_1 is the first-order Bessel function of the first kind. In Figure 6.3, we compare the exact memory kernel (6.68) and the correlation function (6.67) with the results we obtained using the iterative algorithm discussed in Section 6.1.1.

[6] Note that for linear waves the wave momentum $p(x,t)$ is equal to $\partial u(x,t)/\partial t$ (see Eq. (6.58)).

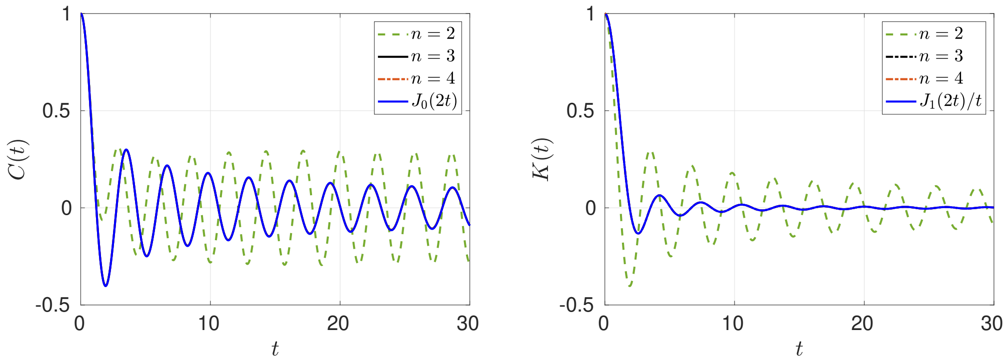

Figure 6.3: Linear wave equation (6.61). Temporal auto-correlation function (6.65) of the momentum $p(x_j,t) = \partial u(x_j,t)/\partial t$ (any location x_j) and MZ memory kernel $K(t)$. Shown are the analytical results (6.66)-(6.68), and the results we obtained by using the recursive algorithm we discussed in Section 6.1.

Note that the system (6.62) is linear. Therefore, we can use the formula (6.21) to compute the coefficients $\{\gamma_1,\dots,\gamma_{n+2}\}$ $(n=6)$. With such coefficients available, we then compute $\{\mu_1,\dots,\mu_{n+2}\}$ using the recurrence relation (6.6), and the MZ memory kernel (6.1). The linear wave is known to be an integrable system. For which, the convergence of the MZ series expansion has been proved theoretically in Chapter 5 since the dynamics of the $e^{t\mathcal{L}}$ and $e^{t\mathcal{QL}}$ is closed within a finite dimensional Hilbert space. In Figure 6.3, we see clearly the convergence of MZ-Faber expansion at the very order n.

Reduced-order stochastic modeling Suppose we are interested in building a consistent reduced-order stochastic model for the wave momentum $p(x_j,t) = \partial u(x_j,t)/\partial t$ at statistical equilibrium. To this end, we employ the spectral expansion technique we discussed in Section 6.2. The auto-correlation function of the process $p(t) = p(x_j,t)$ (at any location x_j), i.e., (6.65), is obtained by solving the MZ equation (6.67) with the kernel computed using the combinatorial algorithm described in Section 6.1.1. Following the first principle modeling paradigm we

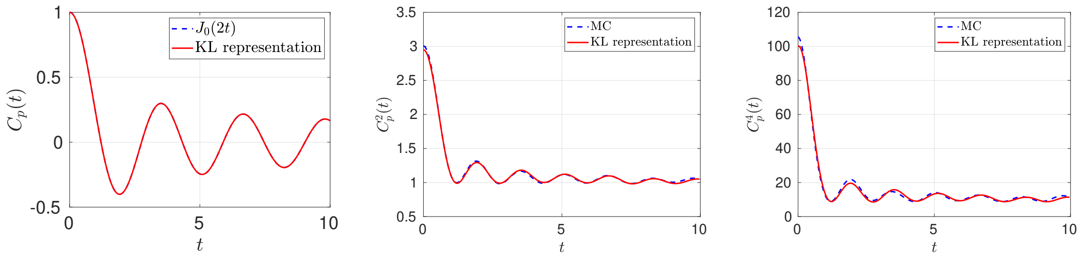

Figure 6.4: Linear wave equation (6.61). Temporal auto-correlation functions (6.70) of the wave momentum.

proposed in Section 6.2, we expand $p(t)$ as

$$p(t) \simeq \sum_{k=1}^{K} \sqrt{\lambda_k} \xi_k(\omega) e_k(t), \tag{6.69}$$

where $(\lambda_k, e_k(t))$ are eigenvalues and eigenfunctions of (6.65). By enforcing consistency of (6.69) with the equilibrium distribution (6.64) at each fixed time we obtain that the random variables $p(t_j)$ are normally distributed with zero mean and variance $1/\gamma$, for all $t_j \in [0, 10]$. In other words $p(t)$ is a centered, stationary Gaussian random process with correlation function (6.65). In Figure 6.4, we plot the auto-correlation functions

$$C_p(t) = \langle p_j(t), p_j(0) \rangle_{eq}, \qquad C_p^2(t) = \langle p_j^2(t), p_j^2(0) \rangle_{eq}, \qquad C_p^4(t) = \langle p_j^4(t), p_j^4(0) \rangle_{eq}. \tag{6.70}$$

Convergence of KL expansions representing high-order correlation functions such as (6.70) is established in Appendix. In Figure 6.4, we see a convergence of the KL representation to the exact result.

6.3.2 Case II: $G(p, \nabla u, u) = \frac{\beta}{4} u_x^4$

In this section, we study the nonlinear wave equation (6.59) with interaction term $G(p, u_x, u) = \beta u_x^4/4$, i.e.,

$$u_{tt} = \alpha u_{xx} + 3\beta u_x^2 u_{xx}, \qquad \alpha, \beta > 0. \tag{6.71}$$

In Figure 6.1 and Figure 6.2 we plot sample solutions of (6.71) corresponding to different initial conditions. It is clearly seen that the nonlinearity $u_x^2 u_{xx}$ breaks the periodicity of traveling wave. This effect is more pronounced if the initial condition is rougher in x, as u_x^2 and u_{xx} are larger in this case, thereby increasing magnitude of the nonlinear term in (6.71). As before, we discretize (6.71) and the Hamiltonian (6.57) with finite differences on a periodic spatial grid (N points in $[0, 2\pi]$). This yields

$$\mathcal{H}_2(\boldsymbol{p}, \boldsymbol{u}) = \sum_{j=0}^{N-1} \frac{h p_j^2}{2} + \sum_{j=0}^{N-1} \frac{h\alpha_1}{2} (u_{j+1} - u_j)^2 + \sum_{j=0}^{N-1} \frac{h\beta_1}{4} (u_{j+1} - u_j)^4, \tag{6.72}$$

where $u_j(t) = u(x_j, t)$ and $p_j(t) = \partial u(x_j, t)/\partial t$ represent the wave amplitude and momentum at location $x_j = hj$ ($j = 0, \ldots, N$, $h = 2\pi/N$), $\alpha_1 = \alpha/h^2$ and $\beta_1 = \beta/h^4$. The discretized equilibrium distribution (6.60) then becomes

$$\rho_{eq}(\boldsymbol{p}, \boldsymbol{u}) \propto \exp\left\{ -\gamma \left(\sum_{j=0}^{N-1} \frac{p_j^2}{2} + \sum_{j=0}^{N-1} \frac{\alpha_1}{2} (u_{j+1} - u_j)^2 + \sum_{j=0}^{N-1} \frac{\beta_1}{4} (u_{j+1} - u_j)^4 \right) \right\}. \tag{6.73}$$

As before, we absorbed the factor h into the parameter γ. Note that the lattice Hamiltonian (6.72) coincides with the Hamiltonian of the Fermi-Pasta-Ulam β-model (6.33), with $m_j = 1$. We emphasize that if a different scheme is used to discretize the wave equation (6.71), then the lattice Hamiltonian (6.72) may not be a conserved quantity.

MZ memory term and auto-correlation functions We choose the wave momentum $p_j(t)$ and the wave displacement $r_j(t) = u_{j+1}(t) - u_j(t)$ as quantities of interest. Moreover, we set $N = 100$ and $\alpha = (2\pi/100)^2$. To study the effects of the nonlinear interaction term, we consider different values of $\beta = \beta_1\alpha^2$, with β_1 ranging from 0.01 to 1. This corresponds to the FPU models with mild and strong nonlinearities, respectively. Based on the structure of the Hamiltonian (6.72) and the equilibrium distribution (6.73), we expect that the dynamics of $p_j(t)$ and $r_j(t)$ will be different for different parameters β. To calculate the temporal auto-correlation function of $p_j(t)$ and $r_j(t)$ at an arbitrary spatial point x_j, we solve the corresponding MZ equations. Such equations are of the form (6.67), where the memory kernel $K(t-s)$ is computed from first-principles (i.e., from the microscopic equations of motion) using the algorithm we presented in Section 6.1.1. In Figure 6.5, we compare the temporal auto-correlation function we obtained for the wave displacement $r_j(t)$ with results of Markov-Chain-Monte-Carlo (MCMC) (10^6 sample paths) for FPU systems with mild nonlinearities ($\beta_1 = 0.01$ and $\beta_1 = 0.1$) at different temperatures ($\gamma = 1$ and $\gamma = 40$). It is seen that the MZ-Faber approximation of the MZ memory kernel yields relatively accurate results for FPU systems with mild nonlinearties at both low ($\gamma = 40$) and high temperature ($\gamma = 1$) as we increase the polynomial order n.

Reduced-order stochastic modeling We employ the spectral approach of Section 6.2 to build stochastic low-dimensional models of the wave momentum $p_j(t)$ and wave displacement $r_j(t) = u_{j+1}(t) - u_j(t)$ at statistical equilibrium. Since we assumed that we are at statistical equilibrium, the statistical properties of the random processes representing $p_j(t)$ and $r_j(t)$ are time-independent. For

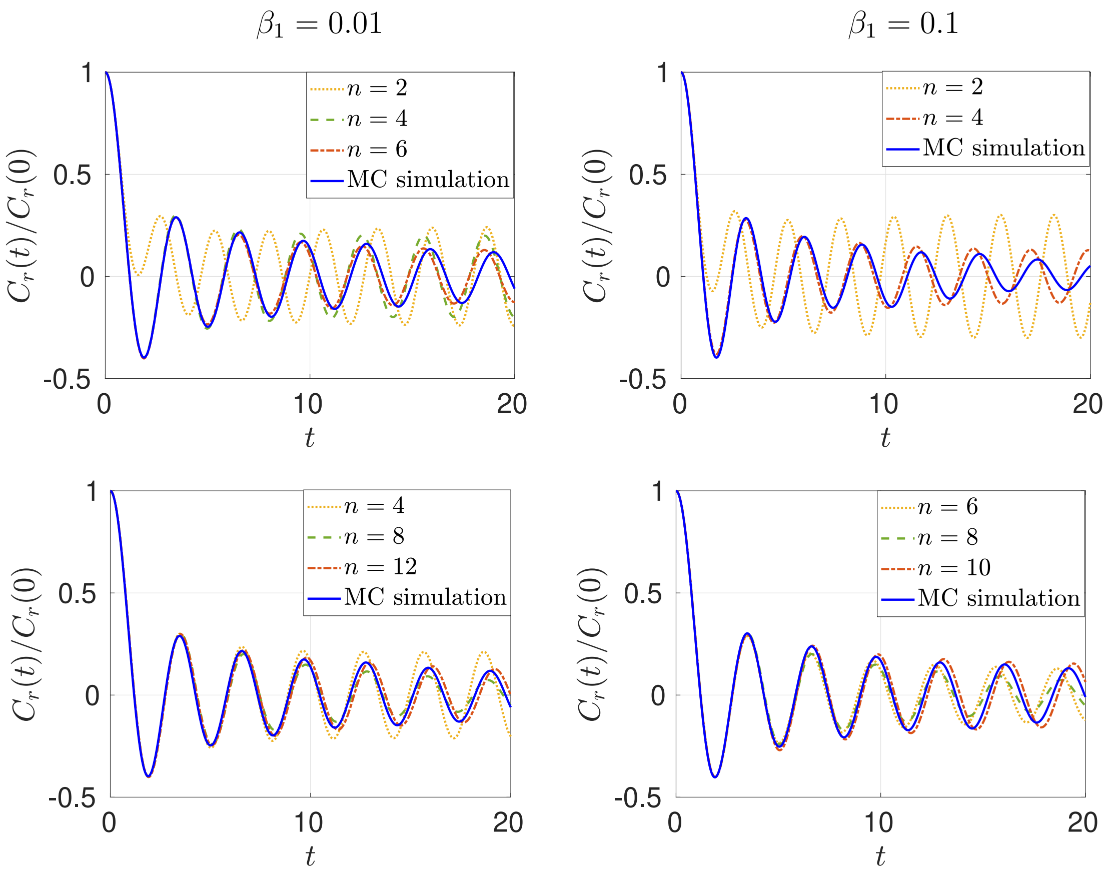

Figure 6.5: Nonlinear wave equation (6.71). Temporal auto-correlation function of the wave displacement $r_j(t)$ for different values of the nonlinear parameter β_1. We compare results we obtained by calculating the MZ memory from first principles using n-th order Faber polynomials (Section 6.1.1) with results from Markov-Chain-Monte-Carlo (10^6 sample paths). The thermodynamic parameter γ is set to 1 (high-temperature) in the first row and to 40 (low-temperature) in the second row.

instance, by integrating (6.73) we obtain the following expression for the one-time PDF of $r_j(t)$

$$r_j(t) \sim e^{-\gamma(\frac{1}{2}\alpha_1 r^2 + \frac{1}{4}\beta_1 r^4)} \qquad \forall t \in [0,T], \quad \forall j = 0,\dots,N-1. \tag{6.74}$$

Clearly, $r_j(t)$ is a stationary non-Gaussian process. To sample the KL expansion of $r_j(t)$ in a way that is consistent with the PDF (6.74) we used the algorithm discussed in [78, 79]. It is straightforward to show that for all $m \in \mathbb{N}$

$$\mathbb{E}\{r_j^{2m}(t)\} = \frac{\displaystyle\int_{-\infty}^{+\infty} r^{2m} e^{-\gamma(\frac{1}{2}r^2 - \frac{1}{4}r^4)} dr}{\displaystyle\int_{-\infty}^{+\infty} e^{-\gamma(\frac{1}{2}r^2 - \frac{1}{4}r^4)} dr} = \frac{\sqrt{2}\gamma^{-\frac{1}{4}-\frac{m}{2}} \Gamma\left(\frac{1}{2}+m\right) U\left(\frac{1}{4}+\frac{m}{2}, \frac{1}{2}, \frac{\gamma}{4}\right)}{e^{\gamma/8} K_{1/4}\left(\frac{\gamma}{8}\right)},$$

where $\Gamma(x)$ is the Gamma function, $K_n(z)$ is the modified Bessel function of the second kind and $U(x,y,z)$ is Tricomi's confluent hypergeometric function. Therefore, for all positive γ and finite m we have that $\mathbb{E}\{r_j^{2m}(t)\} < \infty$, i.e., $r_j(t)$ is L^{2m} process. This condition guarantees convergence of the KL expansion to temporal correlation functions of order greater than two (see Appendix A). In Figure 6.6 we plot the temporal auto-correlation function of various polynomial observables of the nonlinear wave momentum and displacement at an arbitrary spatial point x_j. We compare results we obtained from Markov Chain Monte Carlo simulation (dashed line), with the MZ-KL expansion method based the first-principle memory calculation (continuous line). We also provide results we obtained by using KL expansions with covariance kernel estimated from data (dotted line).

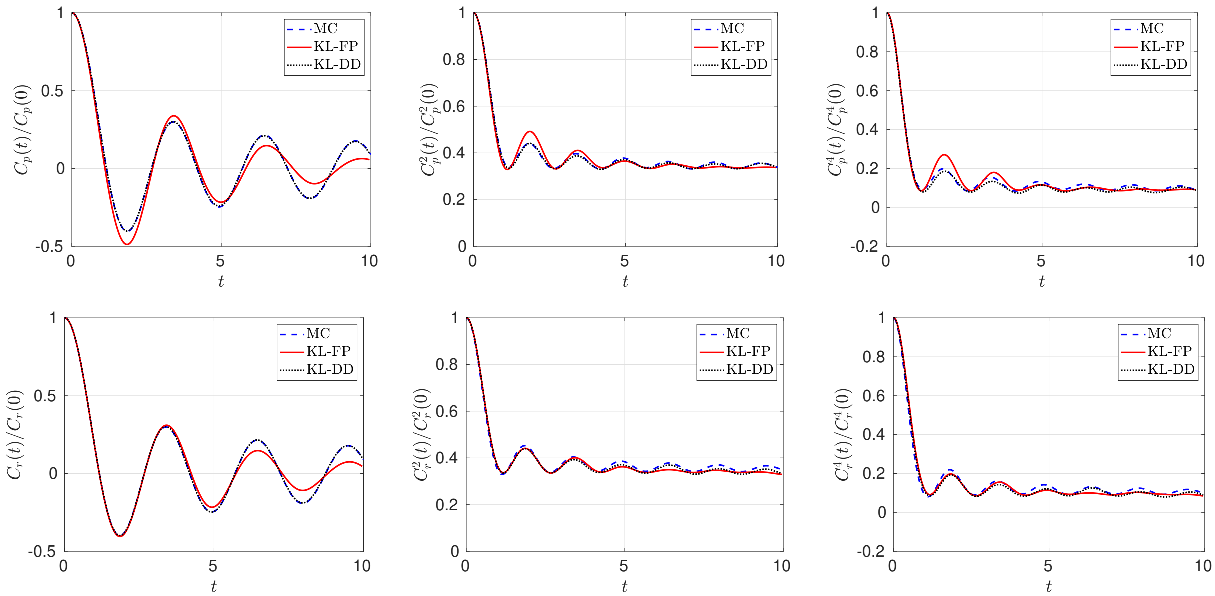

Figure 6.6: Nonlinear wave equaton (6.71). Temporal auto-correlation function of polynomial observables $p_j^m(t)$ (first row) $r_j^m(t)$ (second row) with $m = 1, 2, 4$. We compare results from Markov-Chain-Monte-Carlo simulation (MC), KL expansion based on the first-principle MZ memory kernel calculation (6.67) (KL-FP), and KL expansion based on a data-driven estimate of the temporal auto-correlation function (KL-DD). The parameter γ appearing in (6.73) is set to 40, while $\alpha_1 = \beta_1 = 1$.

6.3.3 Case III: $G(p, \nabla u, u) = \frac{\beta}{4}u^4$

In this section, we consider the nonlinear wave equation with the interaction term $G(p, \nabla u, u) = \frac{\beta}{4}|u|^4$. The Hamiltonian equation of motion for this system is given by

$$\begin{cases} \partial_t u & = p \\ \partial_t p & = \alpha \Delta u - \beta |u|^2 u \end{cases} \tag{6.75}$$

The corresponding nonlinear wave equation is

$$u_{tt} = \alpha \Delta u - \beta |u|^2 u. \tag{6.76}$$

Again, we assume that the nonlinear wave equation evolves from Gibbs initial state $e^{-\gamma\mathcal{H}(p,u)}/Z(\alpha,\beta,\gamma)$ with periodic BCs. The same finite difference scheme can be introduced to approximate the wave equation and the corresponding Gibbs measure. After some simple calculation, we shall get a lattice Hamiltonian system with Hamiltonian

$$\mathcal{H}_3(\boldsymbol{p},\boldsymbol{u}) = \sum_{j=0}^{N-1} \frac{hp_j^2}{2} + \sum_{j=1}^{N-1} \frac{h\alpha_1}{2}(u_{j+1}-u_j)^2 + \sum_{j=0}^{N-1} \frac{h\beta}{4}u_j^4 \tag{6.77}$$

where $\alpha_1 = \alpha/h^2$. The lattice Hamiltonian (6.77) gives the Landau-Ginsburg model [75] with vanishing u^2 term. The equilibrium Gibbs distribution ρ_{eq} becomes

$$\rho_{eq} \propto \exp\left\{-\gamma\left(\sum_{j=0}^{N-1} \frac{p_j^2}{2} + \sum_{j=1}^{N-1} \frac{\alpha_1}{2}(u_{j+1}-u_j)^2 + \sum_{j=0}^{N-1} \frac{\beta}{4}u_j^4\right)\right\} \prod_{j=0}^{N-1} du_j dp_j, \tag{6.78}$$

Different from Case II where a non-canonical transformation can be introduced to decouple the Gibbs distribution (6.73), there is no such non-canonical transformation for Gibbs measure (6.78). In fact, through the nonlinear interaction $\frac{\beta}{4}u_j^4$, u_j becomes dependent on all the other $u_i, i \neq j$. This brings us numerical difficulties when trying to use the first-principle method to get the expansion coefficients of the GLE. To see this, we noticed that an accurate calculation of the expansion coefficients μ_n requires the evaluation of n-point correlation

$$M^n = \langle u_{j_1}^{m_{j_1}}, \cdots, u_{j_n}^{m_{j_n}} \rangle_{eq}$$

where $j_1, \cdots, j_n$ are relevant degree of freedom. For Gibbs measure (6.78), $u_0, \cdots, u_{N-1}$ are dependent random variables with respect to joint distribution

$$\rho_{\boldsymbol{u}} \propto \exp\left\{\sum_{j=1}^{N-1} \frac{\alpha}{2}(u_{j+1} - u_j)^2 + \sum_{j=0}^{N-1} \frac{\beta}{4} u_j^4\right\}$$

To accurately calculate the n-point correlation

$$M^n = \langle u_{j_1}^{m_{j_1}}, \cdots, u_{j_n}^{m_{j_n}} \rangle_{eq} = \langle u_{j_1}^{m_{j_1}}, \cdots, u_{j_n}^{m_{j_n}} \rangle_{\rho_{\boldsymbol{u}}}$$

with respect to $\rho_{\boldsymbol{u}}$ is a hard task since there is no analytic expression for M^n and any sampling method becomes inaccurate when n and N are large. Having said that, it is still possible to use the first principle method to calculate γ_n for some special cases. For instance, if $\beta \ll 1$ is relatively smaller when comparing to α, a technique from perturbative quantum field theory can be used to evaluate $n-$th order moments M_n through Wick's theorem and Feymann diagram. We shall not discuss the perturbation method in current paper.

Calculation of the correlation function For Hamiltonian system (6.77), we choose the tagged particle momenta p_j and its coordinate u_j as quantities of interest. We further specify $h = L/N = 2\pi/100$, $\alpha_1 = (2\pi/100)^2/h^2 = 1$ and $\beta = 1$. Considering the numerical difficulties in calculating GLE expansion coefficients μ_n from the first principle, we choose to calculate the correlation function $C(t)$ by using data-driven method. i.e. the Monte-Carlo simulation. To sample from the Gibbs distribution (6.78), we introduce the following overdamped Langevin dynamics [64, 12]:

$$dP_j = -\partial_{P_j}\mathcal{H}_3 d\tau + \sqrt{\frac{2}{\beta}} dW_{P_j}(t)$$
$$dU_j = -\partial_{U_j}\mathcal{H}_3 d\tau + \sqrt{\frac{2}{\beta}} dW_{U_j}(t), \qquad j = 0, 1, \cdots, N-1, \tag{6.79}$$

where $W_{P_j}(t)$ and $W_{U_j}(t)$ are standard Wiener processes and periodic boundary condition $P_0 = P_N$ and $U_0 = U_N$ is imposed. Here we use capital letter P and U to represent stochastic process generated by the overdampled Langevin equation (6.79). To solve (6.79) numerically, we use the following Euler-Maruyama finite difference scheme:

$$P_j^{n+1} = P_j^n - \Delta t \partial_{P_j}\mathcal{H}_3 + \sqrt{\frac{2\Delta t}{\beta}} G_{P_j}^n$$
$$U_j^{n+1} = U_j^n - \Delta t \partial_{U_j}\mathcal{H}_3 + \sqrt{\frac{2\Delta t}{\beta}} G_{U_j}^n, \qquad j = 0, 1, \cdots, N-1, \tag{6.80}$$

where $G_{P_j}^n$ and $G_{U_j}^n$ are standard Gaussian random variables. $\Delta t = 0.005$ is the step size used in our simulation. We run the simulation for 10^6 steps and then truncated the first 10^4 samples. i.e. we use the solution of (6.79) from $t = 50$ to $t = 5000$. Samples from the equilibrium distribution are obtained by collecting $P_j(t)$ and $U_j(t)$ every 10 steps for $t \in [50, 5000]$.

Reduced-order modelling Now, we use the KL series to represent the dynamics of the coordinate u_j and momenta p_j for a tagged particle j. The procedure to build stochastic process representation of $p_j(t)$ is exactly the same as before since $P(t; \boldsymbol{\omega})$ is a Gaussian process. For $u_j(t)$, however, the situation is slightly

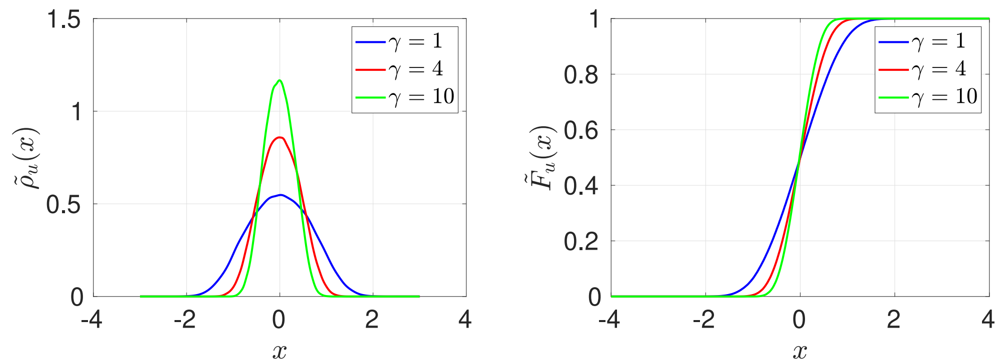

Figure 6.7: Empirical probability density function $\tilde{\rho}_u(x)$ and the cumulative distribution function $\tilde{F}_u(x)$ for $u_j(t)$.

different. The exact form of the marginal distribution for u_j is given by

$$\rho_{u_j}(x) = \int \rho_{\boldsymbol{u}} \prod_{i=0, i\neq j}^{N-1} du_j \propto \int \prod_{i=0, i\neq j}^{N-1} du_i \exp\left\{\sum_{i=1}^{N-1} \frac{\alpha}{2}(u_{i+1} - u_i)^2 + \sum_{i=0}^{N-1} \frac{\beta}{4} u_i^4\right\} \tag{6.81}$$

which cannot be explicitly calculated because of the dependence of u_j and all the rest degree of freedom $\{u_i\}_{i\neq j}$. Hence, in stead of calculating the multiple integral (6.81), we use the emperical probability density $\tilde{\rho}_u(x)$ to replace the exact PDF $\rho_u(x)$ in Phoon's algorithm. While $\tilde{\rho}_u(x)$ can be obtained by standard density estimator from equilibrium samples generated by the overdamped Langevin equation. In Figure 6.7, we display the estimated probability density function (PDF) $\tilde{\rho}_u(x)$ and the cumulative distribution function (CDF) $\tilde{F}_u(x)$. In Figure 6.8, we compare the time autocorrelation function (6.70) generated by the Monte-Carlo simulation and the truncated KL series. The KL approximation result for $C_u^{2m}(t)$ is obtained through formula

$$C_u^{2m}(t) \approx \frac{1}{N} \sum_{n=1}^{N} \left[\left(\sum_{k=1}^{K} \sqrt{\lambda_k^u} R_k(\omega^{(n)}) e_k^u(t)\right) \left(\sum_{k=1}^{K} \sqrt{\lambda_k^u} R_k(\omega^{(n)}) e_k^u(0)\right)\right]^{2m} \tag{6.82}$$

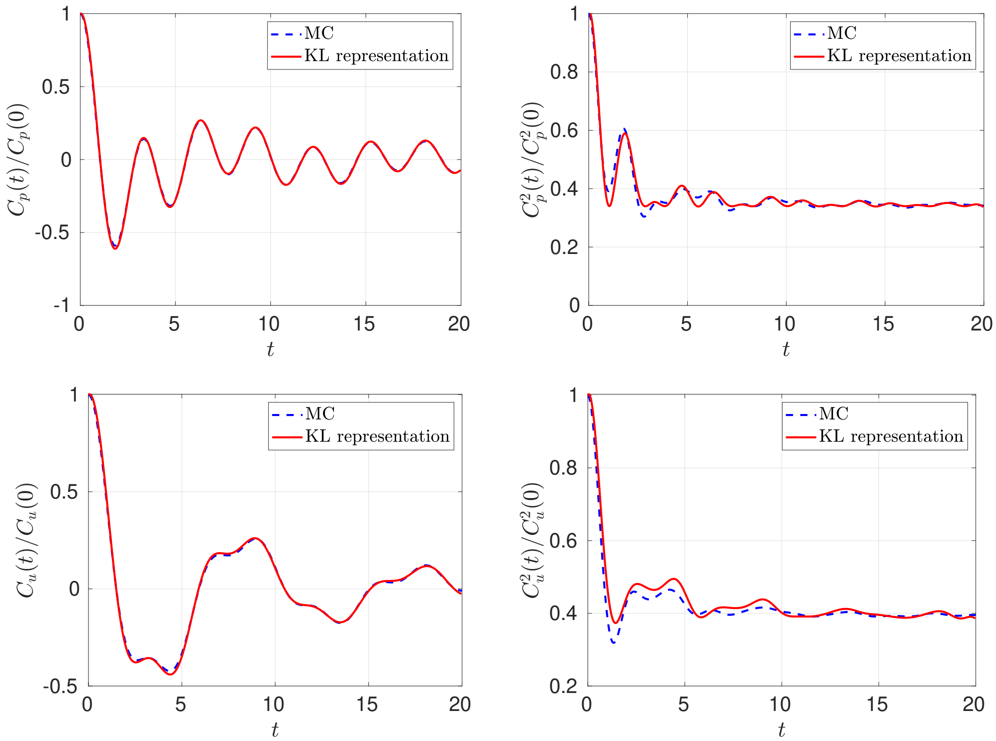

Figure 6.8: Time autocorrelation function of polynomial observable $p_j^{2m}(t)$ (first row) and $u_j^{2m}(t)$ (second row) for nonlinear wave equation (6.76). The thermodynamic parameter $\gamma = 4$.

6.4 Summary

In this chapter, we developed a new method to approximate the Mori-Zwanzig (MZ) memory integral in generalized Langevin equations (GLEs) describing the evolution of smooth observables in high-dimensional nonlinear systems with local interactions. The new method is based on Faber operator series expansions [104], and a formally exact combinatorial algorithm that allows us to compute the expansion coefficients of the MZ memory from first principles, i.e., based on the microscopic equations of motion. We also developed a new stochastic modeling technique that employs Karhunen-Loève expansions to represent the MZ fluctuation term (random noise) for systems in statistical equilibrium. We demonstrated the MZ memory calculation method and the MZ-KL stochastic modeling technique in applications to random wave propagation and prototype problems in classical statistical mechanics such as the Fermi-Pasta-Ulam β-model. We found that the proposed algorithms can accurately capture relaxation to statistical equilibrium in systems with mild nonlinearities, and in strongly nonlinear systems at high-temperature. At low temperature the Faber expansion of the MZ memory kernel is granted to converge only on a time interval that depends on the system and on the observable. In particular, Corollary 3.4.3 in [103] establishes short-time convergence of the MZ-Faber memory approximation for a broad class of nonlinear systems of the form (2.1). This implies that the MZ-Faber cumulant expansion can exhibit short-time convergence, meaning that it produces first-principle results that are accurate only for relatively short integration times.

We conclude by emphasizing that the mathematical techniques we presented can be readily applied to more general systems with local interactions such as particle systems modeling the microscopic dynamics of solids and liquids [102, 58, 59]. This opens the possibility to build new approximation schemes for MZ equations and derive new types of coarse-grained models where the MZ memory is constructed from first-principles and the fluctuation term is modeled stochastically.

Appendix 6.A Auto-correlation function of polynomial observables

In this Appendix we prove that the temporal auto-correlation function of phase space functions of the form $u^n(t) = u^n(\boldsymbol{x}(t, \boldsymbol{x}_0))$, i.e.,

$$\langle u^n(0), u^n(t) \rangle_\rho \qquad n \in \mathbb{N} \tag{6.83}$$

can be represented by replacing $u(t)$ with the KL expansion (6.45), and then sending K to infinity. This result allows us to compute the auto-correlation function of $u^n(t)$ based on the KL expansions of $u(t)$.

Theorem A.1 *Consider a zero-mean stationary stochastic process $u(t)$, $t \in [0,T]$, and assume that it has finite joint moments up to any desired order. Let*

$$u_K(t) = \sum_{k=1}^{K} \sqrt{\lambda_k}\xi_k e_k(t), \tag{6.84}$$

be the truncated Karhunen-Loève expansion of $u(t)$. Then

$$\lim_{K\to\infty} |\langle u^n(0), u^n(t)\rangle_\rho - \langle u_K^n(0), u_K^n(t)\rangle_\rho| \qquad \forall n \in \mathbb{N}, \tag{6.85}$$

i.e., $\langle u_K^n(0), u_K^n(t)\rangle_\rho$ converges uniformly to $\langle u^n(0), u^n(t)\rangle_\rho$ as $K \to \infty$.

Proof. Let us define

$$\begin{aligned}
\delta_K(t) &= |\langle u_K^n(t), u_K^n(0)\rangle - \langle u^n(t), u(0)\rangle| \\
&= |\langle u_K^n(t), u_K^n(0)\rangle - \langle u^n(t), u_K^n(0)\rangle + \langle u^n(t), u_K^n(0)\rangle - \langle u^n(t), u^n(0)\rangle| \\
&= |\langle u_K^n(t) - u^n(t), u_K^n(0)\rangle + \langle u^n(t), u_K^n(0) - u^n(0)\rangle| \\
&\leq |\langle u_K^n(t) - u^n(t), u_K^n(0)\rangle| + |\langle u^n(t), u_K^n(0) - u^n(0)\rangle| \qquad (6.86)
\end{aligned}$$

The first term at the right hand side is of the form

$$a^n - b^n = (a-b)\sum_{i=0}^{n-1} a^i b^{n-1-i}.$$

By using the Cauchy-Schwarz inequality, we obtain

$$
\begin{aligned}
|\langle u_K^n(t) - u^n(t), u_K^n(0)\rangle| =& \left|\langle (u_K(t) - u(t)) \sum_{i=0}^{n-1} u_K^i(t) u^{n-1-i}(t), u_K^n(0)\rangle\right| \\
=& \left|\langle u_K(t) - u(t), u_K^n(0) \sum_{i=0}^{n-1} u_K^i(t) u^{n-1-i}(t)\rangle\right| \\
\leq& \epsilon_K(t) \left\| u_K^n(0) \sum_{i=1}^{n-1} u_K^i(t) u^{n-1-i}(t) \right\|_{L^2},
\end{aligned}
$$

where we defined $\epsilon_K(t) = \|u_K(t) - u(t)\|_{L^2}$. It is well-known that $\epsilon_K(t) \to 0$ as $K \to \infty$ (see, e.g., [62]). By using the generalized Hölder's inequality $\|fg\|_{L^p} \leq \|f\|_{L^q} \|g\|_{L^q}$, where $2p = q$ and the Minkowski inequality, we obtain

$$
\begin{aligned}
|\langle u_K^n(t) - u^n(t), u_K^n(0)\rangle| \leq& \epsilon_K(t) \|u_K^n(0)\|_{L^4} \sum_{i=1}^{n-1} \|u_K^i(t) u^{n-i-1}(t)\|_{L^4} \\
\leq& \epsilon_K(t) \|u_K^n(0)\|_{L^4} \sum_{i=1}^{n-1} \|u_K^i(t)\|_{L^8} \|u^{n-i-1}(t)\|_{L^8} = C_1 \epsilon_K(t),
\end{aligned}
\tag{6.87}
$$

where

$$
C_1 = \|u_K^n(0)\|_{L^4} \sum_{i=1}^{n-1} \|u_K^i(t)\|_{L^8} \|u^{n-i-1}(t)\|_{L^8} < \infty. \tag{6.88}
$$

Similarly, we have

$$
|\langle u^n(t), u_K^n(0) - u^n(0)\rangle| \leq \epsilon_K(0) \|u^n(0)\|_{L^4} \sum_{i=1}^{n-1} \|u_K^i(0)\|_{L^8} \|u^{n-i-1}(0)\|_{L^8} = C_2 \epsilon_K(0). \tag{6.89}
$$

By combining (6.86), (6.87) and (6.89), we obtain

$$
\lim_{K \to +\infty} \delta_K(t) \leq \lim_{K \to +\infty} C_1 \epsilon_K(t) + C_2 \epsilon_K(0) = 0,
$$

which proves the theorem. □

Appendix 6.B Sampling algorithm for non-Gaussian processes

In [78, 79], Phoon et al proposed an iterative algorithm to generate samples from a non-Gaussian stochastic process by using the KL expansion. In this section, we will briefly review this algorithm and present simulation results for the non-Gaussian process $r_j(t)$ appeared in Section 6.3.2. To this end, we consider a centred stationary stochastic process $w(x;\theta)$ with the marginal distribution $w(x_i;\theta) \sim \rho_w(x, \boldsymbol{\xi})$ and covariance function $C(x,s)$, where $\boldsymbol{\xi}$ is the parameter set that characterises the marginal distribution. The corresponding marginal cumulative distribution function is denoted as $F_M(x) = F_M(x, \boldsymbol{\xi})$. The stochastic process $w(x;\theta)$ can be approximated by a truncated KL series

$$w(x;\theta) \simeq \sum_{k=1}^{K} \sqrt{\lambda_k} w_k(\theta) e_k(x)$$

Phoon's algorithm generate samples for random vector $\boldsymbol{w}(\theta) = [w_1(\theta), \cdots, w_K(\theta)]$ such that the empirical mariginal distribution $w(x_i, \theta_n) = \sum_{k=1}^{K} \sqrt{\lambda_k} w_k(\theta_n) e_k(x_i) \sim \tilde{\rho}_w(x, \boldsymbol{\xi})$ approximates $\rho_w(x, \boldsymbol{\xi})$ for all x_i. This is an easy task for Gaussian process since the linear combination of Gaussian random variables is still Gaussian. For non-Gaussian processes, the iterative Algorithm 1 is used to generate $\boldsymbol{w}(\theta)$. In the first step of Algorithm 1, we use numerical integration method to approximate the eigenfunction $e_k(x)$ and eigenvalue λ_k of the KL series with the given covariance function $C(s,x)$. Specifically, we approximate the Fredholm integral equation as follows:

$$\int_D C(s,x) e_k(x) dx \approx \sum_j C(s, x_j) e_k(x_j) \omega_j = \lambda_k e_k(s) \tag{6.90}$$

Algorithm 1: Iterative algorithms for sampling $w(x;\theta)$

KL-decomposition: *Using numerical integration method (see below) to approximate* $e_k(t)$ *and* λ_k *of the KL expansion.*;

Initialization: *Generate* N *sample functions* $w^{(1)}(x;\theta_n)=\sum_{k=1}^{K}\sqrt{\lambda_k}w_k^{(1)}(\theta_n)e_k(x)$ *from the non-Gaussian process, where* N *is the sample number, index* (1) *is the iteration number,* $w_k^{(1)}(\theta_n)$ *are samples of i.i.d standard normal distribution*;

for $1\le m\le M$ **do**

- Estimating the empirical cumulative marginal distribution function $\tilde{F}_M^{(m)}(y|x)=\frac{1}{N}\sum_{n=1}^{N}I(w^{(m)}(x;\theta_n)\le y)$, where I is the indicator.
- Calculate transformed cumulative marginal distribution function as $\eta^{(k)}(x;\theta_n)=F_M^{-1}\tilde{F}_M^{(k)}[w^{(m)}(x;\theta_n)]$.
- Estimate the next generation $w_k^{(m+1)}(\theta)$ as $w_k^{(m+1)}(\theta_n)=\frac{1}{\sqrt{\lambda_k}}\int_D\{\eta^{(m)}(x;\theta_n)-[\sum_n^N\eta^{(m)}(x;\theta_n)]/N\}e_k(x)dx.$
- Standardize $w_k^{(m+1)}(\theta)$ to unit variance.

end

where x_j and ω_j are the suitable choice of nodes and weights for numerical integration in domain D, which can be chosen accordingly for different integration schemes. Since $\langle e_k, e_{k'}\rangle=\delta_{kk'}$. In the grid points $[x_0,x_1,\cdots,x_n]$, the above integral equation can be written as the matrix equation

$$\boldsymbol{C\Omega e}_k=\lambda_k\boldsymbol{e}_k$$

where

$$
\boldsymbol{C} = \begin{bmatrix} C(x_0, x_0) & C(x_0, x_1) & \dots & C(x_0, x_n) \\ C(x_1, x_0) & C(x_1, x_1) & \dots & C(x_1, x_n) \\ \dots & & & \\ C(x_n, x_0) & C(x_0, x_0) & \dots & C(x_n, x_n) \end{bmatrix} \quad \boldsymbol{\Omega} = \begin{bmatrix} \omega_0 & 0 & \dots & 0 \\ 0 & \omega_1 & \dots & 0 \\ \dots & & & \\ 0 & 0 & \dots & \omega_n \end{bmatrix}
$$

and $\boldsymbol{e}_k^T = [e_k(x_0), \cdots, e_k(x_n)]$. By solving the eigenvalue problem $\boldsymbol{C\Omega} e_k = \lambda_k e_k$, one can get the approxmated eigenvalue λ_k and point-wise eigenfunction $e_k(x)$ in $[x_0, \cdots, x_n]^T$. The orginal version of the sampling algorithm in [78, 79] used a Latin Hypercube sampling technique to reduce the correlation between the iterated samples $w_k^{(m+1)}(\theta)$. In the numerical experiments, we have noticed that if the initial samples $\omega_k^{(1)}(\theta_n)$ are chosen from i.i.d standard normal distribution, then the correlation between $\omega_k^{(m+1)}(\theta_n)$ is already very small after several the iterations. Therefore this step is skipped in our implementation to order to speed up the sampling algorithm.

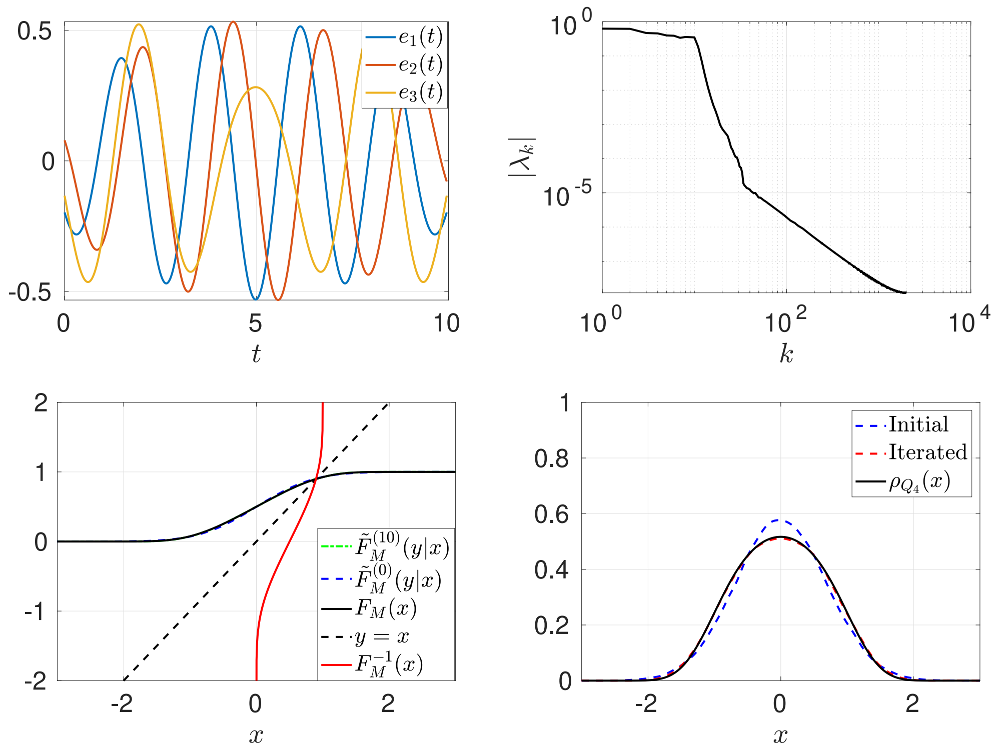

Figure 6.9: KL approximation to a non-Gaussian process $r_j(t)$. In the first row, we show the approximated eigenfunctions (the first three) and eigenvalues for the Fredholm equation (6.90). In the second row, we show the convergence of the sample marginal CDF $\tilde{F}_M(y|x)$, PDF $\tilde{\rho}_{Q_4}(x)$ to the the exact marginal CDF $F_M(x)$ and PDF $\rho_{Q_4}(x)$. The thermodynamic parameter $\gamma = 1$.

Chapter 7

Conclusion

Since the first establishment in 60s, the Mori-Zwanzig framework has gradually became a standard formalism of dimension reduction problems in high-dimensional dynamical systems. Being an abstract and convoluted operator equation, the analysis and approximation of the MZ equation is a daunting task which greatly restricts the development and application of this framework for a rather long time. In this dissertation, we developed a thorough mathematical study on the Mori-Zwanzig equation from both theoretical and numerical points of view. After the introduction of the MZ equation in Chapter 3, we used two chapters (3 and 4) to analyze the semigroup $e^{t\mathcal{G}}$, $e^{t\mathcal{Q}\mathcal{G}\mathcal{Q}}$ and derive from it the prior estimates of the MZ memory integral and the fluctuation force. In particular, we used the Hörmander analysis established for linear hypoelliptic equation in the analysis of the MZ memory integral, and obtained an accurate exponential decaying estimate. As far as we know, this is the first estimate of the MZ memory kernel for stochastic systems driven by white noise. For the computational part, In Chapter 5 we introduced new approximation method for the MZ equation which is based on the orthogonal semigroup expansion theory. The series expansion methods can be proven to be R-superlinearly convergent for linear dynamical systems. Moreover, we also developed a combinatorial computation method which allows us to approximate and solve the MZ equation for nonlinear system from the first principle (Chapter 6). In conclusion, we provided new frameworks to analyze and approximate the Mori-Zwanzig equation. Based on the presented work, further development of the MZ formulation can be expected.

Bibliography

[1] G. Amati, H. Meyer, and T. Schilling. Memory effects in the Fermi–Pasta–Ulam model. *J. Stat. Phys.*, 174(1):219–257, 2019.

[2] N. Aubry, R. Guyonnet, and R. Lima. Spatiotemporal analysis of complex signals: theory and applications. *J. Stat. Phys.*, 64:683–739, 1991.

[3] N. Aubry and R. Lima. Spatiotemporal and statistical symmetries. *J. Stat. Phys.*, 81(3/4):793–828, 1995.

[4] R. J. Baxter. *Exactly solved models in statistical mechanics*. Elsevier, 2016.

[5] A. Bazzani, G. Bassi, and G. Turchetti. Diffusion and memory effects for stochastic processes and fractional Langevin equations. *Physica A*, 324(3-4):530–550, 2003.

[6] M. Berkowitz, J. D. Morgan, D. J. Kouri, and J. A. McCammon. Memory kernels from molecular dynamics. *J. Chem. Phys*, 75(5):2462–2463, 1981.

[7] B. J. Berne and G. D. Harp. On the calculation of time correlation functions. *Adv. Chem. Phys.*, pages 63–227, 1970.

[8] N. Biggs. *Algebraic graph theory*. Cambridge University Press, 1993.

[9] B. Bollobás. *Random Graphs*. Cambridge University Press, 2001.

[10] F. F. Bouchet and M. Corvellec. Invariant measures of the 2D Euler and Vlasov equations. *J. Stat. Mech.: Theory and Experiment*, 2010(08):P08021, 2010.

[11] C. Brennan and D. Venturi. Data-driven closures for stochastic dynamical systems. *J. Comp. Phys.,*, 372:281–298, 2018.

[12] E. Cances, F. Legoll, and G. Stoltz. Theoretical and numerical comparison of some sampling methods for molecular dynamics. *ESAIM: Mathematical Modelling and Numerical Analysis*, 41(2):351–389, 2007.

[13] A. Chertock, D. Gottlieb, and A. Solomonoff. Modified optimal prediction and its application to a particle-method problem. *J. Sci. Comp.*, 37(2):189–201, 2008.

[14] H. Cho, D. Venturi, and G. E. Karniadakis. Karhunen-Loevè expansion for multi-correlated stochastic processes. *Prob. Eng. Mech.*, 34:157–167, 2013.

[15] A. J. Chorin, O. H. Hald, and R. Kupferman. Optimal prediction and the Mori-Zwanzig representation of irreversible processes. *Proc. Natl. Acad. Sci. USA*, 97(7):2968–2973, 2000.

[16] A. J. Chorin, O. H. Hald, and R. Kupferman. Optimal prediction with memory. *Physica D*, 166(3-4):239–257, 2002.

[17] A. J. Chorin, R. Kupferman, and D. Levy. Optimal prediction for Hamiltonian partial differential equations. *J. Comp. Phys.*, 162(1):267–297, 2000.

[18] A. J. Chorin and P. Stinis. Problem reduction, renormalization and memory. *Comm. App. Math. and Comp. Sci.*, 1(1):1–27, 2006.

[19] W. Chu and X. Li. The Mori-Zwanzig formalism for the derivation of a fluctuating heat conduction model from molecular dynamics. *ArXiv:1709.05928*, 2017.

[20] E. B. Davies. Semigroup growth bounds. *J. Operator Theory*, 53(2):225–249, 2005.

[21] S. Denisov, W. Horsthemke, and P. Hänggi. Generalized Fokker-Planck equation: Derivation and exact solutions. *Eur. Phys. J. B*, 68(4):567–575, 2009.

[22] J. M. Dominy and D. Venturi. Duality and conditional expectations in the Nakajima-Mori-Zwanzig formulation. *J. Math. Phys.*, 58(8):082701, 2017.

[23] R. Donninger and B. Schörkhuber. Stable blow up dynamics for energy supercritical wave equations. *Trans. Amer. Math. Soc.*, 366, 2014.

[24] J-P Eckmann and M. Hairer. Non-equilibrium statistical mechanics of strongly anharmonic chains of oscillators. *Commun. Math. Phys.*, 212(1):105–164, 2000.

[25] J-P Eckmann and M. Hairer. Spectral properties of hypoelliptic operators. *Commun. Math. Phys.*, 235(2):233–253, 2003.

[26] J-P Eckmann, C-A Pillet, and Luc Rey-Bellet. Non-equilibrium statistical mechanics of anharmonic chains coupled to two heat baths at different temperatures. *Commun. Math. Phys.*, 201(3):657–697, 1999.

[27] M. Eiermann. On semi-iterative methods generated by Faber polynomials. 56:139–156, 1989.

[28] K.-J. Engel and R. Nagel. *One-parameter semigroups for linear evolution equations*, volume 194. Springer, 1999.

[29] P. Espanol. Hydrodynamics from dissipative particle dynamics. *Phys. Rev. E*, 52(2):1734, 1995.

[30] P. Espanol and P. Warren. Statistical mechanics of dissipative particle dynamics. *EPL*, 30(4):191, 1995.

[31] G. Faber. Über polynomische entwickelunge. *Mathematische Annalen*, 57:389–408, 1903.

[32] F. Flandoli. Dissipativity and invariant measures for stochastic Navier-Stokes equations. *Nonlin. Diff. Eqn. Appl.*, 1(4):403–423, 1994.

[33] J. Florencio and H. M. Lee. Exact time evolution of a classical harmonic-oscillator chain. *Phys. Rev. A*, 31(5):3231, 1985.

[34] D. Givon, R. Kupferman, and O. H. Hald. Existence proof for orthogonal dynamics and the Mori-Zwanzig formalism. *Israel Journal of Mathematics*, 145(1):221–241, 2005.

[35] D. Givon, R. Kupferman, and A. Stuart. Extracting macroscopic dynamics: model problems and algorithms. *Nonlinearity*, 17(6):R55, 2004.

[36] W. Götze. Recent tests of the mode-coupling theory for glassy dynamics. *J. Phys.: cond. matter*, 11(10A):A1, 1999.

[37] J.-P. Hansen and I. R. McDonald. *Theory of simple liquids*. Elsevier, 1990.

[38] B. Helffer and F. Nier. *Hypoelliptic estimates and spectral theory for Fokker-Planck operators and Witten Laplacians*. Springer, 2005.

[39] F. Hérau and F. Nier. Isotropic hypoellipticity and trend to equilibrium for the Fokker-Planck equation with a high-degree potential. *Arch. Ration. Mech. Anal*, 171(2):151–218, 2004.

[40] J. S. Hesthaven, S. Gottlieb, and D. Gottlieb. *Spectral methods for time-dependent problems*. Cambridge Univ. Press, 2007.

[41] C. Hijón, P. Español, E. Vanden-Eijnden, and R. Delgado-Buscalioni. Mori–Zwanzig formalism as a practical computational tool. *Faraday discussions*, 144:301–322, 2010.

[42] P. D. Hislop and I. S. Sigal. *Introduction to spectral theory: With applications to Schrödinger operators*, volume 113. Springer Science & Business Media, 2012.

[43] L. Hörmander. Hypoelliptic second order differential equations. *Acta. Math*, 119(1):147–171, 1967.

[44] T. Hudson and X. Li. Coarse-graining of overdamped langevin dynamics via the Mori-Zwanzig formalism. *arXiv:1810.08175*, 2018.

[45] W. Huisinga, L. Pesce, R. Kosloff, and P. Saalfrank. Faber and Newton polynomial integrators for open-system density matrix propagation. *J. Chem. Phys*, 110(12):5538–5547, 1999.

[46] N. G. Van Kampen and I. Oppenheim. Brownian motion as a problem of eliminating fast variables. *Physica A: Stat. Mech. Appl.*, 138(1-2):231–248, 1986.

[47] A. Karimi and M. R. Paul. Extensive chaos in the Lorenz-96 model. *Chaos*, 20(4):043105(1–11), 2010.

[48] Tosio Kato. *Perturbation theory for linear operators*, volume 132. Springer Science & Business Media, 2013.

[49] H. P. Mc Kean and K. L. Vaninsky. Statistical mechanics of nonlinear wave equations. In *Trends and perspectives in applied mathematics*, pages 239–264. Springer, 1994.

[50] J. Kim and I. Sawada. Dynamics of a harmonic oscillator on the Bethe lattice. *Phys. Rev. E*, 61(3):R2172, 2000.

[51] S. Klainerman. Global existence for nonlinear wave equations. *Comm. Pure and Appl. Math.*, 33(1):43–101, 1980.

[52] P. E. Kloeden and E. Platen. *Numerical solution of stochastic differential equations*, volume 23. Springer Science & Business Media, 2013.

[53] B. O. Koopman. Hamiltonian systems and transformation in Hilbert spaces. *Proc. Natl. Acad. Sci.*, 17(5):315–318, 1931.

[54] R. Kubo. The fluctuation-dissipation theorem. *Reports on Progress in Physics*, 29(1):255, 1966.

[55] H. Kunita. *Stochastic flows and stochastic differential equations*, volume 24. Cambridge university press, 1997.

[56] H. M. Lee. Solutions of the generalized Langevin equation by a method of recurrence relations. *Phys. Rev. B*, 26(5):2547, 1982.

[57] H. Lei, N. Baker, and X. Li. Data-driven parameterization of the generalized Langevin equation. *Proc. Natl. Acad. Sci.*, 113(50):14183–14188, 2016.

[58] Z. Li, , X. Bian, X. Li, and G. E. Karniadakis. Incorporation of memory effects in coarse-grained modeling via the Mori-Zwanzig formalism. *J. Chem. Phys*, 143:243128, 2015.

[59] Z. Li, H. S. Lee, E. Darve, and G. E. Karniadakis. Computing the non-Markovian coarse-grained interactions derived from the Mori-Zwanzig formalism in molecular systems: Application to polymer melts. *J. Chem. Phys*, 146:014104, 2017.

[60] E. N. Lorenz. Predictability - A problem partly solved. In *ECMWF seminar on predictability: Volume 1*, pages 1–18, 1996.

[61] F. Lu, K. Lin, and A. J. Chorin. Data-based stochastic model reduction for the kuramoto–sivashinsky equation. *Physica D*, 340:46–57, 2017.

[62] O. Le Maître and O. M. Knio. *Spectral methods for uncertainty quantification: with applications to computational fluid dynamics*. Springer Science & Business Media, 2010.

[63] C. B. Mendl and H. Spohn. Current fluctuations for anharmonic chains in thermal equilibrium. *J. Stat. Mech.: Theory and Experiment*, 2015(3):P03007, 2015.

[64] C. B. Mendl and H. Spohn. Low temperature dynamics of the one-dimensional discrete nonlinear Schrödinger equation. *J. Stat. Mech. Theory Exp.*, 2015(8):P08028, 2015.

[65] C. Moler and C. Van Loan. Nineteen dubious ways to compute the exponential of a matrix. *SIAM review*, 20(4):801–836, 1978.

[66] C. Moler and C. Van Loan. Nineteen dubious ways to compute the exponential of a matrix, twenty-five years later. *SIAM review*, 45(1):3–49, 2003.

[67] H. Mori. A continued-fraction representation of the time-correlation functions. *Progr. Theor. Phys.*, 34(3):399–416, 1965.

[68] H. Mori. Transport, collective motion, and Brownian motion. *Prog. Theor. Phys.*, 33(3):423–455, 1965.

[69] M. E. J. Newman, S. H. Strogatz, and D. J. Watts. Random graphs with arbitrary degree distributions and their applications. *Phys. Rev. E*, 64:026118, 2001.

[70] P Novati. Solving linear initial value problems by Faber polynomials. *Numerical linear algebra with applications*, 10(3):247–270, 2003.

[71] S. A. Orszag and L. R. Bissonnette. Dynamical properties of truncated Wiener-Hermite expansions. *Phys. Fluids*, 10(12):2603âĂŞ2613, 1967.

[72] M. Ottobre, G. A. Pavliotis, and K. P. Starov. Exponential return to equilibrium for hypoelliptic quadratic systems. *arXiv preprint arXiv:1106.2326*, 2011.

[73] E. J. Parish and K. Duraisamy. A dynamic subgrid scale model for large eddy simulations based on the Mori–Zwanzig formalism. *J. Comp. Phys.*, 349:154–175, 2017.

[74] E. J. Parish and K. Duraisamy. Non-Markovian closure models for large eddy simulations using the Mori-Zwanzig formalism. *Phys. Rev. Fluids*, 2(1):014604, 2017.

[75] G. Parisi. *Statistical field theory*. Addison-Wesley, 1988.

[76] G. A. Pavliotis. *Stochastic processes and applications: diffusion processes, the Fokker-Planck and Langevin equations*, volume 60. Springer, 2014.

[77] Michael E Peskin. *An introduction to quantum field theory*. CRC Press, 2018.

[78] K. K. Phoon, H. W. Huang, and S. T. Quek. Simulation of second-order processes using Karhunen–Loève expansion. *Computers & Structures*, 80(12):1049–1060, 2002.

[79] K. K. Phoon, H. W. Huang, and S. T. Quek. Simulation of strongly non-Gaussian processes using Karhunen–Loève expansion. *Prob. Eng. Mech.*, 20(2):188–198, 2005.

[80] G. Da Prato and J. Zabczyk. *Ergodicity for infinite dimensional systems*, volume 229. Cambridge University Press, 1996.

[81] D. R. Reichman and P. Charbonneau. Mode-coupling theory. *J. Stat. Mech.: Theory and Experiment*, 2005(05):P05013, 2005.

[82] H. Risken. *The Fokker-Planck equation: methods of solution and applications*. Springer-Verlag, second edition, 1989. Mathematics in science and engineering, vol. 60.

[83] I. Snook. *The Langevin and generalised Langevin approach to the dynamics of atomic, polymeric and colloidal systems.* Elsevier, 2006.

[84] P. Stinis. Stochastic optimal prediction for the Kuramoto–Sivashinsky equation. *Multiscale Modeling & Simulation*, 2(4):580–612, 2004.

[85] P. Stinis. A comparative study of two stochastic model reduction methods. *Physica D*, 213:197–213, 2006.

[86] P. Stinis. Higher order Mori-Zwanzig models for the Euler equations. *Multiscale Modeling & Simulation*, 6(3):741–760, 2007.

[87] P. Stinis. Renormalized Mori–Zwanzig-reduced models for systems without scale separation. *Proc. R. Soc. A*, 471(2176):20140446, 2015.

[88] P. K. Suetin and E. V. Pankratiev. *Series of Faber polynomials.* CRC Press, 1998.

[89] L. N. Trefethen. Pseudospectra of linear operators. *SIAM Review*, 39(3):383–406, 1997.

[90] L. N. Trefethen and M. Embree. *Spectra and pseudospectra: the behavior of nonnormal matrices and operators.* Princeton University Press, 2005.

[91] U. Umegaki. Conditional expectation in an operator algebra I. *Tohoku Math. J.*, 6(2):177–181, 1954.

[92] D. Venturi. On proper orthogonal decomposition of randomly perturbed fields with applications to flow past a cylinder and natural convection over a horizontal plate. *J. Fluid Mech.*, 559:215–254, 2006.

[93] D. Venturi. A fully symmetric nonlinear biorthogonal decomposition theory for random fields. *Physica D*, 240(4-5):415–425, 2011.

[94] D. Venturi. The numerical approximation of nonlinear functionals and functional differential equations. *Physics Reports*, 732:1–102, 2018.

[95] D. Venturi, H. Cho, and G. E. Karniadakis. The Mori-Zwanzig approach to uncertainty quantification. In R. Ghanem, D. Higdon, and H. Owhadi, editors, *Handbook of uncertainty quantification.* Springer, 2016.

[96] D. Venturi and G. E. Karniadakis. Convolutionless Nakajima-Zwanzig equations for stochastic analysis in nonlinear dynamical systems. *Proc. R. Soc. A*, 470(2166):1–20, 2014.

[97] D. Venturi, X. Wan, and G. E. Karniadakis. Stochastic low-dimensional modelling of a random laminar wake past a circular cylinder. *J. Fluid Mech.*, 606:339–367, 2008.

[98] C. Villani. Hypocoercivity. *Memoirs of the American Mathematical Society*, 202(950), 2009.

[99] X. Wan and G. E. Karniadakis. Multi-element generalized polynomial chaos for arbitrary probability measures. *SIAM J. Sci. Comput.*, 28(3):901–928, 2006.

[100] R. O. Watts and I. K. Snook. Perturbation theories in non-equilibrium statistical mechanics: Ii. methods based on memory function formalism. *Molecular Physics*, 33(2):443–452, 1977.

[101] V. Wihstutz and MA. Pinsky. *Diffusion processes and related problems in analysis, volume II: Stochastic flows*, volume 27. Springer Science & Business Media, 2012.

[102] Y. Yoshimoto, I. Kinefuchi, T. Mima, A. Fukushima, T. Tokumasu, and S. Takagi. Bottom-up construction of interaction models of non-markovian dissipative particle dynamics. *Phys. Rev. E*, 88(4):043305, 2013.

[103] Y. Zhu, J. Dominy, and D. Venturi. On the estimation of the Mori-Zwanzig memory integral. *J. Math. Phys.*, 59(10):103501, 2018.

[104] Y. Zhu and D. Venturi. Faber approximation of the Mori-Zwanzig equation. *J. Comp. Phys.*, (372):694–718, 2018.

[105] Y. Zhu and D. Venturi. Generalized Langevin equations for systems with local interactions. *arXiv:1906.04918*, 2019.

[106] Y. Zhu and D. Venturi. Hypoellipticity and Mori-Zwanzig theory for SDEs. 2019.

[107] R. Zwanzig. Ensemble methods in the theory of irreversibility. *J. Chem. Phys.*, 33(5):1338–1341, 1960.

[108] R. Zwanzig. Memory effects in irreversible thermodynamics. *Phys. Rev.*, 124:983–992, 1961.

[109] R. Zwanzig. Approximate eigenfunctions of the Liouville operator in classical many-body systems. *Phys. Rev*, 144(1):170, 1966.

[110] R. Zwanzig. Nonlinear generalized Langevin equations. *J. Stat. Phys.*, 9(3):215–220, 1973.