Analytical and Numerical Study of Lindblad Equations

by

Yu Cao

Department of Mathematics

Duke University

Date: ______________________________

Approved:

____

Jianfeng Lu, Advisor

____

Harold Baranger

____

Jonathan Mattingly

____

Sayan Mukherjee

Dissertation submitted in partial fulfillment of the requirements for the degree of
Doctor of Philosophy in the Department of Mathematics
in the Graduate School of Duke University

2020

ABSTRACT

Analytical and Numerical Study of Lindblad Equations

by

Yu Cao

Department of Mathematics

Duke University

Date: ______________________________

Approved:

____

Jianfeng Lu, Advisor

____

Harold Baranger

____

Jonathan Mattingly

____

Sayan Mukherjee

An abstract of a dissertation submitted in partial fulfillment of the requirements for
the degree of Doctor of Philosophy in the Department of Mathematics
in the Graduate School of Duke University

2020

Abstract

Lindblad equations, since introduced in 1976 by Lindblad, and by Gorini, Kossakowski, and Sudarshan, have received much attention in many areas of scientific research. Around the past fifty years, many properties and structures of Lindblad equations have been discovered and identified. In this dissertation, we study Lindblad equations from three aspects: (I) physical perspective; (II) numerical perspective; and (III) information theory perspective.

In Chp. 2, we study Lindblad equations from the physical perspective. More specifically, we derive a Lindblad equation for a simplified Anderson-Holstein model arising from quantum chemistry. Though we consider the classical approach (i.e., the weak coupling limit), we provide more explicit scaling for parameters when the approximations are made. Moreover, we derive a classical master equation based on the Lindbladian formalism.

In Chp. 3, we consider numerical aspects of Lindblad equations. Motivated by the dynamical low-rank approximation method for matrix ODEs and stochastic unraveling for Lindblad equations, we are curious about the relation between the action of dynamical low-rank approximation and the action of stochastic unraveling. To address this, we propose a stochastic dynamical low-rank approximation method. In the context of Lindblad equations, we illustrate a commuting relation between the dynamical low-rank approximation and the stochastic unraveling.

In Chp. 4, we investigate Lindblad equations from the information theory perspective. We consider a particular family of Lindblad equations: primitive Lindblad equations with GNS-detailed balance. We identify Riemannian manifolds in which these Lindblad equations are gradient flow dynamics of sandwiched Rényi divergences. The necessary condition for such a geometric structure is also studied. Moreover, we study the exponential convergence behavior of these Lindblad equations to their equilibria, quantified by the whole family of sandwiched Rényi divergences.

Contents

List of Tables

List of Figures

Notations

$\mathfrak{M}$	The space of all $n \times n$ (complex-valued) matrices
$\mathfrak{A}$	The space of all $n \times n$ Hermitian matrices
$\mathfrak{A}_0$	The space of traceless elements in $\mathfrak{A}$
$\mathfrak{P}$	The space of positive semidefinite matrices
$\mathfrak{P}^+$	The space of strictly positive matrices
$\mathfrak{D}$	The space of density matrices
$\mathfrak{D}^+$	The space of invertible density matrices
$\mathcal{L}^\dagger$	Lindblad super-operator
$\mathcal{R}^\dagger$	Redfield super-operator
$\dagger$	Adjoint
H	Hamiltonian
ρ, σ, and etc.	Positive semidefinite matrices
$\mathbb{I} \equiv \mathbb{I}_n$	Identity matrix with dimension n
Id	Identity super-operator
$\mathcal{P}$	Projection operator
$\mathcal{Q} = \mathrm{Id} - \mathcal{P}$	Projection operator to the orthogonal complement
$[A, B] := AB - BA$	Commutator
$\{A, B\} := AB + BA$	Anti-commutator
$\{\cdot,\cdot\}_{Pois}$	Poisson bracket
$\\|\cdot\\|_{HS}$	Hilbert-Schmidt norm
$\langle A, B\rangle_{HS} = \mathrm{Tr}(A^\dagger B)$	Hilbert-Schmidt inner product

Abbreviations

CME	Classical master equation
QME	Quantum master equation
QMS	Quantum Markov semigroup
SDE	Stochastic differential equation
GNS	Gel'fand-Naimark-Segal
KMS	Kubo-Martin-Schwinger
BKM	Bogoliubov- Kubo-Mori
JKO	Jordan-Kinderlehrer-Otto
LSI	Logarithmic Sobolev inequality

Acknowledgements

I would like to thank my advisor Prof. Jianfeng Lu for his consistent encouragement, help, and guidance. He taught me not only how to approach specific problems, but also how to think as an applied mathematician.

Besides, I would like to thank Prof. Harold Baranger, Prof. Jonathan Mattingly, and Prof. Sayan Mukherjee for serving in my committee, as well as useful advice. Moreover, I would like to thank my collaborators Dr. Yulong Lu and Mr. Lihan Wang.

I also want to thank many other faculties, scholars, and staffs at Duke University (too many to list here). They help to create a friendly and positive environment so that I could enjoy studying and working.

Last but not least, I would like to thank my parents Bing Cao and Luying Cao, and my fiancée Yu Shi for their love and support.

Chapter 1

Introduction to Lindblad equations

The open quantum system studies the interaction between a primary system of interest and a bath system (also known as the environment). Its study plays a significant role in the development of quantum chemistry, quantum optics, quantum information theory, etc. Many interesting models in open quantum systems include Jaynes-Cummings model [132], spin-Boson model [22], molecular electronics [76], Anderson-Holstein model [5, 57, 86].

Typically, a fundamental challenge in studying open quantum systems comes from the vast degrees of freedom of the bath. Therefore, an essential task in open quantum system is to obtain a self-consistent dynamics only including the primary system of interest (in other words, model reduction methods are employed to approximate a sizeable interacting system). Among many self-consistent dynamics, some famous dynamics include quantum master equation (QME) [22], quantum non-Markovian dynamics [21, 22], and Dyson's equation in the non-equilibrium Green's function (NEGF) formalism [56].

Among various QMEs for open quantum systems, the Lindblad equation is perhaps the most popular and widely studied dynamics. Its history dates back into 1976 when it was firstly studied by Lindblad [98] from a functional analysis perspective, and Gorini, Kossakowski, and Sudarshan [81] for N-level systems. Let us consider a continuous time-parameterized semigroup $(\mathcal{P}_t)_{t\geq 0}$ on the space of linear operators on a finite-dimensional Hilbert space $\mathbb{C}^n$ (namely, $\mathcal{P}_t$ is a linear time-evolution operator in the Heisenberg picture). If $\mathcal{P}_t$ is completely positive (CP) and $\mathcal{P}_t(\mathbb{I}_n) = \mathbb{I}_n$ for any $t \geq 0$, then this semigroup dynamics $\mathcal{P}_t$ is called a *quantum Markov semigroup* (QMS). A common and equivalent definition of QMS is that the semigroup $\mathcal{P}_t^\dagger$ is completely positive trace-preserving (CPTP). Denote the generator of the QMS by $\mathcal{L}$, i.e., $\mathcal{P}_t = e^{t\mathcal{L}}$.[1] It is well-known that the generator $\mathcal{L}^\dagger$ (the adjoint operator of $\mathcal{L}$) must have the following form [81, 98]

$$\mathcal{L}^\dagger(\rho) = -i[H, \rho] + \sum_k \Big(L_k \rho L_k^\dagger - \frac{1}{2}\{L_k^\dagger L_k, \rho\}\Big), \tag{1.1}$$

and Lindblad equations refer to the following dynamics

$$\dot{\rho}_t = \mathcal{L}^\dagger(\rho_t). \tag{1.2}$$

In literature, the Lindblad equation is also known as the Lindblad-Gorini-Kossakowski-Sudarshan (LGKS) equation or the Gorini-Kossakowski-Sudarshan-Lindblad (GKSL) equation. As a remark, Lindblad equations and quantum Markov semigroups refer to the same type of dynamics and are used indistinguishably throughout this dissertation. In the above, the Hermitian operator H is the (effective) Hamiltonian, L_k are Lindblad operators, $[\cdot, \cdot]$ is

[1] Quantum master equations also have time-independent generators, while it might not satisfy the CP condition. Therefore, QMS is a special family of QME.

the commutator, and $\{\cdot,\cdot\}$ is the anti-commutator. Throughout this dissertation, the symbol $\mathcal{L}$ is used as the generator in the Heisenberg picture, whereas its adjoint $\mathcal{L}^\dagger$ is the generator in the Schrödinger picture; the same notation convention applies to the Redfield equation for the symbol $\mathcal{R}^\dagger$. Lindblad equations have been widely employed in or connect to many areas of scientific research. For instance, quantum optics [38], superconductivity [93], dissipative quantum computation [137], quantum thermodynamics [3].

As a remark, we would like to mention the classical analog of the Lindblad equation, which is the Fokker-Planck equation for classical open systems. Consider an n-dimensional Brownian particle under an external potential $\mathcal{V}(x)$, where $x \in \mathbb{R}^n$. A famous stochastic dynamics to describe the motion of the Brownian particle is the overdamped Langevin dynamics

$$\mathrm{d}X_t = -\nabla\mathcal{V}(X_t)\ \mathrm{d}t + \sqrt{2}\,\mathrm{d}W_t. \tag{1.3}$$

In the above, the term $-\nabla\mathcal{V}$ represents the force under the potential $\mathcal{V}$, and the term $\mathrm{d}W_t$ represents the diffusion. The following Fokker-Planck equation describes the time evolution of the probability density function $\varrho_t(x)$

$$\partial_t\varrho_t(x) = \mathrm{div}\left(\varrho_t(x)\nabla\mathcal{V}(x) + \nabla\varrho_t(x)\right). \tag{1.4}$$

The unique stationary distribution is proportional to $e^{-\mathcal{V}(x)}$. Therefore the Langevin dynamics (1.3) is often used in the Monte Carlo sampling of log-concave distributions $e^{-\mathcal{V}(x)}$; see, e.g., [49, 136]. Though there is no direct connection between the Fokker-Planck equation and the Lindblad equation above, as far as we know, these two equations resemble in several ways. In Chp. 3, we shall discuss the numerical methods of these two equations under the same framework; In Chp. 4, almost all results for Lindblad equations are motivated by the results for Fokker-Planck equations. We will observe their relationship in details in Chp. 3 and Chp. 4 below.

In this dissertation, we will discuss three aspects of Lindblad equations presented in the following three chapters, respectively.

Physical perspective

In Chp. 2, we consider the (simplified) Anderson-Holstein (AH) model [5, 57, 86], used for electron transport through molecular electronics in quantum chemistry. For the AH model, the Redfield-type quantum master equation has been derived and studied in [62, 65, 66]; however, the Lindblad formalism has not. Therefore, one main result of Chp. 2 is to derive a Lindblad equation for the AH model under the weak coupling limit (see Sec. 2.3 and Sec. 2.4). The primary approach is to use the time-convolutionless equation formalism to obtain the Redfield equation (2.9) under the weak coupling limit (see, e.g., [22, 52]), and then apply the secular approximation to obtain the Lindblad equation in both interaction picture (see (2.19)) and Schrödinger picture (see (2.10)). We notice that Redfield and Lindblad equations are at the same level, and the condition for the secular approximation is, in fact, a redundant condition, compared to the condition used for the weak-coupling limit expansion. Moreover, we also study the semiclassical limit of the resulting Lindblad equation in Sec. 2.5 and compare Redfield and Lindblad equations from the perturbation theory perspective in Sec. 2.6.

Numerical perspective

In many scenarios, Lindblad equations have very high-dimension, and it is thus challenging to solve numerically. If the quantum system has the dimension n, then we need $\mathcal{O}(n^2)$ space at each time to store the (full) density matrix.

There have been extensive studies on stochastic unraveling methods (or say Monte Carlo methods) for Lindblad equations since the last 90s; see, e.g., [25, 50, 78, 120]: the main idea is to find a stochastic process for wave functions ψ_t such that $\mathbb{E}[\psi_t\psi_t^\dagger]$ solves the Lindblad equation (1.2). Since then, these Monte Carlo methods have been adapted and extended to Redfield equations [92], and non-Markovian dynamics [23, 104]. An alternative approach to achieve the model reduction is the (dynamical) low-rank approximation to density matrices [91, 95]. We use $\rho_{LR,t} \approx \rho_t$ with moving basis, and the time-evolution dynamics of $\rho_{LR,t}$ is given by projecting the exact (local) tangent vector to the tangent space of a sub-manifold. In [96], these two ideas are combined to achieve further efficiency.

Though plenty of methods are known, one question is that when we apply both dynamical low-rank approximation and stochastic unraveling method one after the other, can we obtain the same resulting low-rank stochastic unraveling method? Motivated by this, we propose a stochastic dynamical low-rank approximation (SDLR) method for general finite-dimensional SDEs (see Sec. 3.2). Then we apply such a low-rank method to the stochastic unraveling algorithm of Lindblad equations. Under certain conditions, we can indeed observe a commuting relation between the dynamical low-rank approximation method and the stochastic unraveling method (see Sec. 3.5.2 for more details). As a remark, our SDLR method is proposed in a general set-up and has broad applications beyond the purpose of solving Lindblad equations; see Sec. 3.6 for more examples.

Information theory perspective

After the onset of information theory established by Claude Shannon in 1948 [131], entropy has become a critical concept and tool to study both classical and quantum dynamics. There are many examples of entropies, e.g., the relative entropy (also known as the Kullback-Leibler divergence [94]) and Rényi divergences [125]. Information theory has been developing rapidly, and it becomes a vast area of research; in this dissertation, there are two relevant research topics: the first one is the (Wasserstein) information geometry, and the second one is the entropy functional inequality.

Information geometry

The relation between the information geometric structure and the Fokker-Planck equation was disclosed in the celebrated work of Jordan, Kinderlehrer, and Otto [88]: they showed that the Fokker-Planck equation (1.4) could be regarded as gradient flow dynamics of the relative entropy for the Wasserstein metric. Since then, many classical dynamical systems have been studied and are identified as gradient flow dynamics of certain entropies; see, e.g., [29, 63, 64, 100, 115]. Due to its success for classical dynamics and the increasing importance of quantum information theory, it is interesting to generalize such geometric structures to quantum dynamics, which was initiated and developed by a series of works of Carlen and Maas [35–37]. In particular, in [36], they identified a Riemannian manifold such that primitive Lindblad equations with GNS-detailed balance are gradient flow dynamics of the quantum relative entropy. This also motivates the introduction of quantum Wasserstein distance in the Benamou-Brenier formalism [15]; see discussions in Chp. 5.

Inspired by the importance of sandwiched Rényi divergences in quantum information theory [141] and quantum thermodynamics [19], we try to extend the result of Carlen and Maas in [36] to sandwiched Rényi divergences: in Thm. 4.2, we identify Riemannian manifolds such that primitive Lindblad equations with GNS-detailed balance are gradient flow dynamics of sandwiched Rényi divergences. The necessary condition for a Lindblad equation to have a gradient flow form is also discussed in Thm. 4.24, which extends [37, Thm. 2.9].

Functional inequalities and exponential convergence

Apart from the geometric interpretation of dynamical systems, another important question is to quantify their long-time behavior. The entropy production method [8, 16, 20, 67, 109, 133] assisted with the famous log-Sobolev inequality (LSI) (see, e.g., [11, 53, 82–84, 114, 122]) has been a well-known technique to establish the exponential convergence to the equilibrium for various Markov semigroups. It is natural to consider their generalization in the quantum region. In [36], a quantum LSI based on the entropy production is used to prove the exponential convergence of primitive Lindblad equations with GNS-detailed balance. See, e.g., [51, 123] for more about quantum functional inequalities.

In Chp. 4, we introduce the quantum α-LSI, and in Thm. 4.7, we prove the exponential convergence of primitive Lindblad equations with GNS-detailed balance to their equilibria, quantified by sandwiched Rényi divergences. We remark that such an exponential convergence behavior of primitive Lindblad equations under sandwiched Rényi divergences was also studied in [107]. Our results differ from theirs in the sense that we consider the long-time behavior, and they consider the global contraction of quantum states.

Acknowledgment

This dissertation is prepared based on the following journal publications, in collaboration with Prof. Jianfeng Lu and Dr. Yulong Lu.

[27] Y. Cao and J. Lu, *Lindblad equation and its semiclassical limit of the Anderson-Holstein model*, J. Math. Phys. 58 (2017), no. 12, 122105.

[28] Y. Cao and J. Lu, *Stochastic dynamical low-rank approximation method*, J. Comput. Phys. 372 (2018), 564–586.

[30] Y. Cao, J. Lu, and Y. Lu, *Gradient flow structure and exponential decay of the sandwiched Rényi divergence for primitive Lindblad equations with GNS-detailed balance*, J. Math. Phys. 60 (2019), no. 5, 052202.

Other relevant publications or preprints

Below are some relevant publications or preprints, in collaboration with Prof. Jianfeng Lu, Dr. Yulong Lu, Mr. Lihan Wang. However, these manuscripts are not included in this dissertation.

[33] Y. Cao, J. Lu, and L. Wang, *Complexity of randomized algorithms for underdamped Langevin dynamics*, 2020. arXiv: 2003.09906.

[32] Y. Cao, J. Lu, and L. Wang, *On explicit L^2-convergence rate estimate for underdamped Langevin dynamics*, 2019. arXiv: 1908.04746.

[31] Y. Cao and J. Lu, *Tensorization of the strong data processing inequality for quantum chi-square divergences*, Quantum 3 (2019), 199.

[29] Y. Cao, J. Lu, and Y. Lu, *Exponential decay of Rényi divergence under Fokker–Planck equations*, J. Stat. Phys. 176 (2019), no. 5, 1172–1184.

Chapter 2

The Lindblad equation for the Anderson-Holstein model

2.1 Introduction

Multi-level open quantum systems have received much attention due to their broad applications and intriguing phenomena [72, 76, 77]. One of the simplest models is perhaps the Anderson-Holstein model, a two-level open quantum system [5, 86]. The Anderson-Holstein model is a simplistic model for a molecule as the system of interest, represented by a classical nucleus degree of freedom and a two-level electronic degree of freedom, coupled with a bath of Fermions, for instance, a reservoir of electrons. In this chapter, the simplified version of the Anderson-Holstein model discussed in [57] is used as an example for illustrating purpose, which will be explained in more detail in Sec. 2.2. The goal is to understand the approach of quantum master equations, in particular, the Lindbladian formalism, for such systems in the weak-coupling limit and also to study the semiclassical limit of the quantum master equations. Our study here should generalize to other multi-level open quantum systems.

The Anderson-Holstein model has been widely studied via various theoretical and nu-

This chapter is prepared based on [27].

merical approaches, for instance, the Green's function approach [75, 124], equation-of-motion method [73, 74], quantum Monte Carlo method [106], semiclassical approximation [102, 103], non-crossing approximation [41], and quantum master equations [57–60, 62, 65, 66]. From the perspective of the quantum master equation, which is mostly related to the current work, the quantum master equation in the Redfield flavor for the Anderson-Holstein model has been derived in [62, 65, 66]. The semiclassical limit of the Redfield equation, known as the classical master equation (CME), has been considered in [57], which also proposed a numerical method based on the surface hopping. The CME perspective has then been used to study various physical aspects of the Anderson-Holstein model, e.g., broadening, Marcus rate [58–60]. In all these works, the focus has been on the Redfield equation (or Redfield generator). As far as we know, for the Anderson-Holstein model, not much attention has been put into the Lindbladian formulation nor its semiclassical limit, which are the focuses of this chapter.

Recall that the Hamiltonian can fully characterize closed quantum systems; its time-evolution dynamics is characterized by the Schrödinger equation (or the von Neumann equation if we are dealing with density operators). In the framework of the quantum master equation, open quantum systems can be precisely described by the Nakajima-Zwanzig equation with the assistance of a projection operator to a subspace, in which density operators for the closed system are separable [22]. However, in general, it is not easy to retrieve useful information, neither analytically nor numerically. This difficulty is attributed to the memory effect in the Nakajima-Zwanzig equation. It is, of course, critical to research on non-Markovian dynamics; many questions in non-Markovian dynamics are still open, and the mathematical foundation requires further investigation [21]. Often, we take Markovian approximation to simplify the governing equations, which is a valid approximation in the weak coupling regime; see, e.g., [52].

The Markovian approximation yields the Redfield equation using the time-convolutionless (TCL) projection operator method in the weak-coupling limit [22]. Furthermore, with the secular approximation, the Lindblad equation can be obtained from the Redfield equation [22]; the Lindblad equation possesses better mathematical properties such as complete positivity [98]. The essence of the secular approximation is to approximate fast oscillating terms by zero in the sense of averaging on a coarser time scale. This chapter aims at studying both the Redfield equation and the Lindblad equation for multi-level open quantum systems as well as their semiclassical limit, in the context of the Anderson-Holstein model. For these two equations, we obtain the classical master equations (CME) and Lindbladian classical master equations (LCME) in the semiclassical limit. The relations of various models and the asymptotic limit connecting those are summarized in Fig. 2.1.

It is worth mentioning that there is a debate in the literature on which equation better models the open quantum system, especially when the coupling between the system and

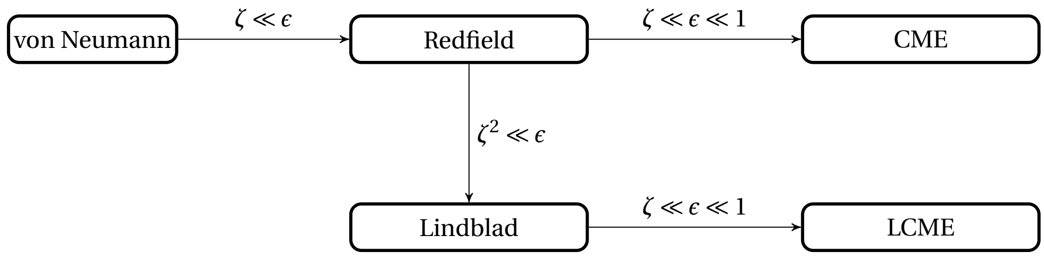


Figure 2.1: This diagram summarizes the conditions needed for approximation and connections between various models. The non-dimensional parameters ζ and ε will be introduced in Sec. 2.2.

the bath is not weak. The underlying discussion focuses on whether complete positivity is necessary for modeling open quantum systems. At least two arguments are supporting the complete positivity in quantum systems: the first one is from the perspective of "total domain"; the second one from "product state" [101]. What is more, one recent research indicates that the lack of complete positivity renders the Redfield equation to be inconsistent with the second law of thermodynamics [7]. Some, however, criticize that we might over-emphasize the importance of complete positivity in modeling open quantum systems. Pechukas proposed that for a composite quantum system with an entangled initial condition, the positivity property might not hold for the reduced dynamics [119]. Shaji and Sudarshan argued that complete positivity is not necessary by carefully examining arguments supporting complete positivity [129]. Negativity, as opposed to positivity, is not only observed in experiments but also can be informative to the coupling with the bath [101]. In our study on the Anderson-Holstein model, imposing complete positivity (and thus Lindblad equation) should be justified as we only consider the weak-coupling regime.

In particular, as will become apparent in our analysis, under the same assumption used in deriving the Redfield equation, the secular approximation for getting the Lindblad equation is, in fact, also justified; hence, the use of Lindblad equation is natural.

In this chapter, we consider the Anderson-Holstein model in the weak-coupling and semiclassical limits. Under the assumption that the coupling strength is weak, we first revisit the derivation of the Redfield equation in Sec. 2.3, and then derive the Lindblad equation in Sec. 2.4. The semiclassical study of both equations is discussed in Sec. 2.5. The perturbation analysis of both equations is presented in Sec. 2.6. Finally, Sec. 2.7 summarizes the main results.

2.2 Anderson-Holstein model

The Anderson-Holstein model under study here describes a two-level system coupled with a bath of many non-interacting electrons (or in general spin-$1/2$ Fermions). For instance, the two-level system can be thought of as a simplistic model for the nuclei degree of freedom of a molecule with two potential energy surfaces depending on the electronic state of the molecule. For simplicity, in the Anderson-Holstein model, the two-level system is in one spatial dimension. One of the potential energy surface is taken to be a harmonic oscillator with frequency ω_s. The potential difference $\mathcal{U}(x)$ is chosen as a linear function of the nucleus position, and thus the other potential surface is also a harmonic oscillator with shifted center and energy [86]. More specifically, the Hamiltonian for the whole system is

$$
\begin{aligned}
H &= H_s + H_b + H_c, \\
H_s &= \frac{\mathbf{p}^2}{2m} + \frac{1}{2} m\omega_s^2 \mathbf{x}^2 + \mathcal{U}(\mathbf{x})\mathbf{d}^\dagger \mathbf{d}, \\
H_b &= \sum_k (E_k - \mu_{\mathcal{F}})\mathbf{c}_k^\dagger \mathbf{c}_k, \\
H_c &= \sum_k V_k (\mathbf{c}_k^\dagger \mathbf{d} + \mathbf{c}_k \mathbf{d}^\dagger),
\end{aligned}
\tag{2.1}
$$

where we follow the notation of [57]: $\mathbf{d}$ and $\mathbf{d}^\dagger$ are the annihilation and creation operators for the two-level electron state of the molecule, $\mathbf{c}_k$ and $\mathbf{c}_k^\dagger$ are the annihilation and creation operators for electron states in the bath, E_k is the energy level of those states, $\mu_{\mathcal{F}}$ is the Fermi level, and V_k is the coupling strength between the molecule and the k-th mode in the bath, assumed to be real. The Hilbert space corresponds to the molecule is thus $L^2(\mathbb{R}) \otimes \mathbb{C}^2 = L^2(\mathbb{R}) \otimes \operatorname{span}\{|0\rangle, |1\rangle\}$, so that we have $\mathbf{d}|1\rangle = |0\rangle$ and $\mathbf{d}^\dagger|0\rangle = |1\rangle$. Thus in (2.1), H_s is the Hamiltonian operator of the "system", H_b is the Hamiltonian of the "bath", and H_c describes the coupling between the system and the bath. The goal is to understand the evolution of the system as an open quantum system (i.e., integrate out the bath degree of freedom). The above is a simplified version of the original Anderson-Holstein model as

(i) only one electron is considered for the molecule, and thus the model excludes the Coulomb interaction between electrons of the molecule;

(ii) the molecule is coupled to only one electrode, and thus only one Fermi level $\mu_{\mathcal{F}}$ is used for the environment.

Remark. In the literature, e.g., [103], sometimes the term "single-level" is used for the above system to emphasize that there is only one on-site electron. We have adopted here the term "two-level" to emphasize that the Hilbert space for the molecular system is $L^2(\mathbb{R}) \otimes \operatorname{span}\{|0\rangle, |1\rangle\}$, and the second component has dimension 2.

To proceed, let us non-dimensionalize the problem according to the following rescaling:

(i) Denote ℓ as the characteristic length scale of x, the position degree of freedom of the system, i.e., if we take $x = \ell\tilde{x}$, $\tilde{x}$ becomes a dimensionless quantity with order $\mathcal{O}(1)$.

(ii) As a consequence, the scaling factor for the molecular energy is $\mathsf{E} = m\omega_s^2\ell^2$. We will use E as the scaling factor for all physical quantities whose dimension is energy (thus including all terms in the Hamiltonian).

(iii) Denote T as the time scale of the evolution of the system. Thus, $t = \mathsf{T}\tilde{t}$ where $\tilde{t} = \mathcal{O}(1)$. Physically, it is reasonable to choose $\mathsf{T} = \frac{1}{\omega_s}$ [130], since this is the time scale of an isolated harmonic oscillator with frequency ω_s.

(iv) As both $\hbar\omega_s$ and E have the energy dimension, the ratio

$$\epsilon := \frac{\hbar\omega_s}{\mathsf{E}}$$

is a dimensionless quantity. In our analysis of the semiclassical limit of the system, we will assume that ϵ is the small parameter.

(v) Let V be the typical interaction strength with dimension as energy, i.e., we assume $\tilde{V}_k := \frac{V_k}{\mathsf{V}} = \mathcal{O}(1)$. The ratio of V and E

$$\zeta := \frac{\mathsf{V}}{\mathsf{E}}$$

is dimensionless. Weak-coupling limit refers to the case $\mathsf{V} \ll \mathsf{E}$ (hence, $\zeta \downarrow 0$): physically, the coupling between the bath and the system is weak, compared to the typical energy scale of the system. Note that by the above rescaling, we have

$$\frac{V_k}{\mathsf{E}} = \zeta \tilde{V}_k.$$

After performing the above rescaling, the non-dimensionalized Hamiltonian becomes

$$\begin{aligned}
H_{non} &= H_{s,non} + H_{b,non} + \zeta H_{c,non}, \\
H_{s,non} &= -\frac{1}{2}\epsilon^2 \nabla_{\tilde{x}}^2 + \frac{1}{2}\tilde{x}^2 + \tilde{\mathcal{U}}(\tilde{x})\mathbf{d}^\dagger \mathbf{d}, \\
H_{b,non} &= \sum_k (\tilde{E}_k - \tilde{\mu_{\mathcal{F}}})\mathbf{c}_k^\dagger \mathbf{c}_k, \\
H_{c,non} &= \sum_k \tilde{V}_k (\mathbf{c}_k^\dagger \mathbf{d} + \mathbf{c}_k \mathbf{d}^\dagger),
\end{aligned}$$

where $\tilde{\mathcal{U}}(\tilde{x}) = \frac{\mathcal{U}(x)}{\mathsf{E}}$, $\tilde{E}_k = \frac{E_k}{\mathsf{E}}$, and $\tilde{\mu_{\mathcal{F}}} = \frac{\mu_{\mathcal{F}}}{\mathsf{E}}$. Moreover, the von Neumann equation becomes

$$i\epsilon\partial_{\tilde{t}}\rho = [H_{non}, \rho]\,,$$

where ρ is the density operator for the closed system.

To simplify notations, by dropping "tilde" and "non", and shifting the energy reference to replace $E_k - \mu_{\mathcal{F}}$ by E_k, we arrive at the Hamiltonian

$$\begin{aligned}
H &= H_s + H_b + \zeta H_c, \\
H_s &= -\frac{1}{2}\epsilon^2\nabla_x^2 + \frac{1}{2}x^2 + \mathcal{U}(x)\mathbf{d}^\dagger\mathbf{d}, \\
H_b &= \sum_k E_k\mathbf{c}_k^\dagger\mathbf{c}_k, \\
H_c &= \sum_k V_k(\mathbf{c}_k^\dagger\mathbf{d} + \mathbf{c}_k\mathbf{d}^\dagger) =: \mathbf{C}^\dagger\mathbf{d} + \mathbf{C}\mathbf{d}^\dagger.
\end{aligned} \tag{2.2}$$

Here, $\mathcal{U}(x)$ is a linear function of the position x with the form (the reason of the specific choice of parametrization will become clear below)

$$\mathcal{U}(x) := \sqrt{2}\mathrm{g}x + \mathrm{g}^2 + \bar{\mathcal{U}}_0,$$

where $\bar{\mathcal{U}}_0$ is known as the renormalized energy, and the weighted annihilation operator $\mathbf{C}$ is defined as

$$\mathbf{C} := \sum_k V_k\mathbf{c}_k.$$

The evolution of the density operator is given by the von Neumann equation

$$i\epsilon\partial_t\rho = [H, \rho]. \tag{2.3}$$

All quantities in (2.2) and (2.3) are dimensionless, and parameters V_k and E_k's are $\mathcal{O}(1)$.

In summary, after non-dimensionalization, the model contains two scaling parameters ϵ and ζ, corresponding to the semiclassical parameter and the coupling strength, respectively. In the rest of the chapter, we will consider the weak-coupling limit and the semiclassical limit. In the weak-coupling limit, we have $\zeta \downarrow 0$, which leads to Redfield and Lindblad equations for fixed ϵ, while the semiclassical limit means $\epsilon \downarrow 0$; see Fig. 2.1 for an overview.

Notice that

$$H_0 := \langle 0|H_s|0\rangle \equiv -\frac{1}{2}\epsilon^2\nabla_x^2 + \frac{1}{2}x^2$$

is a Hamiltonian for a single harmonic oscillator. It is well-known that it has eigenfunctions

$$\phi_k^0 = \mathsf{N}_k^0\,\mathsf{H}_k\left(\frac{x}{\sqrt{\epsilon}}\right)\exp\left(-\frac{x^2}{2\epsilon}\right), \qquad k \in \mathbb{N},$$

where H_k is the k-th Hermite polynomial and N_k^0 is the normalization constant. The corresponding eigenvalue is $\epsilon(k+\frac{1}{2})$. Similarly,

$$H_1 := \langle 1|H_s|1\rangle \equiv -\frac{1}{2}\epsilon^2\nabla_x^2 + \frac{1}{2}(x+\sqrt{2}\mathrm{g})^2 + \bar{\mathcal{U}}_0$$

has eigenfunctions $\phi_k^1 = \phi_k^0(x + \sqrt{2}\mathrm{g})$ and corresponding eigenvalues $\epsilon(k + \frac{1}{2}) + \bar{\mathcal{U}}_0$. Note that this explains the specific parametrization of the linear function $\mathcal{U}(x)$ above. To keep notations and calculations simple, we only consider the Anderson-Holstein model with the renormalized energy $\bar{\mathcal{U}}_0 \equiv 0$. The extension to the general case is straightforward.

2.3 Revisiting the derivation of the Redfield equation

The derivation of the Redfield equation has been well studied and presented in, e.g., [18, 22, 62]. Physically, the Born-Markov approximation is the key to reduce the dynamics of the closed system to a Markovian dynamics of the primary system only. In this section, we will revisit the derivation for the Anderson-Holstein model to set the grounds of our discussion below, using the time-convolutionless equation (TCL) approach following [22, Chp. 9]; we will borrow notations from this reference as well. For fixed ϵ, we consider the weak-coupling limit below, that is, $\zeta \downarrow 0$, while ϵ stays fixed. Hence, $\eta := \frac{\zeta}{\epsilon} \downarrow 0$.

In open quantum systems, it is often more convenient to use the interaction picture (for the uncoupled system and bath), in particular in the weak-coupling limit, since it removes the effect of the fast motion (due to $H_s + H_b$) from the slow motion (due to ζH_c). The operator in the interaction picture is defined as

$$O_I(t) := e^{\frac{i}{\epsilon}(H_s+H_b)t} O e^{-\frac{i}{\epsilon}(H_s+H_b)t}.$$

The von Neumann equation in the interaction picture is

$$\frac{\mathrm{d}}{\mathrm{d}t}\rho_I(t) = -i\frac{\zeta}{\epsilon}\left[H_{c,I}(t), \rho_I(t)\right] =: \eta \mathbf{L}(t)(\rho_I(t)), \tag{2.4}$$

where we have introduced

$$\mathbf{L}(t)(\cdot) := -i\left[H_{c,I}(t), \cdot\,\right] \tag{2.5}$$

as a super-operator acting on density operators. By explicit calculation, we have

$$H_{c,I}(t) = \mathbf{C}_I^\dagger(t)\mathbf{d}_I(t) + \mathbf{d}_I^\dagger(t)\mathbf{C}_I(t),$$

where

$$\mathbf{d}_I(t) = e^{\frac{i}{\epsilon}H_s t}\mathbf{d}e^{-\frac{i}{\epsilon}H_s t}, \qquad \text{and} \qquad \mathbf{C}_I(t) = \sum_k V_k e^{-\frac{i}{\epsilon}E_k t}\mathbf{c}_k.$$

For any trace-class operator A, defined for the whole closed system, we define a projection operator $\mathcal{P}$ by

$$\mathcal{P}(A) := \mathrm{Tr}_b(A) \otimes \rho_{b,eq},$$

where $\mathrm{Tr}_b(\cdot)$ is the partial trace over the bath degree of freedom, $\rho_{b,eq} := e^{-\beta H_b}/Z_b$ is the density operator of the electron bath at thermal equilibrium, with $\beta = \frac{\mathsf{E}}{k_B T}$ (inverse of the rescaled temperature), and Z_b is the partition function. This projection operator $\mathcal{P}$

disentangles the system and the bath, and replaces the bath by the thermal equilibrium. This is a core ingredient in the Born approximation, whose physical reasoning can be found in [18, p. 276]. We also define its orthogonal complement as $\mathcal{Q} := \mathrm{Id} - \mathcal{P}$. For a given density matrix $\rho_I(t)$, $\mathcal{P}\rho_I(t)$ is known as the relevant part, and $\mathcal{Q}\rho_I(t)$ the irrelevant part.

We may formally write down the solution to (2.4) using Green's function as

$$\rho_I(s) = G(t,s)\rho_I(t) = G(t,s)(\mathcal{P}+\mathcal{Q})\rho_I(t), \qquad \text{(for } s \le t\text{)}, \tag{2.6}$$

where $G(t,s) := T_{\rightarrow} \exp\left(-\eta \int_s^t \mathrm{d}s' \, \mathbf{L}(s')\right)$ and $T_{\rightarrow}$ is the anti-chronological time ordering operator. On the other hand, applying the operator $\mathcal{Q}$ to (2.4) gives a differential equation

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{Q}\rho_I(t) = \eta\mathcal{Q}\mathbf{L}(t)\mathcal{P}\rho_I(t) + \eta\mathcal{Q}\mathbf{L}(t)\mathcal{Q}\rho_I(t).$$

By introducing the propagator $\mathcal{G}(t,s) := T_{\leftarrow} \exp\left(\eta \int_s^t \mathrm{d}s' \, \mathcal{Q}\mathbf{L}(s')\right)$, where $T_{\leftarrow}$ is the chronological time-ordering operator, we could derive that

$$\frac{\partial}{\partial s}\left(\mathcal{G}(t,s)\mathcal{Q}\rho_I(s)\right) = \eta\mathcal{G}(t,s)\mathcal{Q}\mathbf{L}(s)\mathcal{P}\rho_I(s).$$

For fixed t, after integrating both sides for the variable s from t_0 to t,

$$
\begin{aligned}
\mathcal{Q}\rho_I(t) =& \ \mathcal{G}(t,t_0)\mathcal{Q}\rho_I(t_0) + \eta \int_{t_0}^t ds\, \mathcal{G}(t,s)\mathcal{Q}\mathbf{L}(s)\mathcal{P}\rho_I(s) \\
\overset{(2.6)}{=}& \ \mathcal{G}(t,t_0)\mathcal{Q}\rho_I(t_0) + \eta \int_{t_0}^t ds\, \mathcal{G}(t,s)\mathcal{Q}\mathbf{L}(s)\mathcal{P}G(t,s)(\mathcal{P}+\mathcal{Q})\rho_I(t) \\
=& \ \mathcal{G}(t,t_0)\mathcal{Q}\rho_I(t_0) + \Sigma(t)(\mathcal{P}+\mathcal{Q})\rho_I(t),
\end{aligned}
$$

where t_0 is the starting time of interest, and $\Sigma(t) := \eta \int_{t_0}^t \mathrm{d}s\ \mathcal{G}(t,s)\mathcal{Q}\mathbf{L}(s)\mathcal{P}G(t,s)$. Assume that at time t_0, the bath is at thermal equilibrium, and the density operator $\rho_I(t_0)$ is separable, i.e., $\rho_I(t_0) = \rho_{s,I}(t_0) \otimes \rho_{b,eq}$. Then $\mathcal{Q}\rho_I(t_0) = 0$, and hence we obtain

$$
\mathcal{Q}\rho_I(t) = (\mathbf{Id} - \Sigma(t))^{-1}\, \Sigma(t)\mathcal{P}\rho_I(t),
$$

if $\mathbf{Id} - \Sigma(t)$ is invertible, which is the case, for instance, when η is so small that $\|\Sigma(t)\| < 1$.

We may also apply the operator $\mathcal{P}$ to (2.4) and get

$$
\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{P}\rho_I(t) &= \eta\mathcal{P}\mathbf{L}(t)\mathcal{P}\rho_I(t) + \eta\mathcal{P}\mathbf{L}(t)\mathcal{Q}\rho_I(t) \\
&= \eta\mathcal{P}\mathbf{L}(t)\mathcal{P}\rho_I(t) + \eta\mathcal{P}\mathbf{L}(t)(\mathbf{Id} - \Sigma(t))^{-1}\Sigma(t)\mathcal{P}\rho_I(t) \\
&= \eta\mathcal{P}\mathbf{L}(t)(\mathbf{Id} - \Sigma(t))^{-1}\mathcal{P}\rho_I(t) \\
&= \Big(\eta\mathcal{P}\mathbf{L}(t)\mathcal{P} + \eta^2\mathcal{P}\mathbf{L}(t)\int_{t_0}^t \mathrm{d}s\ \mathcal{Q}\mathbf{L}(s)\mathcal{P}\Big)\rho_I(t) + \mathcal{O}(\eta^3).
\end{aligned}
$$

In the last step, we perform the asymptotic expansion with respect to η. It could be easily verified that $\mathcal{P}\mathbf{L}(t)\mathcal{P} = 0$ from the definition of $\mathbf{L}$ (2.5). The leading order expansion is

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{P}\rho_I(t) = \eta^2 \int_{t_0}^{t} \mathrm{d}s\ \mathcal{P}\mathbf{L}(t)\mathbf{L}(s)\mathcal{P}\rho_I(t).$$

After replacing $\mathbf{L}$ by its definition (2.5), we arrive at

$$\frac{\mathrm{d}}{\mathrm{d}t}\rho_{s,I}(t) = -\left(\frac{\zeta}{\epsilon}\right)^2 \mathbf{Tr}_b\Big(\int_{t_0}^{t} \mathrm{d}s\Big[H_{c,I}(t), [H_{c,I}(s), \rho_{s,I}(t) \otimes \rho_{b,eq}]\Big]\Big).$$

The leading order expansion of $\frac{\mathrm{d}}{\mathrm{d}t}\rho_{s,I}(t)$ is the same as applying the Born-Markov approximation to the von Neumann equation directly. Change the variable $\tau = t - s$ and push t_0 to approach $-\infty$, then the last equation becomes the *Redfield equation*

$$\frac{\mathrm{d}}{\mathrm{d}t}\rho_{s,I}(t) = -\left(\frac{\zeta}{\epsilon}\right)^2 \mathbf{Tr}_b\left(\int_{0}^{\infty} \mathrm{d}\tau\ \Big[H_{c,I}(t), [H_{c,I}(t-\tau), \rho_{s,I}(t) \otimes \rho_{b,eq}]\Big]\right). \qquad (2.7)$$

There are two formal justifications for pushing t_0 to $-\infty$: if $t_0 = -\infty$, the dynamics does not depend on the initial condition; moreover, if the system evolves from a long time ago, we may as well consider $t_0 = -\infty$.

After opening the double commutator and simplify the equation, we arrive at

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\rho_{s,I}(t) = -\left(\frac{\zeta}{\epsilon}\right)^2 \int_{0}^{\infty} d\tau\ &\big[\mathbf{d}_I(t), \mathbf{d}_I^\dagger(t-\tau)\rho_{s,I}(t)\big]F(t, t-\tau) \\ &+\big[\rho_{s,I}(t)\mathbf{d}_I^\dagger(t-\tau), \mathbf{d}_I(t)\big]G(t, t-\tau) + \text{h.c.},\end{aligned} \qquad (2.8)$$

where time correlation functions

$$F(t,t') := \mathrm{Tr}_b\left(\mathbf{C}_I^\dagger(t)\mathbf{C}_I(t')\rho_{b,eq}\right) = \sum_k V_k^2 \exp\left(\frac{i}{\epsilon}E_k(t-t')\right)\mathsf{f}_{\mathrm{FD}}(E_k),$$

$$G(t,t') := \mathrm{Tr}_b\left(\mathbf{C}_I(t')\mathbf{C}_I^\dagger(t)\rho_{b,eq}\right) = \sum_k V_k^2 \exp\left(\frac{i}{\epsilon}E_k(t-t')\right)(1-\mathsf{f}_{\mathrm{FD}}(E_k)),$$

and where $\mathsf{f}_{\mathrm{FD}}(z) = 1/(1+e^{\beta z})$ is the Fermi-Dirac function.

Transforming (2.8) back into the Schrödinger picture, we end up with

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\rho_s(t) = -\frac{i}{\epsilon}\left[H_s,\rho_s(t)\right] - \left(\frac{\zeta}{\epsilon}\right)^2 \int_0^\infty d\tau\ \Big(\\
\big(\mathbf{d}e^{-\frac{i}{\epsilon}H_s\tau}\mathbf{d}^\dagger e^{\frac{i}{\epsilon}H_s\tau}\rho_s(t) - e^{-\frac{i}{\epsilon}H_s\tau}\mathbf{d}^\dagger e^{\frac{i}{\epsilon}H_s\tau}\rho_s(t)\mathbf{d}\big)F(t,t-\tau) \\
+\big(\rho_s(t)e^{-\frac{i}{\epsilon}H_s\tau}\mathbf{d}^\dagger e^{\frac{i}{\epsilon}H_s\tau}\mathbf{d} - \mathbf{d}\rho_s(t)e^{-\frac{i}{\epsilon}H_s\tau}\mathbf{d}^\dagger e^{\frac{i}{\epsilon}H_s\tau}\big)G(t,t-\tau) + \text{h.c.}\Big).
\end{aligned} \tag{2.9}$$

This is the Redfield equation for the Anderson-Holstein model.

2.4 Derivation of the Lindblad equation

In this section, we aim at deriving the Lindblad equation for the Anderson-Holstein model. The derivation of Lindblad equations from a microscopic point of view has been studied for some cases, see, e.g., [22], though to the best of our knowledge, not for the Anderson-Holstein model. In this section, we will show that under the previous condition

$$\zeta \ll \epsilon, \quad \zeta \ll 1,$$

the Lindblad equation in the Schrödinger picture for the Anderson-Holstein model is

$$\frac{\mathrm{d}}{\mathrm{d}t}\rho_s(t) = -\frac{i}{\epsilon}\big[H_s + \zeta^2\mathscr{H}, \rho_s(t)\big] + \frac{\zeta^2}{\epsilon}\mathscr{D}(\rho_s(t)), \tag{2.10}$$

with a *Lindbladian corrected Hamiltonian*

$$\mathscr{H} = \sum_{\omega\in\mathbb{Z}} b_F(\omega)\mathbf{D}(\omega)\mathbf{D}^\dagger(\omega) - b_G(\omega)\mathbf{D}^\dagger(\omega)\mathbf{D}(\omega), \tag{2.11}$$

and dissipative operator

$$\begin{aligned}\mathscr{D}(\rho_s(t)) = \sum_{\omega\in\mathbb{Z}} & a_F(\omega)\Big(\mathbf{D}^\dagger(\omega)\rho_s(t)\mathbf{D}(\omega) - \frac{1}{2}\left\{\mathbf{D}(\omega)\mathbf{D}^\dagger(\omega), \rho_s(t)\right\}\Big) \\ & + a_G(\omega)\Big(\mathbf{D}(\omega)\rho_s(t)\mathbf{D}^\dagger(\omega) - \frac{1}{2}\left\{\mathbf{D}^\dagger(\omega)\mathbf{D}(\omega), \rho_s(t)\right\}\Big).\end{aligned} \tag{2.12}$$

The coefficients will be given in (2.20) below, and the operators $\mathbf{D}^{(\dagger)}(\omega)$ will be defined in (2.13) below. Note that the above equation has the Lindbladian form as in (1.1) and (1.2).

2.4.1 An alternative representation of the Redfield equation

In the Anderson-Holstein model, it is natural to consider the eigenfunctions of H_s, which form two energy ladders. The evolution of the system can be considered as quantum jumping between different energy levels. Thus, the annihilation and creation operators might be decomposed in terms of the numbers of energy levels that the system jumps. Such a decomposition has been used to derive the Lindblad equation in [22, pp. 125-131].

We shall use this technique to study both the Redfield equation and the Lindblad equation below for the Anderson-Holstein model.

Definition 2.1. For each $\omega \in \mathbb{Z}$, define an operator

$$\mathbf{D}(\omega) := \sum_{k'-k=\omega} \Pi_k^{(0)}\, \mathbf{d}\, \Pi_{k'}^{(1)}, \tag{2.13}$$

where $\Pi_k^{(m)}$ is the projection operator to the quantum state $|\phi_k^m\rangle \otimes |m\rangle \equiv |\phi_k^m, m\rangle$, $m \in \{0, 1\}$, $k, k' \in \mathbb{N}$.

Recall that $|\phi_k^m, m\rangle$ is the eigenfunction of the Hamiltonian H_s discussed at the end of Sec. 2.2. It is straightforward to know that the adjoint operator of $\mathbf{D}(\omega)$ is

$$\mathbf{D}^\dagger(\omega) = \sum_{k'-k=\omega} \Pi_{k'}^{(1)}\, \mathbf{d}^\dagger\, \Pi_k^{(0)}.$$

This definition was used in [22] for a slightly different coupling Hamiltonian, but it is also applicable here in the Anderson-Holstein model. It can also be checked that properties proposed in [22] still hold.

Lemma 2.2. (i) $\mathbf{D}(\omega)$ and $\mathbf{D}^\dagger(\omega)$ are the eigen-operators of H_s, namely,

$$[H_s, \mathbf{D}(\omega)] = -\epsilon\omega \mathbf{D}(\omega), \qquad \left[H_s, \mathbf{D}^\dagger(\omega)\right] = \epsilon\omega \mathbf{D}^\dagger(\omega).$$

(ii) In the interaction picture, $\mathbf{D}_I(\omega,t)$ and $\mathbf{D}_I^\dagger(\omega,t)$ have the following form

$$\mathbf{D}_I(\omega,t) \equiv e^{\frac{i}{\epsilon}H_s t}\mathbf{D}(\omega)e^{-\frac{i}{\epsilon}H_s t} = e^{-i\omega t}\mathbf{D}(\omega),$$

$$\mathbf{D}_I^\dagger(\omega,t) \equiv e^{\frac{i}{\epsilon}H_s t}\mathbf{D}^\dagger(\omega)e^{-\frac{i}{\epsilon}H_s t} = e^{i\omega t}\mathbf{D}^\dagger(\omega).$$

(iii) $\mathbf{d}$ and $\mathbf{d}^\dagger$ can be decomposed into $\mathbf{D}(\omega)$ and $\mathbf{D}^\dagger(\omega)$, respectively. More specifically,

$$\mathbf{d} = \sum_\omega \mathbf{D}(\omega), \qquad \mathbf{d}^\dagger = \sum_\omega \mathbf{D}^\dagger(\omega). \tag{2.14}$$

(iv) Then we can decompose the coupling Hamiltonian H_c as

$$H_c = \sum_\omega \big(\mathbf{D}(\omega)\otimes\mathbf{C}^\dagger + \mathbf{D}^\dagger(\omega)\otimes\mathbf{C}\big)\,,$$

and in the interaction picture

$$H_{c,I}(t) = \sum_\omega \Big(e^{-i\omega t}\mathbf{D}(\omega)\otimes\mathbf{C}_I^\dagger(t) + e^{i\omega t}\mathbf{D}^\dagger(\omega)\otimes\mathbf{C}_I(t)\Big). \tag{2.15}$$

These results directly follow from the definition of $\mathbf{D}(\omega)$. (2.14) is essential in decomposing $\mathbf{d}(\mathbf{d}^\dagger)$ in terms of the level of jumps. The reason that ω is the level of jumping can be observed from the definition that $\mathbf{D}(\omega)$ maps the quantum state $|\phi_{k'}^1, 1\rangle$ to the quantum state $|\phi_k^0, 0\rangle$, where $k = k' - \omega$. To prove the decomposition of $\mathbf{d}$ in terms of $\mathbf{D}(\omega)$, we use the completion relation $\sum_k(\Pi_k^{(0)} + \Pi_k^{(1)}) = \mathrm{Id}$,

$$
\begin{aligned}
\mathbf{d} &= \sum_{k,k'} \left(\Pi_k^{(0)} + \Pi_k^{(1)}\right) \mathbf{d} \left(\Pi_{k'}^{(0)} + \Pi_{k'}^{(1)}\right) = \sum_{k,k'} \Pi_k^{(0)}\, \mathbf{d}\, \Pi_{k'}^{(1)} \\
&= \sum_{\omega} \sum_{k'-k=\omega} \Pi_k^{(0)}\, \mathbf{d}\, \Pi_{k'}^{(1)} = \sum_{\omega} \mathbf{D}(\omega).
\end{aligned}
$$

In the second step, we have used $\mathbf{d}|\phi_{k'}^0, 0\rangle = \mathbf{d}^\dagger|\phi_k^1, 1\rangle = 0$ for any $k, k' \in \mathbb{N}$. It follows after taking the Hermitian conjugate that $\langle\phi_k^1, 1|\mathbf{d} = 0$; that is why there is only one term $\Pi_k^{(0)}\, \mathbf{d}\, \Pi_{k'}^{(1)}$ left. In the third step, re-ordering the double summation is employed to firstly sum over all differences of levels, namely, ω and then sum over all possible combination of $k', k \in \mathbb{N}$ such that $k' - k = \omega$, where the latter sum gives $\mathbf{D}(\omega)$.

Starting from (2.7), replacing $H_{c,I}$ by (2.15), and opening the double commutators, we arrive at an alternative representation of the Redfield equation

$$
\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\rho_{s,I}(t) = -\frac{\zeta^2}{\epsilon} \sum_{\omega,\,\omega'} & e^{-i(\omega'-\omega)t} \left(\mathbf{D}(\omega')\mathbf{D}^\dagger(\omega)\rho_{s,I}(t) - \mathbf{D}^\dagger(\omega)\rho_{s,I}(t)\mathbf{D}(\omega')\right) F(\omega) \\
& + e^{-i(\omega'-\omega)t} \left(\rho_{s,I}(t)\mathbf{D}^\dagger(\omega)\mathbf{D}(\omega') - \mathbf{D}(\omega')\rho_{s,I}(t)\mathbf{D}^\dagger(\omega)\right) G(\omega) + \text{h.c.},
\end{aligned}
\tag{2.16}
$$

where

$$
\begin{aligned}
F(\omega) &\equiv \frac{1}{\epsilon}\int_0^\infty d\tau\, e^{-i\omega\tau} F(t, t-\tau) = \sum_k V_k^2\, \mathsf{f}_{\mathrm{FD}}(E_k) \int_0^\infty d\tau\, e^{i(E_k - \epsilon\omega)\tau}, \\
G(\omega) &\equiv \frac{1}{\epsilon}\int_0^\infty d\tau\, e^{-i\omega\tau} G(t, t-\tau) = \sum_k V_k^2 (1 - \mathsf{f}_{\mathrm{FD}}(E_k)) \int_0^\infty d\tau\, e^{i(E_k - \epsilon\omega)\tau}.
\end{aligned}
\tag{2.17}
$$

These two equations can be viewed as the Laplace transform of time correlation functions with the frequency parameter $i\omega$.

2.4.2 Secular approximation

From (2.16) and (2.17), we observe that $\frac{\mathrm{d}}{\mathrm{d}t}\rho_{s,I}(t) = \mathcal{O}\left(\frac{\zeta^2}{\epsilon}\right)$. This motivates the choice of the relaxation time as $\tau_R \equiv \frac{\epsilon}{\zeta^2}$. Recall that we have assumed $\zeta \ll \epsilon$ and $\zeta \ll 1$, then $\zeta^2 \ll \epsilon$, or equivalently $\tau_R = \frac{\epsilon}{\zeta^2} \gg 1$. Integrating $\rho_{s,I}(t)$ over the time period $[t, t + r\tau_R]$ for $r = \mathcal{O}(1)$, we obtain

$$\rho_{s,I}(t + r\tau_R) - \rho_{s,I}(t) = -\frac{\zeta^2}{\epsilon} \sum_{\omega,\ \omega'} \int_t^{t+r\tau_R} \mathrm{d}s\ e^{-i(\omega'-\omega)s} Op(s) + \text{h.c.}$$

We use the short-hand $Op(s)$, for simplicity, to denote the long term involving operators in (2.16). Then, by changing the variable $s = t + \tau_R s'$,

$$\rho_{s,I}(t + r\tau_R) - \rho_{s,I}(t) = -\sum_{\omega,\ \omega'} e^{-i(\omega'-\omega)t} \int_0^r \mathrm{d}s'\ e^{-i(\omega'-\omega)\tau_R s'} Op(t + \tau_R s') + \text{h.c.}$$

By the Riemann-Lebesgue lemma, if $\omega' \neq \omega$, $(\omega' - \omega)\tau_R = \mathcal{O}(\tau_R) \gg 1$,

$$\int_0^r \mathrm{d}s'\ e^{-i(\omega'-\omega)\tau_R s'} Op(t + \tau_R s') \approx 0.$$

Then terms involving $\omega' - \omega \neq 0$ have negligible contribution. Hence,

$$\rho_{s,I}(t + r\tau_R) - \rho_{s,I}(t) \approx -\sum_{\omega=\omega'} e^{-i(\omega'-\omega)t} \int_0^r \mathrm{d}s'\ e^{-i(\omega'-\omega)\tau_R s'} Op(t + \tau_R s') + \text{h.c.},$$

that is,

$$\rho_{s,I}(t+r\tau_R)-\rho_{s,I}(t)\approx-\frac{\zeta^2}{\epsilon}\sum_{\omega=\omega'}\int_t^{t+r\tau_R}ds\, e^{-i(\omega'-\omega)s}Op(s)+\text{h.c.}$$

This is known as the *secular approximation* [22], which we have justified here in the sense of coarse-grained approximation over the relaxation time. Divide both sides by $r\tau_R$ and then take the limit $r\to 0$,

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\rho_{s,I}(t)\approx-\frac{\zeta^2}{\epsilon}\sum_{\omega=\omega'}& e^{-i(\omega'-\omega)t}\big(\mathbf{D}(\omega')\mathbf{D}^\dagger(\omega)\rho_{s,I}(t)-\mathbf{D}^\dagger(\omega)\rho_{s,I}(t)\mathbf{D}(\omega')\big)F(\omega)\\ &+e^{-i(\omega'-\omega)t}\big(\rho_{s,I}(t)\mathbf{D}^\dagger(\omega)\mathbf{D}(\omega')-\mathbf{D}(\omega')\rho_{s,I}(t)\mathbf{D}^\dagger(\omega)\big)G(\omega)+\text{h.c.}\end{aligned}$$

Then, we arrive at the secular approximation, which is the basis for deriving the Lindblad equation,

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\rho_{s,I}(t)=-\frac{\zeta^2}{\epsilon}\sum_{\omega}&\left(\mathbf{D}(\omega)\mathbf{D}^\dagger(\omega)\rho_{s,I}(t)-\mathbf{D}^\dagger(\omega)\rho_{s,I}(t)\mathbf{D}(\omega)\right)F(\omega)+\\ &+\left(\rho_{s,I}(t)\mathbf{D}^\dagger(\omega)\mathbf{D}(\omega)-\mathbf{D}(\omega)\rho_{s,I}(t)\mathbf{D}^\dagger(\omega)\right)G(\omega)+\text{h.c.}\end{aligned}\tag{2.18}$$

Remark. By checking the previous argument, in fact, the secular approximation is valid when $\frac{\zeta^2}{\epsilon}\ll 1$, that is, $\zeta^2\ll\epsilon$, which appears to be a weaker condition than $\zeta\ll\epsilon$ used for the Born-Markov approximation.

2.4.3 Lindblad equation in the interaction picture

To write (2.18) in a Lindbladian form, we need to decompose coefficients $F(\omega)$ and $G(\omega)$ into their real and imaginary parts. Let $a_F(\omega) := F(\omega) + F(\omega)^*$ and $b_F(\omega) := \frac{F(\omega)-F(\omega)^*}{2i}$, then $F(\omega) \equiv \frac{a_F(\omega)}{2} + ib_F(\omega)$. Similarly, we can decompose $G(\omega) = \frac{a_G(\omega)}{2} + ib_G(\omega)$. With these notations, (2.18) becomes the *Lindblad equation*

$$\frac{\mathrm{d}}{\mathrm{d}t}\rho_{s,I}(t) = -\frac{i\zeta^2}{\epsilon}\left[\mathscr{H}_I, \rho_{s,I}(t)\right] + \frac{\zeta^2}{\epsilon}\mathscr{D}(\rho_{s,I}(t)), \tag{2.19}$$

where the *Lindbladian correction Hamiltonian* $\mathscr{H}_I$ has the form

$$\mathscr{H}_I = \sum_{\omega} b_F(\omega)\mathbf{D}(\omega)\mathbf{D}^\dagger(\omega) - b_G(\omega)\mathbf{D}^\dagger(\omega)\mathbf{D}(\omega),$$

and the dissipative operator $\mathscr{D}$ has the form

$$\begin{aligned}\mathscr{D}(\rho_{s,I}(t)) = \sum_{\omega} a_F(\omega)\Big(&\mathbf{D}^\dagger(\omega)\rho_{s,I}(t)\mathbf{D}(\omega) - \frac{1}{2}\left\{\mathbf{D}(\omega)\mathbf{D}^\dagger(\omega), \rho_{s,I}(t)\right\}\Big) \\ &+ a_G(\omega)\Big(\mathbf{D}(\omega)\rho_{s,I}(t)\mathbf{D}^\dagger(\omega) - \frac{1}{2}\left\{\mathbf{D}^\dagger(\omega)\mathbf{D}(\omega), \rho_{s,I}(t)\right\}\Big).\end{aligned}$$

2.4.4 Lindblad equation in the Schrödinger picture

Transforming back into the Schrödinger picture by $\rho_s(t) = e^{-\frac{i}{\epsilon}H_s t}\rho_{s,I}(t)e^{\frac{i}{\epsilon}H_s t}$, we obtain the *Lindblad equation* in (2.10), (2.11), and (2.12). Note that $\mathscr{H} = e^{-\frac{i}{\epsilon}H_s t}\mathscr{H}_I e^{\frac{i}{\epsilon}H_s t}$ is invariant in different pictures because $\mathbf{D}(\omega)$ and $\mathbf{D}^\dagger(\omega)$ both appear in the same term, and thus the factors $e^{\pm i\omega t}$ will always cancel during the picture transformation. This cancellation also applies to the dissipative operator $\mathscr{D}$.

To understand the Lindbladian corrected Hamiltonian, we note that after some simple computation, it could be shown that

$$\begin{aligned}\mathscr{H} =& \sum_k \Big(\sum_\omega b_F(\omega) \left| \langle \phi_k^0 | \phi_{k+\omega}^1 \rangle \right|^2 \Big) |\phi_k^0, 0\rangle\langle\phi_k^0, 0| \\ &- \sum_k \Big(\sum_\omega b_G(\omega) \left| \langle \phi_k^1 | \phi_{k-\omega}^0 \rangle \right|^2 \Big) |\phi_k^1, 1\rangle\langle\phi_k^1, 1|.\end{aligned}$$

Hence, for the new Hamiltonian, i.e., $H_s + \zeta^2\mathscr{H}$, the set of eigenstates are the same, $|\phi_k^0, 0\rangle$ and $|\phi_k^1, 1\rangle$ for $k \in \mathbb{N}$. In the new Hamiltonian, the energy eigenvalues, however, are perturbed by an order $\mathcal{O}(\zeta^2)$.

Physically, this perturbation of energy comes from the interaction between two quantum states $|0\rangle$ and $|1\rangle$ through the coupling with the bath. More specifically, since the quantum eigenstate $|\phi_k^0, 0\rangle$ interacts with quantum eigenstates $|\phi_{k+\omega}^1, 1\rangle$ (for all possible $\omega \in \mathbb{Z}$), their interaction contributes to the change of energy. The interaction is realized through the bath, hence the perturbed Hamiltonian should be weighted by time correlation functions, namely, terms $b_{F(G)}(\omega)$ and is also proportional to ζ^2.

The effect of Lindblad operators will be further investigated in Sec. 2.6 below in the context of perturbation theory, and it will be shown that Lindblad operators characterize the hopping between quantum states $|0\rangle$ and $|1\rangle$. More specifically, the hopping rate out of the eigenstate $|\phi_k^0, 0\rangle$ is $\frac{\zeta^2}{\epsilon}\sum_\omega a_F(\omega)\big|\left\langle\phi_k^0|\phi_{k+\omega}^1\right\rangle\big|^2$. It is worth pointing out that this expression is quite similar to the perturbed energy eigenvalue as above $\zeta^2\sum_\omega b_F(\omega)\left|\langle\phi_k^0|\phi_{k+\omega}^1\rangle\right|^2$. For the Laplace transform of time correlation functions, namely, $F(\omega)$ and $G(\omega)$ in (2.17), the real part contributes to (weak) hopping, and the imaginary part contributes to the (weak) perturbation to energy eigenvalues.

2.4.5 Coefficients and wide band approximation

It remains to determine the coefficients $a_{F,G}(\omega)$ and $b_{F,G}(\omega)$ in the Lindblad equation. Using oscillatory integral

$$\int_0^\infty \mathrm{d}\tau\, e^{i\omega\tau} = \pi\delta(\omega) + i\,\mathrm{p.\,v.}\left(\frac{1}{\omega}\right),$$

we could obtain that

$$F(\omega) = \sum_k V_k^2 \mathsf{f}_{\mathrm{FD}}(E_k)\Big(\pi\delta(E_k - \epsilon\omega) + i\,\mathrm{p.\,v.}\big(\frac{1}{E_k - \epsilon\omega}\big)\Big),$$

$$G(\omega) = \sum_k V_k^2 (1 - \mathsf{f}_{\mathrm{FD}}(E_k))\Big(\pi\delta(E_k - \epsilon\omega) + i\,\mathrm{p.\,v.}\big(\frac{1}{E_k - \epsilon\omega}\big)\Big).$$

Thus, by matching the definition of $a_{F,G}(\omega)$ and $b_{F,G}(\omega)$,

$$
\begin{aligned}
a_F(\omega) &= 2\pi \sum_k V_k^2 \, \mathsf{f}_{\text{FD}}(E_k) \, \delta(E_k - \epsilon\omega), \\
b_F(\omega) &= \sum_k V_k^2 \, \mathsf{f}_{\text{FD}}(E_k) \, \text{p. v.}\big(\frac{1}{E_k - \epsilon\omega}\big), \\
a_G(\omega) &= 2\pi \sum_k V_k^2 (1 - \mathsf{f}_{\text{FD}}(E_k)) \delta(E_k - \epsilon\omega), \\
b_G(\omega) &= \sum_k V_k^2 (1 - \mathsf{f}_{\text{FD}}(E_k)) \, \text{p. v.}\big(\frac{1}{E_k - \epsilon\omega}\big).
\end{aligned}
\tag{2.20}
$$

Notice that $a_{F,G}(\omega)$ and $b_{F,G}(\omega)$ are (generalized) functions with respect to ω. Even though we only need values at $\omega \in \mathbb{Z}$, these functions are indeed well-defined on $\mathbb{R}$.

In the Anderson-Holstein model, we have assumed that the electron bath is infinitely large, so the continuum approximation appears to be a possible approach to simplify the above coefficients. Assume that $V_k^2 \equiv V^2(E_k)$ is a continuous function of E_k. In the discrete case, suppose the total number of states in the bath is N, then $V^2(E)$ should be inversely proportional to N, to make the overall interaction strength between the system and bath remain at $\mathcal{O}(1)$: let $V^2(E) = \frac{\check{V}^2(E)}{N}$. Let D be the energy band width of the electron bath and $\nu(E)$ be the density of states at the energy level E. In the wide band approximation, to simplify the last equation (2.20), it is assumed that the contribution to the interaction strength from different energy levels of the electron bath is approximately the same; explicitly, assume

$$2\pi\check{V}^2(E)\nu(E) = \Gamma, \qquad \forall E \in [-D, D],$$

where Γ is a constant [24]. Then for a test function $g(\omega)$,

$$\begin{aligned}
&\int_{-\infty}^{\infty} \mathrm{d}\omega\ a_F(\omega)g(\omega) \\
&= 2\pi \sum_k \frac{V_k^2 \mathsf{f}_{\mathrm{FD}}(E_k)}{\epsilon} g\big(E_k/\epsilon\big) \\
&\approx 2\pi \int_{-D}^{D} dE\ \check{V}^2(E)\nu(E)\mathsf{f}_{\mathrm{FD}}(E)\frac{1}{\epsilon} g\big(E/\epsilon\big) \qquad\qquad \text{(continuum approximation)} \\
&= \Gamma \int_{-D/\epsilon}^{D/\epsilon} \mathrm{d}\omega\ \mathsf{f}_{\mathrm{FD}}(\epsilon\omega)g(\omega) \\
&= \int_{-\infty}^{\infty} \mathrm{d}\omega\ \Gamma\chi_{[-D/\epsilon, D/\epsilon]}(\omega)\mathsf{f}_{\mathrm{FD}}(\epsilon\omega)g(\omega).
\end{aligned}$$

Therefore, in the continuum limit,

$$a_F(\omega) = \Gamma\chi_{[-D/\epsilon, D/\epsilon]}(\omega)\mathsf{f}_{\mathrm{FD}}(\epsilon\omega).$$

There are two ways to get rid of the characteristic function: the first way is to assume that $D = \infty$, mentioned in [116]; the second way is to consider $\epsilon \downarrow 0$. In either way, we end up with the approximation

$$a_F(\omega) \approx \Gamma\mathsf{f}_{\mathrm{FD}}(\epsilon\omega).$$

These two conditions are consistent with where the term $a_F(\omega)$ comes. $a_F(\omega)$ is part of the interaction strength, which involves both electronic bath and open quantum system; when the electronic bath is infinitely wide or the open quantum system falls into the semiclassical region, the (generalized) function $a_F(\omega)$ can be approximated in this way. Similarly, we can approximate

$$a_G(\omega) = \Gamma\chi_{[-D/\epsilon, D/\epsilon]}(\omega)(1 - \mathsf{f}_{\text{FD}}(\epsilon\omega)) \approx \Gamma(1 - \mathsf{f}_{\text{FD}}(\epsilon\omega)).$$

As for $b_F(\omega)$ and $b_G(\omega)$, we have not found easy expressions for them.

2.5 Semiclassical limit

The semiclassical limit of the Redfield equation has been proposed and studied in [57, 62]. The system of phase space functions obtained in the semiclassical limit of the Redfield equation by applying the Wigner transform, is called the *classical master equation* (CME). In the first part, we attempt to justify the formal derivation of [57] in a more mathematical way. In the second part, more importantly, we try to study the phase space counterparts of the Lindblad equation by applying the Wigner transform. We call the system of phase space functions the *Lindbladian classical master equation* (LCME).

2.5.1 Wigner transform and phase space functions

Recall that we identify $\epsilon \equiv \hbar\omega_s/\mathsf{E}$ as the semiclassical parameter. Also recall that the Wigner transform of an operator A on $L^2(\mathbb{R})$ is defined by [140, 144]

$$(A_{\mathcal{W}})(x,p) := \int_{\mathbb{R}} A\Big(x+\frac{y}{2}, x-\frac{y}{2}\Big)\exp\Big(-\frac{ipy}{\epsilon}\Big)\ \mathrm{d}y.$$

The subscript $\mathcal{W}$ indicates the Wigner transform. If the operator $A=\rho$ is a density operator, it could be easily shown that $\int \mathrm{d}x\,\mathrm{d}p\,\frac{1}{2\pi\epsilon}\rho_{\mathcal{W}}(x,p)=1$. Hence, if we apply the Wigner transform to density operators ρ, the factor $\frac{1}{2\pi\epsilon}$ is needed to ensure unit mass.

For a single level quantum system, the phase space function for a quantum master equation is clear (see, e.g., [140]). As for multi-level open quantum systems, the definition of phase space functions is not very straightforward, and we need to clarify this concept below. For a general two-level open quantum system, the reduced density matrix can be written in the matrix form as

$$\rho_s(t) = \begin{pmatrix} \rho_{0,0}(t) & \rho_{0,1}(t) \\ \rho_{1,0}(t) & \rho_{1,1}(t) \end{pmatrix}.$$

For a general quantum master equation, it is expected that $\rho_{0,1}(t)$ and $\rho_{1,0}(t)$ do not vanish. However, it could be verified that if at the initial time t_0, $\rho_s(t)$ only has diagonal terms, i.e., $\rho_{0,1}(t_0)=\rho_{1,0}(t_0)=0$, then $\frac{\mathrm{d}}{\mathrm{d}t}\rho_{0,1}(t_0)=\frac{\mathrm{d}}{\mathrm{d}t}\rho_{1,0}(t_0)=0$ for both Redfield equation and

Lindblad equations. Thus, $\rho_{0,1}(t) = \rho_{1,0}(t) = 0$ for all t.

Therefore, in the below, we shall only consider diagonal elements of $\rho_s(t)$, i.e., assume $\rho_s(t) = \rho_0(t)|0\rangle\langle 0| + \rho_1(t)|1\rangle\langle 1|$ for all t. In the matrix form,

$$\rho_s(t) = \begin{pmatrix} \rho_0(t) & 0 \\ 0 & \rho_1(t) \end{pmatrix}.$$

The phase space function obtained by applying the Wigner transform for $\rho_m(t)$ is denoted by $\varrho_m(x,p,t)$, for $m \in \{0,1\}$, namely,

$$\varrho_m(x,p,t) = \frac{1}{2\pi\epsilon}(\rho_m(t))_{\mathcal{W}}.$$

Hence, it is not difficult to show that $\int \mathrm{d}x\,\mathrm{d}p\ \varrho_0(x,p,t) + \varrho_1(x,p,t) = 1$.

We need the following lemma below.

Lemma 2.3 (Semiclassical expansion, [144, pp. 66 - 68])**.** *We have*

(i) $(AB)_{\mathcal{W}} = A_{\mathcal{W}}B_{\mathcal{W}} + \frac{i\epsilon}{2}\{A_{\mathcal{W}}, B_{\mathcal{W}}\}_{Pois} + \mathcal{O}(\epsilon^2)$*, where A and B are two operators and $\{\cdot,\cdot\}_{Pois}$ is the Poisson bracket, defined by*[1]

$$\{f(x,p), g(x,p)\}_{Pois} = \partial_x f(x,p)\partial_p g(x,p) - \partial_p f(x,p)\partial_x g(x,p).$$

Moreover, $(AB)_{\mathcal{W}}$ is linear with respect to both $A_{\mathcal{W}}$ and $B_{\mathcal{W}}$.

[1] We choose to follow the usual convention here, which differs by a negative sign compared to [144].

(ii) In particular, we have $[A,B]_{\mathcal{W}} = i\epsilon\,\{A_{\mathcal{W}}, B_{\mathcal{W}}\}_{Pois} + \mathcal{O}(\epsilon^3)$.

In the lemma below, we provide the Wigner transform of the square of position and momentum operators. This lemma can be verified by direct calculation.

Lemma 2.4 (Quadratic case)**.** *If* $A = \mathbf{x}^2$*, then* $A_{\mathcal{W}} = x^2$*; if* $A = -\epsilon^2\nabla_x^2$*, then* $A_{\mathcal{W}} = p^2$*.*

2.5.2 Classical master equation

As discussed in the last subsection, we shall only consider the reduced density matrix of the form $\rho_s(t) = \rho_0(t)|0\rangle\langle 0| + \rho_1(t)|1\rangle\langle 1|$. From (2.9), we could derive that the time evolution equations for $\rho_0(t)$ and $\rho_1(t)$ are

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\rho_0(t) =& -\frac{i}{\epsilon}[H_0,\rho_0(t)] - \frac{\zeta^2}{\epsilon}\sum_k V_k^2 \int_0^\infty \mathrm{d}\tau\; e^{-iH_1\tau}e^{iH_0\tau}\rho_0(t)e^{iE_k\tau}\mathsf{f}_{\mathrm{FD}}(E_k) \\ &- \rho_1(t)e^{-iH_1\tau}e^{iH_0\tau}e^{iE_k\tau}(1-\mathsf{f}_{\mathrm{FD}}(E_k)) + \text{h.c.}, \\ \frac{\mathrm{d}}{\mathrm{d}t}\rho_1(t) =& -\frac{i}{\epsilon}[H_1,\rho_1(t)] - \frac{\zeta^2}{\epsilon}\sum_k V_k^2 \int_0^\infty \mathrm{d}\tau\; \rho_1(t)e^{-iH_1\tau}e^{iH_0\tau}e^{iE_k\tau}(1-\mathsf{f}_{\mathrm{FD}}(E_k)) \\ &- e^{-iH_1\tau}e^{iH_0\tau}\rho_0(t)e^{iE_k\tau}\mathsf{f}_{\mathrm{FD}}(E_k) + \text{h.c.},\end{aligned}$$

which agree with [59, Eqs. (14) and (15)].

Using Lem. 2.3, we can calculate the equations for the corresponding Wigner transform

$$\begin{aligned}\partial_t\varrho_0(t) = \{\hbar_0, \varrho_0(t)\}_{Pois} - \frac{\zeta^2}{\epsilon}\sum_k V_k^2\int_0^\infty \mathrm{d}\tau\Big(&(e^{-iH_1\tau})_{\mathcal{W}}(e^{iH_0\tau})_{\mathcal{W}}\varrho_0(t)e^{iE_k\tau}\mathsf{f}_{\mathrm{FD}}(E_k) \\ &- \varrho_1(t)(e^{-iH_1\tau})_{\mathcal{W}}(e^{iH_0\tau})_{\mathcal{W}}e^{iE_k\tau}(1-\mathsf{f}_{\mathrm{FD}}(E_k)) + \text{c.c.} + \mathcal{O}(\zeta^2)\Big),\end{aligned}$$

$$\partial_t \varrho_1(t) = \{\hbar_1, \varrho_1(t)\}_{Pois} - \frac{\zeta^2}{\epsilon}\sum_k V_k^2 \int_0^\infty \mathrm{d}\tau \Big(\varrho_1(t)(e^{-iH_1\tau})_{\mathcal{W}}(e^{iH_0\tau})_{\mathcal{W}} e^{iE_k\tau}(1-\mathsf{f}_{\mathrm{FD}}(E_k))$$

$$- (e^{-iH_1\tau})_{\mathcal{W}}(e^{iH_0\tau})_{\mathcal{W}}\varrho_0(t) e^{iE_k\tau}\mathsf{f}_{\mathrm{FD}}(E_k) + \text{c.c.} + \mathcal{O}(\zeta^2)\Big),$$

where $\hbar_0$ and $\hbar_1$ are Wigner transforms of H_0 and H_1, respectively. For clarity, in the last equation, the coordinates (x, p) are suppressed in phase space functions and in $\hbar_0$, $\hbar_1$ as well. After dropping higher order terms $\mathcal{O}(\zeta^2)$, we have

$$\partial_t \varrho_0 = \{\hbar_0, \varrho_0\}_{Pois} - \gamma_{0\to 1}\varrho_0 + \gamma_{1\to 0}\varrho_1,$$

$$\partial_t \varrho_1 = \{\hbar_1, \varrho_1\}_{Pois} + \gamma_{0\to 1}\varrho_0 - \gamma_{1\to 0}\varrho_1,$$

where the hopping rates

$$\gamma_{0\to 1} = \frac{\zeta^2}{\epsilon}\sum_k V_k^2 \int_0^\infty \mathrm{d}\tau\ (e^{-iH_1\tau})_{\mathcal{W}}(e^{iH_0\tau})_{\mathcal{W}} e^{iE_k\tau}\mathsf{f}_{\mathrm{FD}}(E_k) + \text{c.c.},$$

$$\gamma_{1\to 0} = \frac{\zeta^2}{\epsilon}\sum_k V_k^2 \int_0^\infty \mathrm{d}\tau\ (e^{-iH_1\tau})_{\mathcal{W}}(e^{iH_0\tau})_{\mathcal{W}} e^{iE_k\tau}(1-\mathsf{f}_{\mathrm{FD}}(E_k)) + \text{c.c.}$$

Notice that hopping rates are phase space functions of coordinates x and p. They describe how fast the jumping between states $|0\rangle$ and $|1\rangle$ occurs, depending on the phase space coordinates (x, p).

Remark. Compared with the result in [57], the rates $\gamma_{0\to 1}$ and $\gamma_{1\to 0}$ that we have above are considerably more complicated. Here we provide a heuristic simplification of the expression, though we do not know how to justify the argument on a more rigorous level.

Eq. (60) in [48] shows that if H is a Hamiltonian for harmonic oscillators, then

$$(e^{iHt})_{\mathcal{W}} = \cos(t/2)^{-1} e^{2i\hbar \tan(t/2)},$$

where $\hbar = (H)_{\mathcal{W}}$. Since H_0 and H_1 are Hamiltonians for harmonic oscillators, by the last equation,

$$\begin{aligned}(e^{-iH_1\tau})_{\mathcal{W}}(e^{iH_0\tau})_{\mathcal{W}} &= \cos(\tau/2)^{-2} e^{-2i(\hbar_1-\hbar_0)\tan(\tau/2)} \\ &= \cos(\tau/2)^{-2} e^{-2i\mathcal{U}(x)\tan(\tau/2)} \\ &\overset{\tau\downarrow 0}{\approx} e^{-i\mathcal{U}(x)\tau}.\end{aligned}$$

In the last step, we use $\cos(\tau/2) \to 1$ and $\tan(\tau/2) \to \tau/2$ as $\tau \to 0$. Then, hopping rates can be written as

$$\gamma_{0\to 1} = \frac{\zeta^2}{\epsilon} \sum_k V_k^2 \int_0^\infty \mathrm{d}\tau \ \cos(\tau/2)^{-2} e^{-2i\mathcal{U}(x)\tan(\tau/2)} e^{iE_k\tau} \mathsf{f}_{\text{FD}}(E_k) + \text{c.c.}$$

Heuristically, if we approximate $(e^{-iH_1\tau})_{\mathcal{W}}(e^{iH_0\tau})_{\mathcal{W}}$ by $e^{-i\mathcal{U}(x)\tau}$, i.e., assume that the integral is mostly contributed from $\tau \approx 0$, then

$$\begin{aligned}\gamma_{0\to 1} &= \frac{\zeta^2}{\epsilon} 2\pi \sum_k V_k^2 \mathsf{f}_{\text{FD}}(E_k)\delta(E_k - \mathcal{U}(x)) \\ &= \frac{\zeta^2}{\epsilon}\Gamma \ \mathsf{f}_{\text{FD}}(\mathcal{U}(x))\chi_{[\mathcal{U}^{-1}(-D),\mathcal{U}^{-1}(D)]}(x) \\ &\approx \frac{\zeta^2}{\epsilon}\Gamma \ \mathsf{f}_{\text{FD}}(\mathcal{U}(x)), \qquad \text{(let } D=\infty, \text{ i.e., wide band approximation)}\end{aligned}$$

which becomes the result in [57]. $\mathcal{U}^{-1}(x)$ is the inverse function of $\mathcal{U}(x)$; since $\mathcal{U}(x)$ is a linear function, $\mathcal{U}^{-1}(x)$ is well-defined. The second and the third step of the last equation use similar computation as in the wide band approximation in Sec. 2.4.5. This heuristic computation gives a nice simple expression for the rates, but it should be pointed out that we do not know how to justify the crucial approximation of $(e^{-iH_1\tau})_{\mathcal{W}}(e^{iH_0\tau})_{\mathcal{W}}$ by $e^{-i\mathcal{U}(x)\tau}$.

2.5.3 Lindbladian classical master equation

Recall that if we assume that $\rho_s(t_0)$ is diagonal, then the reduced density operator $\rho_s(t) = \rho_0(t)|0\rangle\langle 0| + \rho_1(t)|1\rangle\langle 1|$, without terms involving $|1\rangle\langle 0|$ nor $|0\rangle\langle 1|$. Then the Lindblad equation can be written more explicitly as

$$
\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\rho_0(t) = &-\frac{i}{\epsilon}\Big[H_0 + \zeta^2\sum_{\omega} b_F(\omega)\langle 0|\mathbf{D}(\omega)|1\rangle\langle 1|\mathbf{D}^\dagger(\omega)|0\rangle, \rho_0(t)\Big] \\
&+\frac{\zeta^2}{\epsilon}\sum_{\omega} a_G(\omega)\langle 0|\mathbf{D}(\omega)|1\rangle\rho_1(t)\langle 1|\mathbf{D}^\dagger(\omega)|0\rangle \\
&-\frac{\zeta^2}{\epsilon}\sum_{\omega}\frac{a_F(\omega)}{2}\begin{pmatrix}\langle 0|\mathbf{D}(\omega)|1\rangle\langle 1|\mathbf{D}^\dagger(\omega)|0\rangle\rho_0(t) \\ \\ +\rho_0(t)\langle 0|\mathbf{D}(\omega)|1\rangle\langle 1|\mathbf{D}^\dagger(\omega)|0\rangle\end{pmatrix},
\end{aligned}
$$

$$
\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\rho_1(t) = &- \frac{i}{\epsilon}\Big[H_1 - \zeta^2 \sum_\omega b_G(\omega)\langle 1|\mathbf{D}^\dagger(\omega)|0\rangle\langle 0|\mathbf{D}(\omega)|1\rangle, \rho_1(t)\Big] \\
&+ \frac{\zeta^2}{\epsilon}\sum_\omega a_F(\omega)\langle 1|\mathbf{D}^\dagger(\omega)|0\rangle\rho_0(t)\langle 0|\mathbf{D}(\omega)|1\rangle \\
&- \frac{\zeta^2}{\epsilon}\sum_\omega \frac{a_G(\omega)}{2}\begin{pmatrix}\langle 1|\mathbf{D}^\dagger(\omega)|0\rangle\langle 0|\mathbf{D}(\omega)|1\rangle\rho_1(t) \\ \\ + \rho_1(t)\langle 1|\mathbf{D}^\dagger(\omega)|0\rangle\langle 0|\mathbf{D}(\omega)|1\rangle\end{pmatrix}.
\end{aligned}
$$

The system of time-evolution equations obtained by applying the Wigner transform to the Lindblad equation is given by, after some straightforward calculations,

$$
\begin{aligned}
\partial_t\varrho_0(t) = &\left\{\hbar_0 + \zeta^2\mathcal{H}_0, \varrho_0(t)\right\} \\
&+ \frac{\zeta^2}{\epsilon}\sum_\omega a_G(\omega)\left(\langle 0|\mathbf{D}(\omega)\mathbf{D}^\dagger(\omega)|0\rangle\right)_{\mathcal{W}}\varrho_1(t) \\
&- \frac{\zeta^2}{\epsilon}\sum_\omega a_F(\omega)\left(\langle 0|\mathbf{D}(\omega)\mathbf{D}^\dagger(\omega)|0\rangle\right)_{\mathcal{W}}\varrho_0(t) + \mathcal{O}(\zeta^2, \zeta^2\epsilon^2),
\end{aligned}
\tag{2.21}
$$

$$
\begin{aligned}
\partial_t\varrho_1(t) = &\left\{\hbar_1 - \zeta^2\mathcal{H}_1, \varrho_1(t)\right\} \\
&+ \frac{\zeta^2}{\epsilon}\sum_\omega a_F(\omega)\left(\langle 1|\mathbf{D}^\dagger(\omega)\mathbf{D}(\omega)|1\rangle\right)_{\mathcal{W}}\varrho_0(t) \\
&- \frac{\zeta^2}{\epsilon}\sum_\omega a_G(\omega)\left(\langle 1|\mathbf{D}^\dagger(\omega)\mathbf{D}(\omega)|1\rangle\right)_{\mathcal{W}}\varrho_1(t) + \mathcal{O}(\zeta^2, \zeta^2\epsilon^2),
\end{aligned}
\tag{2.22}
$$

where

$$
\begin{aligned}
\mathcal{H}_0 &= \sum_\omega b_F(\omega)\left(\langle 0|\mathbf{D}(\omega)\mathbf{D}^\dagger(\omega)|0\rangle\right)_{\mathcal{W}}, \\
\mathcal{H}_1 &= \sum_\omega b_G(\omega)\left(\langle 1|\mathbf{D}^\dagger(\omega)\mathbf{D}(\omega)|1\rangle\right)_{\mathcal{W}}.
\end{aligned}
$$

The error terms with order $\mathcal{O}(\zeta^2)$ come from the Wigner transform for hopping terms, and the error terms with order $\mathcal{O}(\zeta^2\epsilon^2)$ come from the Wigner transform of commutators.

2.6 Comparison of the Redfield equation and the Lindblad equation from perturbation theory

In this section, we intend to use perturbation theory to understand the similarity and the difference of the Redfield equation (2.16) and the Lindblad equation (2.10) by considering the hopping between different quantum states. The mathematical tool for perturbation theory has been widely studied, and the most famous formula from that is perhaps the Fermi Golden Rule.

The main finding is that if the quantum system is prepared at a pure state $|\phi_i^0, 0\rangle$, then its hopping rates to other quantum states $|\phi_f^0, 0\rangle$ and $|\phi_f^1, 1\rangle$ are the same up to the first order $\mathcal{O}(\frac{\zeta^2}{\epsilon})$, for both Redfield and Lindblad equations. By linearity of Redfield and Lindblad equations, this conclusion is also true if initially ρ_m $(m = 0, 1)$ are diagonal.

2.6.1 Perturbation of Redfield and Lindblad equations

As mentioned above, the Lindblad equation can be derived by using the secular approximation to the Redfield equation, under the condition that $\zeta^2 \ll \epsilon$. Under this condition, the hopping term in the Redfield equation (2.16) could be considered as a small perturbation, that is, $\frac{\zeta^2}{\epsilon}$ is regarded as the perturbation parameter below.

In the Schrödinger picture, the Redfield equation has the form, by (2.16),

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\rho_s(t) = -\frac{i}{\epsilon}[H_s, \rho_s(t)] - \frac{\zeta^2}{\epsilon}\sum_{\omega,\omega'} &\left(\mathbf{D}(\omega')\mathbf{D}^\dagger(\omega)\rho_s(t) - \mathbf{D}^\dagger(\omega)\rho_s(t)\mathbf{D}(\omega')\right) F(\omega) \\ &+ \left(\rho_s(t)\mathbf{D}^\dagger(\omega)\mathbf{D}(\omega') - \mathbf{D}(\omega')\rho_s(t)\mathbf{D}^\dagger(\omega)\right) G(\omega) \\ &+\text{h.c.}\end{aligned}$$

After rewriting it in terms of $a_{F(G)}(\omega)$ and $b_{F(G)}(\omega)$, it becomes

$$\frac{\mathrm{d}}{\mathrm{d}t}\rho_s(t) = -\frac{i}{\epsilon}[H_s, \rho_s(t)] + \frac{\zeta^2}{\epsilon}\mathcal{R}^\dagger(\rho_s(t)),$$

where the operator $\mathcal{R}^\dagger$ has the form

$$\begin{aligned}\mathcal{R}^\dagger(\rho_s(t)) = &-\frac{1}{2}\sum_{\omega,\omega'} a_F(\omega)\left(\mathbf{D}(\omega')\mathbf{D}^\dagger(\omega)\rho_s(t) - \mathbf{D}^\dagger(\omega)\rho_s(t)\mathbf{D}(\omega') + \text{h.c.}\right) \\ &-\frac{1}{2}\sum_{\omega,\omega'} a_G(\omega)\left(\rho_s(t)\mathbf{D}^\dagger(\omega)\mathbf{D}(\omega') - \mathbf{D}(\omega')\rho_s(t)\mathbf{D}^\dagger(\omega) + \text{h.c.}\right) \\ &- i\sum_{\omega,\omega'} b_F(\omega)\left(\mathbf{D}(\omega')\mathbf{D}^\dagger(\omega)\rho_s(t) - \mathbf{D}^\dagger(\omega)\rho_s(t)\mathbf{D}(\omega') - \text{h.c.}\right) \\ &- i\sum_{\omega,\omega'} b_G(\omega)\left(\rho_s(t)\mathbf{D}^\dagger(\omega)\mathbf{D}(\omega') - \mathbf{D}(\omega')\rho_s(t)\mathbf{D}^\dagger(\omega) - \text{h.c.}\right).\end{aligned} \tag{2.23}$$

If we only keep those terms with $\omega' = \omega$, it becomes the Lindblad equation in (2.10).

Then, consider the perturbation for the system initially at $|\phi_i^0, 0\rangle$ and ending at $|\phi_f^0, 0\rangle$

or $|\phi_f^1, 1\rangle$. We are interested in finding the transition between different quantum states. Because of the symmetry of this problem, it suffices to look at the transition from $|\phi_i^0, 0\rangle$ to $|\phi_f^0, 0\rangle$ and to $|\phi_f^1, 1\rangle$.

Assume that

$$\rho_s(t) = \rho_{s,0}(t) + \frac{\zeta^2}{\epsilon}\rho_{s,1}(t) + \mathcal{O}\Big(\big(\frac{\zeta^2}{\epsilon}\big)^2\Big).$$

Then by collecting terms of the same order,

$$\mathcal{O}(1): \qquad \frac{\mathrm{d}}{\mathrm{d}t}\rho_{s,0}(t) = -\frac{i}{\epsilon}\big[H_s, \rho_{s,0}(t)\big], \qquad \rho_{s,0}(0) = |\phi_i^0, 0\rangle\langle\phi_i^0, 0| \equiv \Pi_i^{(0)};$$

$$\mathcal{O}\Big(\frac{\zeta^2}{\epsilon}\Big): \qquad \frac{\mathrm{d}}{\mathrm{d}t}\rho_{s,1}(t) = -\frac{i}{\epsilon}\big[H_s, \rho_{s,1}(t)\big] + \frac{\zeta^2}{\epsilon}\mathcal{R}^\dagger(\rho_{s,0}(t)), \qquad \rho_{s,1}(0) = 0.$$

For the leading order, the solution is simply $\rho_{s,0}(t) = |\phi_i^0, 0\rangle\langle\phi_i^0, 0| \equiv \Pi_i^{(0)}$. For the first order, after some computing,

$$\begin{aligned}\mathcal{R}^\dagger(\rho_{s,0}(t)) = -\frac{1}{2}\sum_{\omega,\omega'} a_F(\omega)\big(&\Pi_{i+\omega-\omega'}^{(0)}\mathbf{d}\Pi_{i+\omega}^{(1)}\mathbf{d}^\dagger\Pi_i^{(0)} - \Pi_{i+\omega}^{(1)}\mathbf{d}^\dagger\Pi_i^{(0)}\mathbf{d}\Pi_{i+\omega'}^{(1)} \\ &+ \Pi_i^{(0)}\mathbf{d}\Pi_{i+\omega}^{(1)}\mathbf{d}^\dagger\Pi_{i+\omega-\omega'}^{(0)} - \Pi_{i+\omega'}^{(1)}\mathbf{d}^\dagger\Pi_i^{(0)}\mathbf{d}\Pi_{i+\omega}^{(1)}\big) \\ -i\sum_{\omega,\omega'} b_F(\omega)\big(&\Pi_{i+\omega-\omega'}^{(0)}\mathbf{d}\Pi_{i+\omega}^{(1)}\mathbf{d}^\dagger\Pi_i^{(0)} - \Pi_{i+\omega}^{(1)}\mathbf{d}^\dagger\Pi_i^{(0)}\mathbf{d}\Pi_{i+\omega'}^{(1)} \\ &- \Pi_i^{(0)}\mathbf{d}\Pi_{i+\omega}^{(1)}\mathbf{d}^\dagger\Pi_{i+\omega-\omega'}^{(0)} + \Pi_{i+\omega'}^{(1)}\mathbf{d}^\dagger\Pi_i^{(0)}\mathbf{d}\Pi_{i+\omega}^{(1)}\big).\end{aligned}$$

The terms involving $a_G(\omega)$ and $b_G(\omega)$ vanish, since there is no hopping from the state $|1\rangle$

to $|0\rangle$ in our set-up. To study the transition between different quantum states, let

$$\lambda_f^{(m)}(t) \equiv \left\langle \phi_f^0, 0\right| \rho_{s,m}(t) \left|\phi_f^0, 0\right\rangle, \qquad \theta_f^{(m)}(t) \equiv \left\langle \phi_f^1, 1\right| \rho_{s,m}(t) \left|\phi_f^1, 1\right\rangle,$$

where $m = 0, 1, 2, \cdots$ represents the perturbation order. For the leading order, apparently, $\lambda_f^{(0)}(t) = \delta_{i,f}$ and $\theta_f^{(0)}(t) = 0$ for $f \in \mathbb{N}$.

As for the first-order, it could be obtained that

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\lambda_f^{(1)}(t) &= \frac{\zeta^2}{\epsilon} \left\langle \phi_f^0, 0\right| \mathcal{R}^\dagger(\rho_{s,0}(t)) \left|\phi_f^0, 0\right\rangle \\
&= \frac{\zeta^2}{\epsilon}\Big(-\sum_{\omega,\omega'} a_F(\omega)\delta_{i,f}\delta_{\omega,\omega'} \left|\left\langle \phi_i^0 | \phi_{i+\omega}^1\right\rangle\right|^2 \Big) \\
&= -\frac{\zeta^2}{\epsilon}\sum_{\omega} a_F(\omega) \left|\left\langle \phi_i^0 | \phi_{i+\omega}^1\right\rangle\right|^2 \delta_{i,f},
\end{aligned} \tag{2.24}$$

and

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\theta_f^{(1)}(t) &= \frac{\zeta^2}{\epsilon} \left\langle \phi_f^1, 1\right| \mathcal{R}^\dagger(\rho_{s,0}(t)) \left|\phi_f^1, 1\right\rangle \\
&= \frac{\zeta^2}{\epsilon}\sum_{\omega,\omega'} a_F(\omega)\delta_{\omega,\omega'}\delta_{f,i+\omega} \left|\left\langle \phi_i^0 | \phi_f^1\right\rangle\right|^2 \\
&= \frac{\zeta^2}{\epsilon} a_F(f-i) \left|\left\langle \phi_i^0 | \phi_f^1\right\rangle\right|^2 .
\end{aligned}$$

As we observe, for the diagonal elements of $\mathcal{R}^\dagger(\rho_{s,0}(t))$, the only contribution comes from terms with $\omega = \omega'$. Therefore, if this condition $\omega' = \omega$ is imposed beforehand, the results for $\frac{\mathrm{d}}{\mathrm{d}t}\lambda_f^{(1)}(t)$ and $\frac{\mathrm{d}}{\mathrm{d}t}\theta_f^{(1)}(t)$ do not change. This implies that if we apply the perturbation analysis to the Lindblad equation, then we should obtain exactly the same set of equations for the diagonal elements up to the first order.

It could be noticed that the rate of probability leaving $|\phi_i^0, 0\rangle$ is the summation of rates of probability entering $|\phi_f^1, 1\rangle$ for varying $f \in \mathbb{N}$. More specifically, the rate of leaving $|\phi_i^0, 0\rangle$ is $\frac{\zeta^2}{\epsilon}\sum_\omega a_F(\omega)\left|\langle\phi_i^0|\phi_{i+\omega}^1\rangle\right|^2$, and it hops to the state $|\phi_f^1, 1\rangle$ with rate $\frac{\zeta^2}{\epsilon}a_F(\omega)\left|\langle\phi_i^0|\phi_{i+\omega}^1\rangle\right|^2$ if $f = i+\omega$ for some $\omega \in \mathbb{Z}$.

2.6.2 More general initial condition

A slightly more general initial condition is to assume that ρ_0 and ρ_1 are diagonal, i.e., $\rho_0(0) = \sum_k \lambda_k |\phi_k^0\rangle\langle\phi_k^0|$ and $\rho_1(0) = \sum_k \theta_k |\phi_k^1\rangle\langle\phi_k^1|$. Then the leading order is

$$\lambda_k^{(0)}(t) = \lambda_k, \qquad \theta_k^{(0)}(t) = \theta_k, \qquad \forall k \in \mathbb{N},\ t \geq 0.$$

Then we need to compute the first order,

$$\begin{aligned}
\mathcal{R}^\dagger(\rho_{s,0}(t)) = &-\frac{1}{2}\sum_{\omega,\omega',i}\lambda_i a_F(\omega)\big(\Pi_{i+\omega-\omega'}^{(0)}\mathbf{d}\Pi_{i+\omega}^{(1)}\mathbf{d}^\dagger\Pi_i^{(0)} - \Pi_{i+\omega}^{(1)}\mathbf{d}^\dagger\Pi_i^{(0)}\mathbf{d}\Pi_{i+\omega'}^{(1)} + \text{h.c.}\big) \\
&-i\sum_{\omega,\omega',i}\lambda_i b_F(\omega)\big(\Pi_{i+\omega-\omega'}^{(0)}\mathbf{d}\Pi_{i+\omega}^{(1)}\mathbf{d}^\dagger\Pi_i^{(0)} - \Pi_{i+\omega}^{(1)}\mathbf{d}^\dagger\Pi_i^{(0)}\mathbf{d}\Pi_{i+\omega'}^{(1)} - \text{h.c.}\big) \\
&-\frac{1}{2}\sum_{\omega,\omega',i}\theta_i a_G(\omega)\big(\Pi_i^{(1)}\mathbf{d}^\dagger\Pi_{i-\omega}^{(0)}\mathbf{d}\Pi_{i-\omega+\omega'}^{(1)} - \Pi_{i-\omega'}^{(0)}\mathbf{d}\Pi_i^{(1)}\mathbf{d}^\dagger\Pi_{i-\omega}^{(0)} + \text{h.c.}\big) \\
&-i\sum_{\omega,\omega',i}\theta_i b_G(\omega)\big(\Pi_i^{(1)}\mathbf{d}^\dagger\Pi_{i-\omega}^{(0)}\mathbf{d}\Pi_{i-\omega+\omega'}^{(1)} - \Pi_{i-\omega'}^{(0)}\mathbf{d}\Pi_i^{(1)}\mathbf{d}^\dagger\Pi_{i-\omega}^{(0)} - \text{h.c.}\big).
\end{aligned}$$

After some computation,

$$
\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}t}\lambda_f^{(1)}(t) = \frac{\zeta^2}{\epsilon}\left\langle \phi_f^0, 0\right| \mathcal{R}^\dagger(\rho_{s,0}(t)) \left|\phi_f^0, 0\right\rangle \\
&= \frac{\zeta^2}{\epsilon}\Big(-\sum_{\omega,\omega',i} \lambda_i a_F(\omega)\delta_{i,f}\delta_{\omega,\omega'} \left|\left\langle \phi_i^0 | \phi_{i+\omega}^1\right\rangle\right|^2 + \sum_{\omega,\omega',i} \theta_i a_G(\omega)\delta_{i-\omega,f}\delta_{\omega,\omega'} \left|\left\langle \phi_f^0 | \phi_i^1\right\rangle\right|^2 \Big) \\
&= -\frac{\zeta^2}{\epsilon}\sum_{\omega} \lambda_f a_F(\omega) \left|\left\langle \phi_f^0 | \phi_{f+\omega}^1\right\rangle\right|^2 + \frac{\zeta^2}{\epsilon}\sum_{i} \theta_i a_G(i-f) \left|\left\langle \phi_f^0 | \phi_i^1\right\rangle\right|^2,
\end{aligned}
\tag{2.25}
$$

and

$$
\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}t}\theta_f^{(1)}(t) = \frac{\zeta^2}{\epsilon}\left\langle \phi_f^1, 1\right| \mathcal{R}^\dagger(\rho_{s,0}(t)) \left|\phi_f^1, 1\right\rangle \\
&= \frac{\zeta^2}{\epsilon}\Big(\sum_{\omega,\omega',i} \lambda_i a_F(\omega)\delta_{\omega,\omega'}\delta_{f,i+\omega} \left|\left\langle \phi_i^0 | \phi_f^1\right\rangle\right|^2 - \sum_{\omega,\omega',i} \theta_i a_G(\omega)\delta_{i,f}\delta_{\omega,\omega'} \left|\left\langle \phi_i^1 | \phi_{i-\omega}^0\right\rangle\right|^2 \Big) \\
&= \frac{\zeta^2}{\epsilon}\sum_{i} \lambda_i a_F(f-i) \left|\left\langle \phi_i^0 | \phi_f^1\right\rangle\right|^2 - \frac{\zeta^2}{\epsilon}\sum_{\omega} \theta_f a_G(\omega) \left|\left\langle \phi_f^1 | \phi_{f-\omega}^0\right\rangle\right|^2.
\end{aligned}
\tag{2.26}
$$

Hence, the hopping rate from $|\phi_i^0, 0\rangle$ to $|\phi_f^1, 1\rangle$ is $\frac{\zeta^2}{\epsilon} a_F(f-i)\left|\langle \phi_i^0|\phi_f^1\rangle\right|^2$, and the hopping rate from $|\phi_i^1, 1\rangle$ to $|\phi_f^0, 0\rangle$ is $\frac{\zeta^2}{\epsilon} a_G\left(i-f\right)\left|\langle \phi_f^0|\phi_i^1\rangle\right|^2$. Because $\lambda_k(t)$ and $\theta_k(t)$ have an interpretation as the probability at the state $|\phi_k^0, 0\rangle$ or $|\phi_k^1, 1\rangle$, respectively, (2.25) and (2.26) can be interpreted as the *Kolmogorov's backward equation* for a continuous time Markov Chain. To ensure that the first-order perturbation is valid, we assume $t \ll \frac{\epsilon}{\zeta^2}$.

The above computation shows that if initially $\rho_0(0)$ and $\rho_1(0)$ are diagonal, then both

Redfield and Lindblad equations lead to the same transition rates up to the first order. Since the operator $\mathcal{R}^\dagger$ is linear, this result is just a natural extension of that in the last subsection.

2.6.3 Interpretation of hopping rates in the LCME

We continue to assume that $\rho_m(0)$ ($m = 0, 1$) are diagonal, and in this subsection, we will consider the Lindblad equation only. It could be easily verified that $\rho_m(t)$ are diagonal for any time $t \geq 0$. Hence,

$$\rho_0(t) = \sum_k \lambda_k(t) |\phi_k^0\rangle\langle\phi_k^0|.$$

Consider the jumping leaving $|0\rangle$ to $|1\rangle$ in the LCME, that is,

$$\frac{\zeta^2}{\epsilon} \sum_\omega a_F(\omega) \left(\langle 0|\mathbf{D}(\omega)\mathbf{D}^\dagger(\omega)|0\rangle\right)_{\mathcal{W}} \varrho_0(x, p, t)$$

in (2.21). By the definition of $\mathbf{D}^{(\dagger)}(\omega)$ in (2.13), and by Lem. 2.3, we could obtain

$$\begin{aligned}
&\frac{\zeta^2}{\epsilon} \sum_\omega a_F(\omega) \left(\langle 0|\mathbf{D}(\omega)\mathbf{D}^\dagger(\omega)|0\rangle\right)_{\mathcal{W}} \varrho_0(x, p, t) \\
=& \frac{\zeta^2}{\epsilon} \sum_\omega a_F(\omega) \left(\langle 0|\mathbf{D}(\omega)\mathbf{D}^\dagger(\omega)|0\rangle \frac{1}{2\pi\epsilon}\rho_0\right)_{\mathcal{W}} + \mathcal{O}(\zeta^2) \\
=& \frac{\zeta^2}{\epsilon} \sum_\omega a_F(\omega) \sum_k \left|\langle\phi_k^0|\phi_{k+\omega}^1\rangle\right|^2 \lambda_k(t) \frac{1}{2\pi\epsilon} \left(|\phi_k^0\rangle\langle\phi_k^0|\right)_{\mathcal{W}} + \mathcal{O}(\zeta^2) \\
=& \sum_k \left(\frac{\zeta^2}{\epsilon} \sum_\omega a_F(\omega) \left|\langle\phi_k^0|\phi_{k+\omega}^1\rangle\right|^2\right) \lambda_k(t) \frac{1}{2\pi\epsilon} \left(|\phi_k^0\rangle\langle\phi_k^0|\right)_{\mathcal{W}} + \mathcal{O}(\zeta^2).
\end{aligned}$$

This equation shows that the hopping rate from the state $|0\rangle$ to $|1\rangle$ in the semiclassical limit is the summation of the contribution from each state $|\phi_k^0, 0\rangle$. More specifically, the probability at the state $|\phi_k^0, 0\rangle$ in the semiclassical sense is $\lambda_k(t)\frac{1}{2\pi\epsilon}\left(|\phi_k^0\rangle\langle\phi_k^0|\right)_{\mathcal{W}}$; the hopping rate out of the state $|\phi_k^0\rangle$ is $\frac{\zeta^2}{\epsilon}\sum_\omega a_F(\omega)\left|\langle\phi_k^0|\phi_{k+\omega}^1\rangle\right|^2$; their product is exactly the contribution from the quantum state $|\phi_k^0, 0\rangle$.

Notice that the hopping rate obtained here is consistent with (2.24) obtained by perturbation theory. This matches our intuition and it connects the hopping rate in the LCME with the hopping rate from the Lindblad equation.

2.6.4 Discussion on the Franck-Condon blockade

Using the wide band approximation, the hopping rate could be explicitly computed using Franck-Condon factors [59, 90]. The Franck-Condon factor is

$$\left\langle\phi_n^0|\phi_m^1\right\rangle = \frac{(-1)^{n-N}\sqrt{N!}}{\sqrt{M!}}\exp\Big(-\frac{\mathsf{g}^2}{2\epsilon}\Big)\Big(\frac{\mathsf{g}}{\sqrt{\epsilon}}\Big)^{M-N}\mathsf{L}_N^{(M-N)}\Big(\frac{\mathsf{g}^2}{\epsilon}\Big),$$

where $\mathsf{L}_k^{(k')}(x)$ is the generalized Laguerre polynomial, $N = \min(n, m)$, and $M = \max(n, m)$. Then the hopping rate from $|\phi_n^0, 0\rangle$ to $|\phi_m^1, 1\rangle$ is

$$\frac{\zeta^2\Gamma}{\epsilon}\frac{1}{1+e^{\beta\epsilon(m-n)}}\frac{N!}{M!}\exp\Big(-\frac{\mathsf{g}^2}{\epsilon}\Big)\Big(\frac{\mathsf{g}^2}{\epsilon}\Big)^{M-N}\Big(\mathsf{L}_N^{(M-N)}\big(\mathsf{g}^2/\epsilon\big)\Big)^2.$$

We observe that the hopping rate is small due to the weak phonon-reservoir coupling (i.e., $\frac{\zeta^2}{\epsilon} \ll 1$). The Franck-Condon factor $\exp(-\mathrm{g}^2/\epsilon)$ becomes exponentially small as $\mathrm{g}^2/\epsilon \to \infty$, which is known as the Franck-Condon blockade [90, 124] for large on-the-molecule electron-phonon coupling (i.e., $\mathrm{g} \gg 1$). From the last expression, we notice that the Franck-Condon blockade occurs when the ratio of the electron-phonon coupling rate g^2 and the semiclassical parameter ϵ is large; in particular, even for finite $\mathrm{g} = \mathcal{O}(1)$, if $\epsilon \to 0$, such an exponentially small hopping rate also appears.

2.7 Conclusion

In this chapter, we have revisited the derivation of the Redfield equation by time-convolutionless equation (TCL). Consequently, we have derived the Lindblad equation using the secular approximation, in the context of the Anderson-Holstein model. As an analogy to the classical master equation, the Lindbladian classical master equation (LCME) has been introduced, and its form is given in (2.21) and (2.22). The comparison between Redfield and Lindblad equations from the perspective of perturbation theory is considered: Redfield and Lindblad equations yield the same hopping rate in the first order. In other words, the dynamics of their diagonal elements of the reduced density operator ρ_s are the same, to the first order. The condition of the derivation and the perturbation result both suggest that the Lindblad equation might be a better candidate for studying the Anderson-

Holstein model than the Redfield equation. Our reasons are listed as follows. First, from the derivation, they are at the same level, i.e., both under the weak-coupling limit $\zeta \ll \epsilon$. Second, the condition of deriving the Lindblad equation from the Redfield equation is redundant, and there is, in fact, no further constraint in approximating the Redfield equation by the Lindblad equation. Third, they have the same hopping rates between eigenstates up to the first order in perturbation theory. Finally, the analysis for the Lindblad equation is more accessible.

Usually, for multi-level open quantum systems, the density operator $\rho_s(t)$ does not necessarily have vanishing $\rho_{0,1}(t)$ and $\rho_{1,0}(t)$. These two off-diagonal operators will cause more challenge in both analysis and numerics. As we have shown, for the Redfield equation, if we start from diagonal density operators, i.e., $\rho_s(t) = \rho_0(t)|0\rangle\langle 0| + \rho_1(t)|1\rangle\langle 1|$ at some time $t = t_0$, then it remains to be in this form for all time t. If we start from a general density operator $\rho_s(t)$, then analyzing the time-evolution of off-diagonal terms (i.e., $\rho_{0,1}(t)$ and $\rho_{1,0}(t)$) would be challenging, and it could be our next stage of research. For the Lindblad equation, the same phenomenon appears. This approach of only considering diagonal operators with vanishing $\rho_{0,1}(t)$ and $\rho_{1,0}(t)$, has its restriction, but it could lead to simple equations after applying the Wigner transform. The semiclassical limits for both Redfield and Lindblad equations have a similar form.

Chapter 3

Stochastic dynamical low-rank approximation method with application to high-dimensional Lindblad equations

3.1 Introduction

Many problems in computational physics are challenging to solve due to the *curse of dimensionality*, such as high dimensional master equations and many-body quantum dynamics. In an attempt to resolve this difficulty, many ideas have been proposed: for instance, model reduction methods [6,40], Monte Carlo methods [50,78,120,135]. In many situations, several methods of dimension reduction need to be combined. For instance, after applying the Monte Carlo method to some deterministic dynamics by simulating a stochastic differential equation (SDE), the dimension of that SDE may still be very large. Then, it is attractive to further apply model reduction methods in order to capture the main dynamical flows. This is one of our motivations to study model reduction methods for high-dimensional SDEs arising from high-dimensional PDEs or matrix ODEs. In particular, we are interested in two crucial physical dynamics: Fokker-Planck equations [118] and Lindblad equations [81,98].

This chapter is prepared based on [28].

Fokker-Planck equations and Lindblad equations are the governing master equations to describe system evolution for open classical and quantum systems, respectively, under the Markovian approximation. Both are challenging to solve when the dimension becomes large. In order to resolve this problem, it is standard to consider Monte Carlo methods based on SDEs; Monte Carlo methods utilize the statistical average of sampling trajectories to obtain the quantity of interest. In the quantum case for Lindblad equations, some familiar names for such Monte Carlo methods include *unraveling*, and *stochastic wave-function method* [50, 78].

More specifically, suppose we would like to solve the Fokker-Planck equation $\partial_t \mu_t = \mathcal{A}_t^\dagger \mu_t$, where μ_t is the probability distribution (or measure to be more general) and $\mathcal{A}_t^\dagger$ is a time-parametrized operator mapping a probability distribution to its tangent space. In the particle-based methods, one simulates an SDE X_t with the infinitesimal generator $\mathcal{A}_t$ and with initial condition X_0 drawn from μ_0; thus, the distribution of X_t is exactly μ_t. Similarly, as the quantum analog, let us consider the Lindblad equation (see (1.1) and (1.2)). One may solve it by sampling an SDE X_t such that $\mathbb{E}\big[X_t X_t^\dagger\big]$ is precisely the solution of the Lindblad equation (see Lem. 3.8 below for more details). There are various choices of SDEs, e.g., quantum state diffusion (QSD) [78, 120] and linear quantum state diffusion (LQSD) [25, 120]. While it is also possible to use other stochastic processes such as jumping processes [25, 50], we will limit the scope of our consideration to diffusion-type

Monte Carlo methods, throughout this chapter.

The Monte Carlo method for both Fokker-Planck and Lindblad equations can be described under the same framework:

> Given a $\mathbb{C}^n$-valued SDE X_t, one would like to approximate $\mathbb{E}\left[f(X_t)\right] \equiv \int f \, \mathrm{d}\mu_t$ for a collection of prescribed functions $f \in \mathfrak{F}$, where μ_t is the distribution of X_t.

In the case of Fokker-Planck equations, $\mathfrak{F}$ could be a collection of smooth functions; in the case of Lindblad equations, $\mathfrak{F}$ could be a singleton set $\left\{f(x) = xx^\dagger\right\}$ where $x \in \mathbb{C}^n$. In the sequel, we shall consider $f(x) = xx^\dagger$ only, which turns out to be an interesting and useful choice: in the case of Fokker-Planck equations, choosing this test function $f(x) = xx^\dagger$ means one would like to calculate the second moment of the measure μ_t; in the case of Lindblad equations, choosing this f means one would like to compute the density matrix.

To reduce the computational complexity, a popular approach is the model reduction, that is, to retrieve the dynamics by only capturing the evolution of a lower-dimensional object. For our case, there are two directions:

(i) find a low-rank approximation for X_t, or,

(ii) find a low-rank approximation for μ_t.

In literature, several methods take the first approach, in the flavor of Karhunen-Loève expansion (KLE): for instance, proper orthogonal method (POD) [40, 126], dynamical orthogonal (DO) method [112, 126–128], and dynamical bi-orthogonal method (DyBO) [45, 46]. It is clear that in our context, the realization of randomness in X_t is not essential, whereas the distribution μ_t is the key for accurate approximation. Hence, it is natural to consider the low-rank approximation for μ_t, i.e., on the space of probability measures. Then the problem is formulated as follows:

> Given a collection of prescribed test functions $\mathfrak{F}$ and the time evolution equation of probability measures $\partial_t \mu_t = \mathcal{A}_t^{\dagger} \mu_t$ on $\mathbb{C}^n$, one would like to find a low-rank approximation $\mu_{LR,t} \approx \mu_t$ such that
>
> $$\sup_{f\in\mathfrak{F}} \| \int f \ \mathrm{d}\mu_t - f \ \mathrm{d}\mu_{LR,t} \|$$
>
> is small.

As a remark, in Sec. 3.2, we will use the space of finite signed measures, instead of probability measures to avoid the technicality; please see the discussion in Sec. 3.2 for details.

Our work is motivated by extending the (deterministic) *dynamical low-rank approximation*, introduced by Koch and Lubich in [91] for matrix ODEs, to the stochastic case. The main idea in the dynamical low-rank approximation has been illustrated in the context of matrix ODE [91], summarized as follows.

Consider a matrix ODE system $M_t \in \mathbb{C}^{n\times n}$,

$$\frac{\mathrm{d}}{\mathrm{d}t} M_t = F(t, M_t).$$

The dynamical low-rank approximation method in [91] consists of two steps. Firstly, identify a sub-manifold $\mathcal{M}_r \subset \mathbb{C}^{n\times n}$ and approximate the matrix ODE solution M_t by $M_{LR,t} \in \mathcal{M}_r$ for all t; secondly, the time-evolution of $M_{LR,t}$ is given by

$$\frac{\mathrm{d}}{\mathrm{d}t} M_{LR,t} = \operatorname{argmin}_{v\in\mathcal{T}_{M_{LR,t}}\mathcal{M}_r} \mathcal{D}\big(v, F(t, M_{LR,t})\big), \tag{3.1}$$

where $\mathcal{T}_{M_{LR,t}}\mathcal{M}_r$ is the tangent space of $\mathcal{M}_r$ at the current location $M_{LR,t}$ and $\mathcal{D}(v_1, v_2) := \|v_1 - v_2\|$ is a metric on the tangent space. Thus, the evolution is constructed as close as possible to the solution of the matrix ODE above by projecting $F(t, M_{LR,t})$ onto the tangent space $\mathcal{T}_{M_{LR,t}}\mathcal{M}_r$.

In our proposed method, we adopt this idea to the space of finite signed measures on $\mathbb{C}^n$ with bounded second moment, denoted by $\mathcal{M}$[1]. The subspace $\mathcal{M}_r$, in this case, is defined as the space of finite signed measures supported on a linear subspace of $\mathbb{C}^n$ with dimension at most r. Then we hope to approximate μ_t by $\mu_{LR,t} \in \mathcal{M}_r$. The time-evolution equation of the low-rank approximation is given by

$$\partial_t \mu_{LR,t} \equiv \mathcal{A}^\dagger_{LR,t}\mu_{LR,t} := \operatorname{argmin}_{\nu\in\mathcal{T}_{\mu_{LR,t}}\mathcal{M}_r} \mathcal{D}_{\mathfrak{F}}\left(\nu,\ \mathcal{A}^\dagger_t\mu_{LR,t}\right),$$

where $\mathcal{D}_{\mathfrak{F}}$ is a pseudometric defined in (3.5) below, and $\mathcal{T}_{\mu_{LR,t}}\mathcal{M}_r$ is the tangent space of

[1]We slightly abuse the notation and use $\mathcal{M}$ (as well as $\mathcal{M}_r$) for two different spaces, in order to show the connection between our method and the dynamical low-rank approximation method [91].

$\mathcal{M}_r$ at $\mu_{LR,t}$. We will refer this method as the *stochastic dynamical low-rank approximation* method (or SDLR in abbreviation).

As a concrete example, let $\mathfrak{F} = \left\{ f(x) = xx^\dagger \right\}$ be a singleton set, consisting only one test function ($f : \mathbb{C}^n \to \mathfrak{P}_n \subset \mathbb{C}^{n\times n}$, where $\mathfrak{P}_n$ is the space of $n \times n$ positive semidefinite matrices). Assume that the time evolution equation $\partial_t \mu_t = \mathcal{A}_t^\dagger \mu_t$ is the Fokker-Planck equation of an SDE of the form

$$\mathrm{d}X_t = a(X_t, t)\,\mathrm{d}t + \sum_{j=1}^{N} b_j(X_t, t)\,\mathrm{d}W_j, \tag{3.2}$$

where $X_t \in \mathbb{C}^n$, a and b_j are functions $\mathbb{C}^n \times [0, T] \to \mathbb{C}^n$, and W_j are independent real-valued standard Brownian motions. With some additional assumptions and restrictions, one could obtain a low-rank dynamics given in Thm. 3.2, which is one of the main results in this chapter.

As already mentioned above, the stochastic dynamical low-rank approximation is also motivated by developing efficient methods for the Lindblad equations. In that context, the deterministic low-rank approximation has been studied by Le Bris and Rouchon to find a low-rank approximation of Lindblad equations [95]. In the subsequent work [96], Le Bris, Rouchon, and Roussel also introduced an unraveling scheme for the low-rank QME obtained in [95]. The unraveling of Lindblad equations and its connection with the low-rank approximation will be discussed in Sec. 3.5. In particular, as another main result of this chapter, we establish a commuting diagram of the unraveling and the low-rank approximation, with the proposed SDLR method.

The rest of the chapter is organized as follows. In Sec. 3.2, we will formulate the stochastic dynamical low-rank approximation method in the space of finite signed measures. In Sec. 3.3, we will provide a concrete example, in which we will derive a low-rank dynamics for the Fokker-Planck equation, as well as the low-rank dynamics of the SDE for that Fokker-Planck equation. We will also compare our method with the DO method at the end of Sec. 3.3. In Sec. 3.4, we will verify the consistency of our low-rank approximation, and present a Grönwall-type error bound. In Sec. 3.5, we will establish a commuting relation between the action of dynamical low-rank approximation and the action of unraveling. Then numerical examples will be presented in Sec. 3.6 to demonstrate the performance. In Sec. 3.7, we will give a concise summary.

3.2 Stochastic dynamical low-rank approximation

As we recalled in the introduction, the dynamical low-rank approximation method [91], developed for deterministic ODE dynamics, involves the identification of an approximate sub-manifold and projection onto the tangent space by solving a minimization problem. In this section, we will adopt this idea to formulate a dynamical low-rank approximation method in the space of finite signed measures on $\mathbb{C}^n$ with bounded second moment. This low-rank approximation method offers an abstract framework, for instance, to approximate both Fokker-Planck and Lindblad equations via low-rank dynamics, combined with the particle methods. A concrete example and the corresponding low-rank dynamics will be given in Sec. 3.3 below.

3.2.1 Problem set-up and low-rank approximation

Consider the measure space $(\mathbb{C}^n, \mathcal{B})$, where $\mathcal{B}$ is the σ-algebra of Borel sets on $\mathbb{C}^n$. Denote $\mathcal{M}$ as the collection of finite signed measures with bounded second moment:

$$\mathcal{M} := \left\{ \mu \text{ is a finite signed measure} \;\middle|\; \int |x|^2 \, |\mu|(\,\mathrm{d}x) < \infty \right\},$$

where the positive measure $|\mu|$ is the variation of the measure μ.

Consider a given differentiable trajectory $\mu_t \in \mathcal{M}$ solving

$$\partial_t \mu_t = \mathcal{A}_t^\dagger \mu_t,$$

where $\mathcal{A}_t^\dagger : \mathcal{M} \to \mathcal{M}$ is a given time-dependent operator. In the context of Fokker-Planck equation, $\mathcal{A}_t^\dagger$ is the adjoint operator of the infinitesimal generator of the corresponding SDE. In the context of Lindblad equation, $\mathcal{A}_t^\dagger$ is the adjoint operator of the infinitesimal generator of the diffusion-type unraveling scheme of that Lindblad equation (see Sec. 3.5).

The low-rank approximation of $\mathcal{M}$, denoted by $\mathcal{M}_r$, is a subset of $\mathcal{M}$, containing all measures with support on a r-dimensional linear subspace of $\mathbb{C}^n$. Such low-rankness is used to deal with the problem of high dimensionality of $\mathbb{C}^n$. As a remark, the low-rankness that we explore here is not in the sense of taking an ansatz of the measure in the space $\mathcal{M}$ as a linear combination of a few prescribed measures as a basis, which would be a usual Galerkin approximation in the space of measures. Instead, the low-rankness here means that the measure μ is mostly concentrated on a r-dimensional linear subspace of $\mathbb{C}^n$, where $r \ll n$. Intuitively, this approximation would work well for some dissipative dynamics for which the measure is contracted to some low-dimensional space as time evolves; see Sec. 3.6 for numerical demonstration.

Let us characterize the structure of $\mathcal{M}_r$. For any $\mu_{LR} \in \mathcal{M}_r$, by definition, it is supported on a r-dimensional linear subspace, whose orthonormal basis is $\{U_1, U_2, \cdots, U_r\}$. Then one could define a linear mapping $U : \mathbb{C}^r \to \mathbb{C}^n$ by

$$U : y \in \mathbb{C}^r \to \sum_{j=1}^{r} U_j y_j, \qquad \text{or in the matrix form } U = \begin{pmatrix} U_1 & U_2 & \cdots & U_r \end{pmatrix}.$$

Let us denote the r-dimensional *Stiefel manifold* on $\mathbb{C}^n$ by

$$V_r(\mathbb{C}^n) := \left\{ U : \mathbb{C}^r \to \mathbb{C}^n \text{ is linear, and } U^\dagger U = \mathbb{I}_r \right\}.$$

Then $\mathcal{M}_r$ could be viewed as a collection of measures in $\mathcal{M}$ with support on Ran(U), where Ran(U) is the range of the linear operator $U \in V_r(\mathbb{C}^n)$. The restriction of μ_{LR} on Ran(U) can be represented as a finite signed measure on $\mathbb{C}^r$, given by the pullback

$$\theta(E) := \mu_{LR}(UE) = \mu_{LR}(x \in \text{Ran}(U) : x \in UE)$$
$$= \mu_{LR}(x \in \text{Ran}(U) : U^\dagger x \in E),$$

for any Borel set $E \subset \mathbb{C}^r$. Hence, for any measurable set $F \subset \text{Ran}(U) \in \mathcal{B}$,

$$\mu_{LR}(F) = \mu_{LR}(UU^\dagger F) = \theta(U^\dagger F).$$

Thus, there exists a one-to-one correspondence between $\mathcal{M}_r$ and $V_r(\mathbb{C}^n) \oplus \mathcal{M}_{\mathbb{C}^r}$, where $\mathcal{M}_{\mathbb{C}^r}$ denotes the space of finite signed measures on $\mathbb{C}^r$, with bounded second moment. Hence, the low-rank dynamics on $\mathcal{M}_r$ is equivalent to the dynamics of a pair $\big(U_t, \theta_t\big) \in V_r(\mathbb{C}^n) \oplus \mathcal{M}_{\mathbb{C}^r}$.

Remark. We use in the general framework finite signed measure instead of probability measure to avoid subtleties arising from the geometry of probability measures (due to the positivity); see, e.g., [4, 99]. In practice, we will guarantee that the resulting dynamics yields probability measures by imposing more constraints on the low-rank approximation, see Sec. 3.3 for details.

3.2.2 Tangent space projection

It is well-known that the tangent space of $V_r(\mathbb{C}^n)$ at $U \in V_r(\mathbb{C}^n)$ is given by (see, e.g., [85, Thm. 1.2])

$$\mathcal{T}_U V_r(\mathbb{C}^n) = \left\{ iHU : H^\dagger = H \text{ is a linear operator on } \mathbb{C}^n \right\}. \tag{3.3}$$

Thus, for a differentiable trajectory $U_t \in V_r(\mathbb{C}^n)$, we have $\frac{\mathrm{d}}{\mathrm{d}t} U_t = iH_t U_t$ for some time-dependent Hermitian matrix H_t.

Consider any differentiable trajectory θ_t that

$$\partial_t \theta_t = \mathcal{A}^\dagger_{\theta,t} \theta_t, \tag{3.4}$$

where $\mathcal{A}^\dagger_{\theta,t} : \mathcal{M}_{\mathbb{C}^r} \to \mathcal{M}_{\mathbb{C}^r}$ is some operator. The tangent space of $\mathcal{M}_r$ at (U_t, θ_t) is fully characterized by

$$\mathcal{T}_{U_t} V_r(\mathbb{C}^n) \oplus \mathcal{T}_{\theta_t} \mathcal{M}_{\mathbb{C}^r}.$$

It is straightforward to adopt the idea of tangent space projection in (3.1) to our situation. Consider a natural pseudometric on $\mathcal{M}$

$$\mathcal{D}_{\mathfrak{F}}(\nu_1, \nu_2) := \sup_{f \in \mathfrak{F}} \left\| \int f \ \mathrm{d}\nu_1 - f \ \mathrm{d}\nu_2 \right\|, \qquad \nu_{1,2} \in \mathcal{M}. \tag{3.5}$$

Recall that $\mathfrak{F}$ is a collection of test functions and $\|\cdot\|$ is any suitable norm. In Sec. 3.3, we will choose

$$\mathfrak{F} = \left\{ f(x) = xx^\dagger \right\},$$

and the norm $\|\cdot\|$ as the Hilbert-Schmidt norm. An equivalent choice is that

$$\mathfrak{F} = \Big\{ f(x) = \langle x, Ox \rangle \;\Big|\; \text{Hermitian matrix } O \text{ satisfies } \|O\|_{HS} \leq 1 \Big\},$$

and the norm is simply the absolute value. It is not hard to observe that these two choices are equivalent: from the quantum perspective, finding a good approximation of the density matrix $\int f \,\mathrm{d}\nu \equiv \mathbb{E}_\nu\big[xx^\dagger\big]$ is equivalent to finding a good approximation of all observations as $\langle O \rangle_{\text{avg}} := \mathbb{E}_\nu\big[\langle x, Ox\rangle\big] = \mathrm{Tr}\left(\mathbb{E}_\nu\big[Oxx^\dagger\big]\right) = \big\langle O, \mathbb{E}_\nu\big[xx^\dagger\big]\big\rangle_{HS}$ where the observable O is a Hermitian matrix.

The tangent space projection of $\partial_t \mu_t = \mathcal{A}_t^\dagger \mu_t$ to the tangent space $\mathcal{T}_{\mu_{LR,t}}\mathcal{M}_r$ is then given by

$$\begin{aligned}
&\partial_t \mu_{LR,t} \equiv \mathcal{A}_{LR,t}^\dagger \mu_{LR,t} \\
:=\; &\operatorname{argmin}_{\nu \in \mathcal{T}_{\mu_{LR,t}}\mathcal{M}_r} \mathcal{D}_{\mathfrak{F}}\left(\nu,\; \mathcal{A}_t^\dagger \mu_{LR,t}\right) \\
=\; &\operatorname*{argmin}_{\substack{\widetilde{\mathcal{A}}_{LR,t}^\dagger: \\ \widetilde{\mathcal{A}}_{LR,t}^\dagger \mu_{LR,t} \in \mathcal{T}_{\mu_{LR,t}}\mathcal{M}_r}} \sup_{f \in \mathfrak{F}} \Big\| \int f \;\mathrm{d}\big(\widetilde{\mathcal{A}}_{LR,t}^\dagger \mu_{LR,t}\big) - f \;\mathrm{d}\big(\mathcal{A}_t^\dagger \mu_{LR,t}\big) \Big\|.
\end{aligned} \tag{3.6}$$

In the second line, the minimization problem is reformulated from finding tangent vectors ν to finding differential operators $\widetilde{\mathcal{A}}_{LR,t}^\dagger$.

Equivalently, using the adjoint operators $\widetilde{\mathcal{A}}_{LR,t}$ and $\mathcal{A}_t$, it can be written as

$$\begin{aligned}\partial_t \mu_{LR,t} &= \operatorname*{argmin}_{\substack{\widetilde{\mathcal{A}}_{LR,t}: \\ \widetilde{\mathcal{A}}^{\dagger}_{LR,t}\mu_{LR,t} \in \mathcal{T}_{\mu_{LR,t}}\mathcal{M}_r}} \sup_{f\in\mathfrak{F}} \left\| \int \left(\widetilde{\mathcal{A}}_{LR,t} f\right)\, \mathrm{d}\mu_{LR,t} - \left(\mathcal{A}_t f\right)\, \mathrm{d}\mu_{LR,t} \right\| \\ &= \operatorname*{argmin}_{\substack{\widetilde{\mathcal{A}}_{LR,t}: \\ \widetilde{\mathcal{A}}^{\dagger}_{LR,t}\mu_{LR,t} \in \mathcal{T}_{\mu_{LR,t}}\mathcal{M}_r}} \sup_{f\in\mathfrak{F}} \left\| \mathbb{E}_{\mu_{LR,t}} \left[\widetilde{\mathcal{A}}_{LR,t} f - \mathcal{A}_t f \right] \right\|.\end{aligned} \tag{3.7}$$

In Sec. 3.3, we shall parametrize the infinitesimal generator $\widetilde{\mathcal{A}}_{LR,t}$ by some functions to simplify the minimization problem, and also to avoid the vagueness of generic infinitesimal generator for generic stochastic processes.

3.3 SDLR method for SDEs with N driving Brownian motions and a test function $f(x) = xx^{\dagger}$

Based on the framework in the last section, we may explore various low-rank dynamics, by different $\mathcal{D}_{\mathfrak{F}}$ in (3.5). In this section, the test function space $\mathfrak{F}$ is taken to be a singleton with the only element $f(x) = xx^{\dagger}$, and the norm is the Hilbert-Schmidt norm.

To apply the method developed above, we impose some further restrictions on $\mathcal{T}_{\mu_{LR,t}}\mathcal{M}_r$ (thus on the choice of low-rank dynamics). We will make some further comments on these restrictions after we derive the resulting low-rank dynamics by applying the SDLR method.

(i) In the tangent space of $V_r(\mathbb{C}^n)$, for the Hermitian matrix H_t in (3.3), we assume

$$H_t = \mathcal{Q}_{U_t} H_t \mathcal{P}_{U_t} + \text{h.c.}, \tag{3.8}$$

where the projection operator $\mathcal{P}_{U_t} := U_t U_t^\dagger$ and the projection operator to the orthogonal complement $\mathcal{Q}_{U_t} := \mathbb{I}_n - \mathcal{P}_{U_t}$.

(ii) In the tangent space of $\mathcal{M}_{\mathbb{C}^r}$, consider only those $\mathcal{A}_{\theta,t}^\dagger$ corresponding to the Fokker-Planck equation of some SDE on $\mathbb{C}^r$ (cf. (3.4)). This means that we parametrize $\mathcal{A}_{LR,t}^{(\dagger)}$ by a collection of functions. The exact form will be given below in Lem. 3.1. Because the space of all possible infinitesimal generators is opaque and quite large, hence, choosing infinitesimal generators of a particular form is necessary in practice.

Following the second constraint above, let us denote $Y_t(\omega)$ as the corresponding SDE on $\mathbb{C}^r$ whose infinitesimal generator is $\mathcal{A}_{\theta,t}$. Consider the following family of stochastic dynamics with H, A and B's as parameters:

$$\begin{cases} \dot{U}_t = iH_t U_t, \qquad \text{with } H_t^\dagger = H_t,\ H_t = \mathcal{Q}_{U_t} H_t \mathcal{P}_{U_t} + \text{h.c.}, \\ \mathrm{d}Y = A(Y,t)\,\mathrm{d}t + \displaystyle\sum_{j=1}^{M} B_j(Y,t)\,\mathrm{d}W_j. \end{cases} \tag{3.9}$$

Define $X_{LR,t}(\omega) := U_t Y_t(\omega)$, and denote $\mu_{LR,t}$ as the probability measure induced by the random variable $X_{LR,t}(\omega)$. It is straightforward to check that $X_{LR,t}(\omega)$ satisfies the SDE

$$\mathrm{d}X_{LR} = a_{LR}\left(X_{LR}, t\right)\,\mathrm{d}t + \sum_{j=1}^{M} b_{LR,j}\left(X_{LR}, t\right)\,\mathrm{d}W_j,$$

where, for $1 \leq j \leq M$,

$$\begin{cases} a_{LR}(X_{LR,t}, t) = iH_t X_{LR,t} + U_t A(U_t^{\dagger} X_{LR,t}, t), \\ b_{LR,j}(X_{LR,t}, t) = U_t B_j(U_t^{\dagger} X_{LR,t}, t). \end{cases}$$

Note that M does not have to be the same as N, and $\mathrm{d}W_j$ in (3.9) are not necessarily the same Brownian motions in (3.2); we use $\mathrm{d}W_j$ for both to save notations. Since we are interested in the error in the weak sense, how the randomness is achieved does not matter; what is essential is the infinitesimal generator, independent of the particular realization of Brownian motions.

Lemma 3.1. *For the SDE of X_t in* (3.2)*, the infinitesimal generator, acting on $f(x) = xx^{\dagger}$, is the following*

$$\big(\mathcal{A}_t f\big)(x) = x a^{\dagger}(x,t) + a(x,t)x^{\dagger} + \sum_{j=1}^{N} b_j(x,t) b_j^{\dagger}(x,t).$$

Hence, for the low-rank dynamics $\mu_{LR,t}$, $\mathcal{A}_{LR,t}$ is a family of generators parametrized by

H_t, $A(y,t)$ *and* $B_j(y,t)$ *(*$1 \le j \le M$*), in the following form*

$$\big(\mathcal{A}_{LR,t} f\big)(x) = x\left(iH_t x + U_t A(U_t^\dagger x, t)\right)^\dagger + \left(iH_t x + U_t A(U_t^\dagger x, t)\right) x^\dagger$$
$$+ \sum_{j=1}^{M} U_t B_j(U_t^\dagger x, t) B_j^\dagger(U_t^\dagger x, t) U_t^\dagger.$$

The proof is straightforward, by applying the Itô's formula to $f(x) = xx^\dagger$. We are now ready to apply (3.6) and (3.7) to find the time-evolution of the optimal low-rank dynamics $\mu_{LR,t}$.

Theorem 3.2. *Assume that* $\int f(y)\,\theta_t(\,\mathrm{d}y) \equiv \mathbb{E}_{Y_t\sim\theta_t}\big[Y_t Y_t^\dagger\big]$ *is an invertible* $r \times r$ *matrix for any* $t \in [0,T]$*, then*

(i) The following choices of U_t *and* θ_t *give an optimal low-rank dynamics*

$$\begin{cases} M = N, \qquad A(y,t) = U_t^\dagger a(U_t y, t), \qquad B_j(y,t) = U_t^\dagger b_j(U_t y, t), \\ H_t = (-i)\mathcal{Q}_{U_t}\Big(\mathbb{E}_{\theta_t}\big[a(U_t y, t)(U_t y)^\dagger\big] \\ \qquad + \sum_{j=1}^N \mathbb{E}_{\theta_t}\big[b_j(U_t y, t) b_j^\dagger(U_t y, t)\big]\Big) U_t \mathbb{E}_{\theta_t}[yy^\dagger]^{-1} U_t^\dagger + h.c. \end{cases} \tag{3.10}$$

(ii) The solution is optimal in the sense that, for any given U_t *and* θ_t*,*

$$\min_{\widetilde{\mathcal{A}}_{LR,t}^\dagger} \mathcal{D}_{\mathfrak{F}}\big(\widetilde{\mathcal{A}}_{LR,t}^\dagger \mu_{LR,t}, \mathcal{A}_t^\dagger \mu_{LR,t}\big) = \mathcal{D}_{\mathfrak{F}}\big(\mathcal{A}_{LR,t}^\dagger \mu_{LR,t},\ \mathcal{A}_t^\dagger \mu_{LR,t}\big)$$
$$= \Big\| \sum_{j=1}^{N} \mathcal{Q}_{U_t} \mathbb{E}_{\theta_t}\big[b_j(U_t y, t) b_j^\dagger(U_t y, t)\big] \mathcal{Q}_{U_t} \Big\|_{HS},$$

where the adjoint of $\mathcal{A}_{LR,t}^\dagger$ (i.e., $\mathcal{A}_{LR,t}$) is given by (3.10)*, and where $\widetilde{\mathcal{A}}_{LR,t}^\dagger$ is the adjoint operator of $\widetilde{\mathcal{A}}_{LR,t}$, parametrized by A, B_j and H_t as in Lem. 3.1.*

(iii) The system of time-evolution equations of U_t and θ_t is given by the following,

$$\begin{cases} \mathrm{d}Y = U_t^\dagger a(U_tY,t)\ \mathrm{d}t + \sum_{j=1}^N U_t^\dagger b_j(U_tY,t)\ \mathrm{d}W_j\,, \\ \dfrac{\mathrm{d}U_t}{\mathrm{d}t} = \mathcal{Q}_{U_t}\Bigg(\mathbb{E}\big[a(U_tY,t)Y^\dagger\big] \\ \qquad\qquad + \sum_{j=1}^N \mathbb{E}\big[b_j(U_tY,t)b_j^\dagger(U_tY,t)\big]U_t\Bigg)\mathbb{E}\big[YY^\dagger\big]^{-1}. \end{cases} \tag{3.11}$$

If rank $r = n$, then $\mathcal{Q}_{U_t} = 0$ and $\mathcal{P}_{U_t} = \mathbb{I}_n$. Consequently, $\frac{\mathrm{d}U}{\mathrm{d}t} = 0$ and the infinitesimal generator for the SDE of $X_{LR,t}$ is the same as that for X_t. That means, when full rank is used, the original SDE is recovered; equivalently, the Fokker-Planck equation for μ_t is recovered.

Remark. Since the time evolution of U_t and θ_t can be fully recovered from solving an ODE-SDE coupled system of Y and U (with multiple replica of Y), we shall refer (3.11) as the resulting dynamical low-rank approximation as well, when no confusion arises. After all, we will not solve the equation $\partial_t\mu_{LR,t} = \mathcal{A}_{LR,t}^\dagger\mu_{LR,t}$ directly. Instead, we shall use Monte Carlo methods, i.e., solving (3.11) to estimate $\mu_{LR,t}$ by the empirical measure.

Proof. The main idea of the proof is to use stationary conditions with respect to H, A and

B_j to derive the low-rank dynamics. Recall that $\mathcal{P}_U := UU^\dagger$ and $\mathcal{Q}_U := \mathbb{I}_n - UU^\dagger$. With fixed time t, by (3.6) and (3.7), we know

$$\mathcal{D}_{\mathfrak{F}}\big(\mathcal{A}_{LR,t}^\dagger \mu_{LR,t},\ \mathcal{A}_t^\dagger \mu_{LR,t}\big) = \Big\|\mathbb{E}_{\mu_{LR,t}}\big[\mathcal{A}_{LR,t} f - \mathcal{A}_t f\big]\Big\|_{HS} =: \|C\|_{HS},$$

where

$$\begin{aligned} C := \mathbb{E}_{\mu_{LR,t}} &\Big[x(iH_t x + U_t A(U_t^\dagger x, t) - a(x,t))^\dagger + \text{h.c.}\Big] \\ &+ \mathbb{E}_{\mu_{LR,t}}\left[\sum_{j=1}^{M} U_t B_j(U_t^\dagger x, t) B_j^\dagger(U_t^\dagger x, t) U_t^\dagger - \sum_{j=1}^{N} b_j(x,t) b_j^\dagger(x,t)\right]. \end{aligned}$$

To minimize $\mathcal{D}_{\mathfrak{F}}\big(\mathcal{A}_{LR,t}^\dagger \mu_{LR,t},\ \mathcal{A}_t^\dagger \mu_{LR,t}\big)$, it is equivalent to minimize $\|C\|_{HS}^2 = \langle C, C\rangle_{HS}$. The first order stationary conditions of $\langle C, C\rangle_{HS}$ with respect to A, B_j ($1 \le j \le M$) and H_t give

$$\begin{cases} 0 = \Big\langle U_t^\dagger C U_t,\ \mathbb{E}_{\theta_t}\big[(\delta A) y^\dagger + y(\delta A)^\dagger\big]\Big\rangle_{HS}, \\ 0 = \Big\langle U_t^\dagger C U_t,\ \mathbb{E}_{\theta_t}\big[(\delta B_j) B_j^\dagger + B_j(\delta B_j)^\dagger\big]\Big\rangle_{HS}, \\ 0 = \mathrm{Tr}\,\big(\big[C, \mathbb{E}_{\mu_{LR,t}}[xx^\dagger]\big]\,\delta H\big), \end{cases} \tag{3.12}$$

where δA and δB_j are perturbations of functions A and B_j, respectively, and δH is a Hermitian matrix as the perturbation of H.

Consider the third condition in (3.12). Let us complete the basis of $\mathbb{C}^n$ by extending $\mathrm{Ran}(U_t)$, denoted by $\{U_1, U_2, \cdots, U_r, U_{r+1}, \cdots, U_n\}$. Since δH is arbitrary among all possible perturbations, consider the special choice $\delta H = \lambda U_j \langle U_k, \cdot \rangle + \text{h.c.}$ where $j \le r$ and $k > r$. Note that when $j, k \le r$ or $j, k > r$, by our restriction in (3.8), $\langle U_j, H_t U_k \rangle = 0$. Hence as the perturbation of H, δH should preserve this property. That means, we cannot choose δH with nonzero entries for $j, k \le r$ nor $j, k > r$. Denote $D \equiv \left[C, \mathbb{E}_{\mu_{LR,t}}[xx^\dagger]\right]$, which is anti-Hermitian, that is, $D^\dagger = -D$. Plugging the expression of δH into the third condition in (3.12), we could easily compute that

$$\begin{aligned}
0 &= \lambda \langle U_k, DU_j \rangle + \lambda^* \langle U_j, DU_k \rangle \\
&= \lambda \langle U_k, DU_j \rangle + \lambda^* \left\langle U_k, D^\dagger U_j \right\rangle^* = \lambda \langle U_k, DU_j \rangle - \text{c.c.}
\end{aligned}$$

Hence, $\operatorname{Im}\left(\lambda \langle U_k, DU_j \rangle\right) = 0$. Since $\lambda \in \mathbb{C}$ is arbitrary, $\langle U_k, DU_j \rangle = 0$ for all $j \le r$ and $k > r$. That is to say, $\mathcal{Q}_{U_t} D \mathcal{P}_{U_t} = 0$. By plugging the expression of $D \equiv \left[C, \mathbb{E}_{\mu_{LR,t}}[xx^\dagger]\right]$, we could compute that

$$\begin{aligned}
0 &= \mathcal{Q}_{U_t} C U_t \mathbb{E}_{\theta_t}\left[yy^\dagger\right] U_t^\dagger \\
&= \mathbb{E}_{\theta_t}\left[i\mathcal{Q}_{U_t} H_t U_t yy^\dagger - \mathcal{Q}_{U_t} a(U_t y, t) y^\dagger - \mathcal{Q}_{U_t} \sum_{j=1}^{N} b_j(U_t y, t) b_j^\dagger(U_t y, t) U_t \right] \mathbb{E}_{\theta_t}\left[yy^\dagger\right] U_t^\dagger \\
&= i\mathcal{Q}_{U_t} H_t U_t \mathbb{E}_{\theta_t}\left[yy^\dagger\right]^2 U_t^\dagger - \mathcal{Q}_{U_t} \mathbb{E}_{\theta_t}\left[a(U_t y, t) y^\dagger\right] \mathbb{E}_{\theta_t}\left[yy^\dagger\right] U_t^\dagger \\
&\qquad - \mathcal{Q}_{U_t} \sum_{j=1}^{N} \mathbb{E}_{\theta_t}\left[b_j(U_t y, t) b_j^\dagger(U_t y, t)\right] U_t \mathbb{E}_{\theta_t}[yy^\dagger] U_t^\dagger.
\end{aligned}$$

After multiplying both sides by $U_t \mathbb{E}_{\theta_t}\big[yy^\dagger\big]^{-2}$ on the right, the last equation yields

$$i\mathcal{Q}_{U_t} H_t U_t = \mathcal{Q}_{U_t}\Big(\mathbb{E}_{\theta_t}\big[a(U_t y,t)y^\dagger\big] + \sum_{j=1}^{N}\mathbb{E}_{\theta_t}\big[b_j(U_t y,t)b_j^\dagger(U_t y,t)\big]U_t\Big)\mathbb{E}_{\theta_t}[yy^\dagger]^{-1}.$$

After multiplying $U_t^\dagger$ on the right, and then dividing both sides by i, one could obtain

$$\begin{aligned}\mathcal{Q}_{U_t} H_t \mathcal{P}_{U_t} = (-i)\mathcal{Q}_{U_t}\Big(&\mathbb{E}_{\theta_t}\big[a(U_t y,t)(U_t y)^\dagger\big] \\ &+ \sum_{j=1}^{N}\mathbb{E}_{\theta_t}\big[b_j(U_t y,t)b_j^\dagger(U_t y,t)\big]\Big)U_t\mathbb{E}_{\theta_t}[yy^\dagger]^{-1}U_t^\dagger.\end{aligned}$$

Thus, we have already obtained the expression of $H_t \equiv \mathcal{Q}_{U_t} H_t \mathcal{P}_{U_t} + \text{h.c.}$ (cf. (3.10)). As a remark, up to here, we have not yet used any information nor assumption about A and B_j.

Then, we plug the expression of H_t into C,

$$\begin{aligned}C = \mathcal{P}_{U_t}\mathbb{E}_{\theta_t}\big[\big(U_t A(y,t) - a(U_t y,t)\big)y^\dagger U_t^\dagger\big] + \text{h.c.} + \sum_{j=1}^{M} U_t\mathbb{E}_{\theta_t}\big[B_j(y,t)B_j^\dagger(y,t)\big]U_t^\dagger \\ - \sum_{j=1}^{N}\Big(\mathcal{P}_{U_t}\mathbb{E}_{\theta_t}\big[b_j(U_t y,t)b_j^\dagger(U_t y,t)\big]\mathcal{P}_{U_t} + \mathcal{Q}_{U_t}\mathbb{E}_{\theta_t}\big[b_j(U_t y,t)b_j^\dagger(U_t y,t)\big]\mathcal{Q}_{U_t}\Big).\end{aligned}$$

One could observe that $C = \mathcal{P}_{U_t} C \mathcal{P}_{U_t} + \mathcal{Q}_{U_t} C \mathcal{Q}_{U_t}$ with

$$\begin{aligned}\mathcal{P}_{U_t} C \mathcal{P}_{U_t} = \mathcal{P}_{U_t}\mathbb{E}_{\theta_t}\big[\big(U_t A(y,t) - a(U_t y,t)\big)y^\dagger U_t^\dagger\big] + \text{h.c.} \\ + \sum_{j=1}^{M} U_t\mathbb{E}_{\theta_t}\big[B_j(y,t)B_j^\dagger(y,t)\big]U_t^\dagger - \sum_{j=1}^{N}\mathcal{P}_{U_t}\mathbb{E}_{\theta_t}\big[b_j(U_t y,t)b_j^\dagger(U_t y,t)\big]\mathcal{P}_{U_t},\end{aligned}$$

$$\mathcal{Q}_{U_t} C \mathcal{Q}_{U_t} = -\sum_{j=1}^{N} \mathcal{Q}_{U_t} \mathbb{E}_{\theta_t} \left[b_j(U_t y, t) b_j^\dagger(U_t y, t) \right] \mathcal{Q}_{U_t}.$$

Notice that $\mathcal{Q}_{U_t} C \mathcal{Q}_{U_t}$ does not depend on any parameter we choose, that is, it is independent of A, B_j and H_t. Recall that we would like to minimize

$$\langle C, C \rangle_{HS} = \langle \mathcal{P}_{U_t} C \mathcal{P}_{U_t}, \mathcal{P}_{U_t} C \mathcal{P}_{U_t} \rangle_{HS} + \langle \mathcal{Q}_{U_t} C \mathcal{Q}_{U_t}, \mathcal{Q}_{U_t} C \mathcal{Q}_{U_t} \rangle_{HS}. \tag{3.13}$$

The second term is non-negative and we cannot minimize it further. As for the first term, by choosing

$$M = N, \qquad A(y,t) = U_t^\dagger a(U_t y, t), \qquad B_j(y,t) = U_t^\dagger b_j(U_t y, t),$$

one could easily verify that $\mathcal{P}_{U_t} C \mathcal{P}_{U_t} = 0$. With such choices,

$$\|C\|_{HS} = \left\| \sum_{j=1}^{N} \mathcal{Q}_{U_t} \mathbb{E}_{\theta_t} \left[b_j(U_t y, t) b_j^\dagger(U_t y, t) \right] \mathcal{Q}_{U_t} \right\|_{HS}.$$

Hence, the above choices must be optimal (although it does not imply uniqueness). However, the choice of H is indeed unique under our restrictions. One could straightforwardly verify that this solution satisfies the first two equations in the first order stationary conditions in (3.12); that is to say, the above choice yields $U_t^\dagger C U_t = 0$. Though we don't have to use the first two parts in (3.12) to derive the low-rank dynamics, it is still nice to observe the consistency.

Thus, we have proved the first and the second part of this theorem. The third part follows easily from the first part. □

Let us come back to the two restrictions we made at the beginning of this section.

- The reason to impose the condition $H_t = \mathcal{Q}_{U_t} H_t \mathcal{P}_{U_t} +$h.c. is to remove the redundant degrees of freedom. It is easy to observe that $\mathcal{Q}_{U_t} H_t \mathcal{Q}_{U_t}$ is redundant and does not play any role in $\dot{U}_t$. Thus, we might as well let $\mathcal{Q}_{U_t} H_t \mathcal{Q}_{U_t} \equiv 0$. Further if we assume $\mathcal{P}_{U_t} H_t \mathcal{P}_{U_t} \equiv 0$, then it directly implies that $U_t^\dagger \dot{U}_t = 0$, which is used as the orthogonal constraint in [91]. Conversely, if $U_t^\dagger \dot{U}_t = 0$, then $\mathcal{P}_{U_t} H_t \mathcal{P}_{U_t} = 0$, which shows that such a constraint is similar to the constraint $U_t^\dagger \dot{U}_t = 0$ for matrix ODEs.

  Another reason comes from the above proof. The first order stationary condition with respect to δH yields a unique expression for H_t, which indicates that redundancy has been removed via the constraint $H_t = \mathcal{Q}_{U_t} H_t \mathcal{P}_{U_t} +$ h.c.

- The main reason for imposing constraints in the tangent space of $\mathcal{M}_{\mathbb{C}^r}$ is that the Fokker-Planck type generator automatically helps to preserve the positivity of measure. The whole tangent space is rather big and too opaque to handle, since it might involve generators for other stochastic processes, e.g., jump processes. The mixture of jump process and diffusion makes it more challenging to derive a simple low-rank dynamics; this could be an interesting future research direction, however.

Remark. It is a good place to compare our approach with the dynamical orthogonal (DO) method [127]. Using our notations, the ansatz in the DO method is

$$X_{LR,t}(\omega) = \bar{X}_t + U_t Y_t(\omega),$$

where $\bar{X}_t$ is deterministic and for all t, $\mathbb{E}[Y_t(\omega)] = 0$, $U_t^\dagger \dot{U}_t = 0$, $U_t^\dagger U_t = \mathbb{I}_r$. By re-deriving the low-rank dynamics following the proof in [127], one could obtain that

$$\begin{cases} \dfrac{\mathrm{d}}{\mathrm{d}t}\bar{X}_t = \mathbb{E}\big[a(X_{LR,t}, t)\big], \\ \quad \dot{U}_t = \mathcal{Q}_{U_t}\mathbb{E}\big[a(X_{LR,t}, t)Y_t^\dagger\big]\mathbb{E}\big[Y_t Y_t^\dagger\big]^{-1}, \\ \quad \mathrm{d}Y_t = U_t^\dagger \left(a(X_{LR,t}, t) - \mathbb{E}\big[a(X_{LR,t}, t)\big]\right) \mathrm{d}t + \displaystyle\sum_{j=1}^N U_t^\dagger b_j(X_{LR,t}, t)\, \mathrm{d}W_j. \end{cases} \tag{3.14}$$

Compared with (3.11), the time-evolution equations of Y_t and U_t are almost the same except the followings:

- The extra term $U_t^\dagger \mathbb{E}[a(X_{LR,t}, t)]$ in the DO method (3.14), due to the zero-mean constraint in Y_t;
- The non-trivial difference that in (3.11), we have a term

  $$\sum_{j=1}^N \mathbb{E}[b_j(X_{LR,t}, t) b_j^\dagger(X_{LR,t}, t)] U_t,$$

  which could be understood as the Itô correction term due to the second moment.

Some numerical experiments comparing the SDLR method and the DO method will be presented in Sec. 3.6: a high-dimensional geometric Brownian motion and a stochastic Burgers' equation.

3.4 Error analysis of the SDLR method

In this section, we provide error analysis for the low-rank dynamics that we have derived in Sec. 3.3. Thm. 3.4 below indicates that the low-rank dynamics in (3.10) (or (3.11)) is optimal under our ansatz, in the following sense: If X_t is itself an SDE, whose range is supported on a rank r subspace, then under some additional assumptions, $X_{LR,t} = X_t$ by choosing the rank-r dynamics. In other words, our low-rank dynamics is consistent. We also prove a Grönwall-type bound for the error propagation in Thm. 3.5.

3.4.1 Consistency: recovering the low-rank dynamics

Lemma 3.3. *Suppose μ_t is the measure induced by X_t, which solves an SDE of the form as in* (3.2) *with initial condition $X_0 \sim \mu_0$. Assume that μ_t is supported on a r-dimensional linear subspace, whose basis forms a linear operator $V_t : \mathbb{C}^r \to \mathbb{C}^n$ with $V_t^\dagger V_t = \mathbb{I}_r$. Then, for $x \in \mathrm{Ran}(V_t)$, the coefficients of the SDE satisfy*

$$a(x,t) = \Big(\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{P}_{V_t}\Big)x + \mathcal{P}_{V_t}a(x,t), \qquad \textit{and} \qquad b_j(x,t) = \mathcal{P}_{V_t}b_j(x,t), \qquad \forall j,$$

and hence,

$$\sum_{j=1}^{N} \mathcal{Q}_{V_t} \mathbb{E}_{\mu_t} \left[b_j(x,t) b_j^\dagger(x,t) \right] \mathcal{Q}_{V_t} = 0.$$

Proof. Since the range of X_t is $\mathrm{Ran}(V_t)$ by assumption, $V_t V_t^\dagger X_t = X_t$. By taking the derivative for both sides (note that V_t is deterministic),

$$\begin{aligned} \mathrm{d}X_t &= \mathrm{d}\big(\mathcal{P}_{V_t} X_t\big) = \Big(\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{P}_{V_t}\Big) X_t \,\mathrm{d}t + \mathcal{P}_{V_t}\,\mathrm{d}X_t \\ &= \left(\Big(\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{P}_{V_t}\Big) X_t + \mathcal{P}_{V_t} a(X_t,t)\right)\,\mathrm{d}t + \sum_{j=1}^{N} \mathcal{P}_{V_t} b_j(X_t,t)\,\mathrm{d}W_j. \end{aligned} \tag{3.15}$$

By matching it with the SDE of X_t, we obtain $a(X_t,t) = \big(\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{P}_{V_t}\big)\, X_t + \mathcal{P}_{V_t} a(X_t,t)$ and $b_j(X_t,t) = \mathcal{P}_{V_t} b_j(X_t,t)$. □

Theorem 3.4. *Besides the same assumptions in Lem. 3.3, we further assume that* $\frac{\mathrm{d}}{\mathrm{d}t}V_t = iF_tV_t$ *with some Hermitian matrix* F_t *which satisfies* $F_t = \mathcal{Q}_{V_t} F_t \mathcal{P}_{V_t} +$ *h.c. If* $\mu_0 = \mu_{LR,0}$ *and* $V_0 = U_0$*, then we have*

$$\mu_t = \mu_{LR,t}, \quad \textit{and} \quad V_t = U_t,$$

for all $t \in [0,T]$*, as long as the low-rank dynamics in* (3.10) *exists (in particular,* $\mathbb{E}_{\theta_t}[yy^\dagger]$ *remains invertible).*

Proof. Fix time t and assume $X_t = X_{LR,t}$ in distribution and $V_t = U_t$. Next, we shall show that $X_{LR,t}$ and X_t satisfy the same SDE locally and $\frac{\mathrm{d}}{\mathrm{d}t}V_t = \frac{\mathrm{d}}{\mathrm{d}t}U_t$. By Thm. 3.2, we could straightforwardly compute that

$$\begin{aligned}
\mathrm{d}X_{LR,t} &= \mathrm{d}\big(U_t Y_t\big) = \Big(\dot{U}_t Y_t\Big)\,\mathrm{d}t + U_t\,\mathrm{d}Y_t \\
&= \dot{U}_t U_t^\dagger X_{LR,t}\,\mathrm{d}t + U_t\Big(U_t^\dagger a(X_{LR,t},t)\ \mathrm{d}t + \sum_{j=1}^N U_t^\dagger b_j(X_{LR,t},t)\,\mathrm{d}W_j\Big) \\
&= \Big(\dot{U}_t U_t^\dagger X_{LR,t} + U_t U_t^\dagger a(X_{LR,t},t)\Big)\,\mathrm{d}t + \sum_{j=1}^N U_t U_t^\dagger b_j(X_{LR,t},t)\,\mathrm{d}W_j \\
&= \Big(\Big(\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{P}_{U_t}\Big)X_{LR,t} + \mathcal{P}_{U_t} a(X_{LR,t},t)\Big)\ \mathrm{d}t + \sum_{j=1}^N \mathcal{P}_{U_t} b_j(X_{LR,t},t)\,\mathrm{d}W_j.
\end{aligned}$$

In the last step, we have used the fact that $U_t \dot{U}_t^\dagger X_{LR,t} = 0$ due to the restriction on the Hermitian matrix H_t (cf. (3.8)). By comparing the last equation with (3.15) and by $V_t = U_t$, we obtain that the SDEs of X_t and $X_{LR,t}$ coincide, as long as $\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{P}_{V_t} = \frac{\mathrm{d}}{\mathrm{d}t}\mathcal{P}_{U_t}$. That means, we still need to verify $\dot{V}_t = \dot{U}_t$. By (3.11),

$$\begin{aligned}
\dot{U}_t &= \mathcal{Q}_{U_t}\Big(\mathbb{E}\big[a(X_{LR},t)Y^\dagger\big] + \sum_{j=1}^N \mathbb{E}\big[b_j(X_{LR},t)b_j^\dagger(X_{LR},t)\big]U_t\Big)\mathbb{E}[YY^\dagger]^{-1} \\
&\overset{\text{Lem. 3.3}}{=} \mathcal{Q}_{U_t}\mathbb{E}\big[a(X_{LR},t)Y^\dagger\big]\mathbb{E}[YY^\dagger]^{-1} \qquad (\text{ by } U_t = V_t) \\
&\overset{\text{Lem. 3.3}}{=} \mathcal{Q}_{U_t}\mathbb{E}\left[\Big(\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{P}_{V_t}\Big)X_{LR,t}Y^\dagger + \mathcal{P}_{V_t}a(X_{LR,t},t)Y^\dagger\right]\mathbb{E}[YY^\dagger]^{-1} \\
&= \mathcal{Q}_{U_t}\Big(\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{P}_{V_t}\Big)U_t = iF_t V_t = \dot{V}_t.
\end{aligned}$$

Therefore, we have $\dot{V}_t = \dot{U}_t$, and the conclusion holds. $\square$

3.4.2 Error bound

We shall consider the following error

$$\mathcal{E}_t := \left\| \mathbb{E}_{\mu_t}\left[xx^\dagger\right] - \mathbb{E}_{\mu_{LR,t}}\left[xx^\dagger\right] \right\|_{HS},$$

which measures the difference of second moments for two measures μ_t and $\mu_{LR,t}$ in the Hilbert-Schmidt norm. Recall that we chose the single test function $f(x) = xx^\dagger$.

Theorem 3.5. *Assume that:*

(i) Throughout the time-evolution of the low-rank dynamics for $t \in [0, T]$,

$$\mathcal{D}_{\mathfrak{F}}\left(\mathcal{A}_{LR,t}^\dagger \mu_{LR,t},\ \mathcal{A}_t^\dagger \mu_{LR,t}\right) \equiv \left\| \sum_{j=1}^{N} \mathcal{Q}_{U_t} \mathbb{E}_{\theta_t}\left[b_j(U_t y, t) b_j^\dagger(U_t y, t)\right] \mathcal{Q}_{U_t} \right\|_{HS} \leq \epsilon^2.$$

Recall that θ_t is the pullback of the restriction of $\mu_{LR,t}$ to $\mathrm{Ran}(U_t)$.

(ii) There exists a function γ_t such that, for $f(x) = xx^\dagger$,

$$\left\| \mathbb{E}_{\mu_t}\left[\mathcal{A}_t f\right] - \mathbb{E}_{\mu_{LR,t}}\left[\mathcal{A}_t f\right] \right\|_{HS} \leq \gamma_t \mathcal{E}_t \equiv \gamma_t \left\| \mathbb{E}_{\mu_t}\left[f\right] - \mathbb{E}_{\mu_{LR,t}}\left[f\right] \right\|_{HS}.$$

Then, the error satisfies the integral inequality

$$\mathcal{E}_t \leq \mathcal{E}_0 + \epsilon^2 t + \int_0^t \gamma_s \mathcal{E}_s \, \mathrm{d}s,$$

and thus by the Grönwall's inequality,

$$\mathcal{E}_t \le \big(\mathcal{E}_0 + \epsilon^2 t\big) \exp\left(\int_0^t \gamma_s \,\mathrm{d}s\right).$$

Proof. The proof follows from the standard error analysis for time evolution equations:

$$\begin{aligned}
\mathcal{E}_t &\equiv \Big\|\mathbb{E}_{\mu_t}[f] - \mathbb{E}_{\mu_{LR,t}}[f]\Big\|_{HS} \\
&\le \mathcal{E}_0 + \Big\| \int_{\mathbb{C}^n} f(x)\big(\mu_t - \mu_0\big)(\mathrm{d}x) - f(x)\big(\mu_{LR,t} - \mu_{LR,0}\big)(\mathrm{d}x) \Big\|_{HS} \\
&= \mathcal{E}_0 + \Big\| \int_{\mathbb{C}^n} f(x)\Big(\int_0^t \mathcal{A}_s^\dagger \mu_s \,\mathrm{d}s\Big)(\mathrm{d}x) - f(x)\Big(\int_0^t \mathcal{A}_{LR,s}^\dagger \mu_{LR,s} \,\mathrm{d}s\Big)(\mathrm{d}x) \Big\|_{HS} \\
&= \mathcal{E}_0 + \Big\| \int_0^t \Big(\int_{\mathbb{C}^n} f(x)\big(\mathcal{A}_s^\dagger \mu_s\big)(\mathrm{d}x) - f(x)\big(\mathcal{A}_{LR,s}^\dagger \mu_{LR,s}\big)(\mathrm{d}x)\Big) \,\mathrm{d}s \Big\|_{HS} \\
&= \mathcal{E}_0 + \Big\| \int_0^t \Big(\int_{\mathbb{C}^n} \left(\mathcal{A}_s f\right)(x) \mu_s(\mathrm{d}x) - \left(\mathcal{A}_{LR,s} f\right)(x) \mu_{LR,s}(\mathrm{d}x)\Big) \,\mathrm{d}s \Big\|_{HS} \\
&\le \mathcal{E}_0 + \int_0^t \Big\| \mathbb{E}_{\mu_s}[\mathcal{A}_s f] - \mathbb{E}_{\mu_{LR,s}}[\mathcal{A}_s f] + \mathbb{E}_{\mu_{LR,s}}[\mathcal{A}_s f] - \mathbb{E}_{\mu_{LR,s}}[\mathcal{A}_{LR,s} f] \Big\|_{HS} \,\mathrm{d}s \\
&\le \mathcal{E}_0 + \int_0^t \Big\| \mathbb{E}_{\mu_s}[\mathcal{A}_s f] - \mathbb{E}_{\mu_{LR,s}}[\mathcal{A}_s f] \Big\|_{HS} \,\mathrm{d}s \\
&\qquad\qquad + \int_0^t \Big\| \mathbb{E}_{\mu_{LR,s}}[\mathcal{A}_s f] - \mathbb{E}_{\mu_{LR,s}}[\mathcal{A}_{LR,s} f] \Big\|_{HS} \,\mathrm{d}s \\
&= \mathcal{E}_0 + \int_0^t \Big\| \mathbb{E}_{\mu_s}[\mathcal{A}_s f] - \mathbb{E}_{\mu_{LR,s}}[\mathcal{A}_s f] \Big\|_{HS} \,\mathrm{d}s + \int_0^t \mathcal{D}_{\mathfrak{F}}\Big(\mathcal{A}_s^\dagger \mu_{LR,s}, \mathcal{A}_{LR,s}^\dagger \mu_{LR,s}\Big) \,\mathrm{d}s \\
&\le \mathcal{E}_0 + \int_0^t \gamma_s \mathcal{E}_s \ \mathrm{d}s + \int_0^t \epsilon^2 \,\mathrm{d}s.
\end{aligned}$$

Thus, we have proved the integral inequality. The rest of the conclusion follows from the Grönwall's inequality. □

Remark. The estimate above holds for arbitrary ϵ. However, in practice, the above bound is most useful when ϵ is small, which takes place when the diffusivity function $\mathcal{Q}_{U_t} b_j(U_t y, t) = \mathcal{O}(\epsilon)$. For instance, SDEs with small noise fall into this type. When ϵ is large, of course, this indicates that the low-rank approximation fails. One might need to use a higher rank to get an accurate approximation of the dynamics, or even use full rank (i.e., solving the original dynamics).

From the proof, the assumption (ii) in Thm. 3.5 is natural in order to use the Grönwall's inequality. To better illustrate the assumption, we provide a concrete example with an explicit form of γ_t.

Example 3.6 (Choice of γ_t for linear drift and diffusion functions)**.** If the SDE in (3.2) has linear drift and diffusion functions, that is,

$$a(x,t) = \mathbf{\Lambda}_t x, \qquad b_j(x,t) = \mathbf{B}_t^{(j)} x,$$

where bold letters $\mathbf{\Lambda}$ and $\mathbf{B}^{(j)}$ are time-dependent matrices on $\mathbb{C}^n$. By Lem. 3.1, one could straightforwardly find that

$$\gamma_t = 2\|\mathbf{\Lambda}_t\|_2 + \sum_{j=1}^{N} \|\mathbf{B}_t^{(j)}\|_2^2$$

satisfies the assumption.

Another example of γ_t for unraveling scheme will be given in (3.18) below.

If the rank $r = n$, then $\mathcal{Q}_{U_t} = 0$ for all t, so that $\epsilon = 0$. If we further let $\mathcal{E}_0 = 0$, then $\mathcal{E}_t = 0$ for all t. This result is consistent with the intuition, since, in this case, the low-rank dynamic is exactly the original SDE. Though the rank r does not appear explicitly in the error estimate, it is implicitly hidden inside ϵ. Loosely speaking, the larger the rank r, the smaller the ϵ. Even though the relation between ϵ and r is not analytically given, Thm. 3.5 can still be useful in practical simulation and in designing adaptive schemes. Numerically, what we need is to set up error tolerance ϵ, and then compute $\left\| \sum_{j=1}^{N} \mathcal{Q}_{U_t} \mathbb{E}_{\theta_t} \left[b_j(U_t y, t) b_j^{\dagger}(U_t y, t) \right] \mathcal{Q}_{U_t} \right\|_{HS}$ on-the-fly; if this quantity is close to ϵ^2, then it indicates that the rank chosen is not large enough anymore, and we should adaptively increase the rank r in order to control the error. This idea has been used in [95] for numerically solving Lindblad equations by the deterministic low-rank approximation.

3.5 Connections to the unraveling of Lindblad equations

In this section, we consider solving Lindblad equations only, and we will discuss the relationship between the unraveling method and the dynamical low-rank approximation method. In particular, we establish a commuting diagram for the action of unraveling and the action of dynamical low-rank approximation under certain conditions.

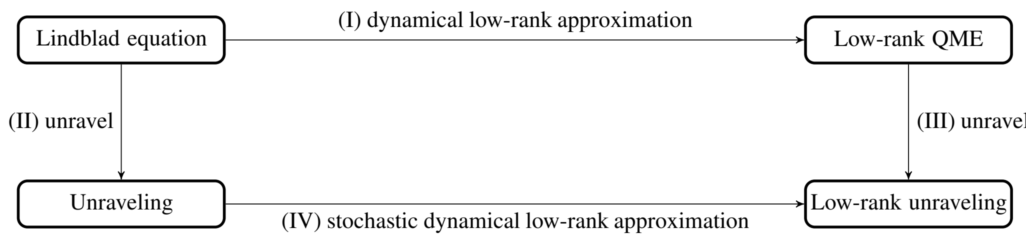


Figure 3.1: Diagram for unraveling and dynamical low-rank approximation for Lindblad equations.

3.5.1 Lindblad equation and stochastic unraveling

Often, the Lindblad equation is challenging to solve numerically due to its high dimension. There are two major dimension-reduction approaches from the literature, summarized in Fig. 3.1: (stochastic) unraveling method (II), and (deterministic) dynamical low-rank approximation method (I).

Step (I): Dynamical low-rank approximation method. This method for Lindblad equations has been studied by Le Bris and Rouchon [95]. The ansatz $\rho_{LR,t} = \widetilde{U}_t \widetilde{\sigma}_t \widetilde{U}_t^\dagger$ is used, where $\widetilde{U}_t$ satisfies the orthonormality constraint $\widetilde{U}_t^\dagger \widetilde{U}_t = \mathbb{I}_r$, and $\widetilde{\sigma}_t$ is a $r \times r$ strictly positive matrix with trace one. Tilde is used to distinguish U_t and σ_t for deterministic dynamical low-rank approximation and those in the SDLR method. The dynamical low-rank approximation method would lead to a coupled ODE system for $\widetilde{U}_t$ and $\widetilde{\sigma}_t$, which approximates the Lindblad equation of ρ_t. We shall revisit the result from [95] below in Thm. 3.7, while dropping the trace-preserving constraint for $\widetilde{\sigma}$. We will discuss the trace-preserving constraint further in Sec. 3.5.3.

Theorem 3.7 (Adapted from [95] with modification)**.** *Consider the subspace*

$$\widetilde{\mathcal{M}}_r := \left\{ \widetilde{U}\widetilde{\sigma}\widetilde{U}^\dagger : \widetilde{U}^\dagger\widetilde{U} = \mathbb{I}_r,\ \widetilde{\sigma} > 0 \right\},$$

to approximate the manifold of positive matrices. Then, by the dynamical low-rank approximation, i.e., by solving

$$\dot{\rho}_{LR,t} = \mathrm{argmin}_{v \in \mathcal{T}_{\rho_{LR,t}}\mathcal{M}_r} \|\mathcal{L}^\dagger(\rho_{LR,t}) - v\|_{HS},$$

with further restriction on the tangent space that $\frac{\mathrm{d}}{\mathrm{d}t}\widetilde{U}_t = iH_t\widetilde{U}_t$*, where the Hermitian matrix* $H_t = \mathcal{Q}_{\widetilde{U}_t} H_t \mathcal{P}_{\widetilde{U}_t} + h.c.$*, we have the unique low-rank dynamics,*

$$\begin{cases} \dfrac{\mathrm{d}}{\mathrm{d}t}\widetilde{\sigma}_t = \widetilde{U}_t^\dagger \mathcal{L}^\dagger(\rho_{LR,t})\widetilde{U}_t, \\ \dfrac{\mathrm{d}}{\mathrm{d}t}\widetilde{U}_t = \mathcal{Q}_{\widetilde{U}_t}\mathcal{L}^\dagger(\rho_{LR,t})\widetilde{U}_t\widetilde{\sigma}_t^{-1}. \end{cases} \tag{3.16}$$

The proof is almost the same as the derivation in [95]. It should be noticed that if the condition $H_t = \mathcal{Q}_{U_t} H_t \mathcal{P}_{U_t} + \text{h.c.}$ is not imposed, the low-rank dynamics is not unique as additional degrees of freedom exist [95]. However, the major difference comes from the trace-preserving condition in [95], which considers the subspace $\widetilde{\mathcal{M}}_r := \{\widetilde{U}\widetilde{\sigma}\widetilde{U}^\dagger : \widetilde{U}^\dagger\widetilde{U} = \mathbb{I}_r,\ \widetilde{\sigma} > 0,\ \mathrm{Tr}(\widetilde{\sigma}) = 1\}$. Consequently, $\lambda\mathbb{I}_r$ term in [95, Eq. (5)] does not appear in (3.16).

Step (II): Unraveling. Recall that unraveling means a stochastic wave-equation that recovers the evolution of density matrices ρ_t. More specifically, it seeks for a stochastic process

$X_t(\omega)$ on $\mathbb{C}^n$, such that $\rho_t := \mathbb{E}\big[X_t(\omega)X_t^\dagger(\omega)\big]$ solves the Lindblad equation (see (1.1) and (1.2)). Multiple choices exist for the unraveling stochastic process: quantum state diffusion (QSD) [78, 120], linear quantum state diffusion (LQSD) [25, 120], and quantum jump process [50]. We restrict the stochastic unraveling to diffusion-type in the discussion below, that is, we shall only consider unraveling of the form as in (3.2). Still various choices exist, while the coefficients satisfy the relation stated in the following Lemma.

Lemma 3.8. *$X_t(\omega)$ in* (3.2) *is a stochastic unraveling of a Lindblad equation* (1.2) *iff*

$$a(x,t)x^\dagger + xa^\dagger(x,t) + \sum_{j=1}^{N} b_j(x,t)b_j^\dagger(x,t) = \mathcal{L}^\dagger\big(xx^\dagger\big), \ \ \forall x \in \mathbb{C}^n, \ \forall t \in [0,T]. \tag{3.17}$$

Among various options of unraveling, two popular choices are:

- Linear quantum state diffusion (LQSD) refers to the choice

$$\begin{cases} a(x) = \Big(-iH - \dfrac{1}{2}\displaystyle\sum_k L_k^\dagger L_k\Big)x, \\ b_{k,1}(x) = \dfrac{1}{\sqrt{2}} L_k x, \\ b_{k,2}(x) = \dfrac{i}{\sqrt{2}} L_k x. \end{cases}$$

  So that the autonomous SDE is given by

$$\mathrm{d}X_t = a(x)\,\mathrm{d}t + \sum_k \left(b_{k,1}(x)\,\mathrm{d}W_{k,1} + b_{k,2}(x)\,\mathrm{d}W_{k,2}\right),$$

where $W_{k,1}$ and $W_{k,2}$ are independent standard (real-valued) Brownian motions. If we have N Lindblad operators, then we have $2N$ (real-valued) Brownian motions in the SDE.

- Quantum state diffusion (QSD) refers to the choice

$$\begin{cases} a(x) = \Big(-iH + \sum_k \langle x, L_k^\dagger x\rangle L_k - \frac{1}{2}L_k^\dagger L_k - \frac{1}{2}|\langle x, L_k^\dagger x\rangle|^2 \Big) x, \\ b_{k,1}(x) = \frac{1}{\sqrt{2}}\big(L_k - \langle x, L_k x\rangle\big) x, \\ b_{k,2}(x) = \frac{i}{\sqrt{2}}\big(L_k - \langle x, L_k x\rangle\big) x. \end{cases}$$

One could easily verify that the above two choices satisfy (3.17).

Step (III): In [96], Le Bris, Rouchon, and Roussel considered an unraveling scheme resulting from the (deterministic) low-rank approximation to Lindblad equations. As the low-rank approximation preserves the structure of the equation, the unraveling is similar to the above discussions. This low-rank unraveling could be used in control variate for the Monte Carlo method for Lindblad equations; see Sec. 3.5.4 below for detail.

3.5.2 A commuting diagram for the dynamical low-rank approximation method and the unraveling method

The route in [96] is from Lindblad equations to low-rank QMEs and then to the unraveling of low-rank QMEs (step (I) to step (III) in Fig. 3.1). It is thus natural to consider an alternative route: i.e., finding the unraveling of Lindblad equations and then applying the SDLR method that we developed in Sec. 3.2 and Sec. 3.3. In Fig. 3.1, this refers to the route from step (II) to step (IV). One immediate question is whether these two routes commute, in the sense that they end up with the same low-rank unraveling. Due to the non-uniqueness of unraveling schemes, in general, the answer is negative. Perhaps, a more specific and reasonable question to ask is that given the unraveling of Lindblad equations, after applying the SDLR method, whether its statistical average recovers the low-rank QME (such as that derived in [95] using the deterministic low-rank approximation).

The answer is definite with slight modification in the (deterministic) dynamical low-rank approximation method in [95]. See Thm. 3.7 for details of the modification in constraints, as well as the resulting low-rank QME. Moreover, it turns out that such a commuting relation does not depend on the particular unraveling scheme chosen for Lindblad equations (see Thm. 3.9).

Theorem 3.9 (Commuting diagram)**.** *For any diffusion-type unraveling scheme of Lindblad equations (see Lem. 3.8), the low-rank unraveling obtained by applying the SDLR method, is an unraveling scheme of the low-rank QME* (3.16)*, obtained via the (deterministic) dynamical low-rank approximation method.*

Proof. By applying the SDLR (step (IV) in Fig. 3.1) to an unraveling scheme of Lindblad equations and using (3.17), one could obtain the following result.

$$\begin{cases} \dot{U}_t = \mathcal{Q}_{U_t}\mathcal{L}^\dagger\big(\mathbb{E}\big[U_tY_tY_t^\dagger U_t^\dagger\big]\big)U_t\mathbb{E}\big[Y_tY_t^\dagger\big]^{-1}, \\ \mathrm{d}Y_t = U_t^\dagger a(U_tY_t,t)\ \mathrm{d}t + \sum_{j=1}^N U_t^\dagger b_j(U_tY_t,t)\ \mathrm{d}W_j. \end{cases}$$

Next, we will verify that the above coupled ODE-SDE system does play the role as an unraveling of the low-rank QME in (3.16), which is the step (III) in Fig. 3.1. By denoting $\sigma_t := \mathbb{E}\big[Y_tY_t^\dagger\big]$, and $\rho_{LR,t} := \mathbb{E}\big[X_{LR,t}X_{LR,t}^\dagger\big] \equiv U_t\sigma_tU_t^\dagger$, one could find that

$$\begin{aligned} \dot{\sigma}_t =&\ \mathbb{E}\Big[Y_ta^\dagger(U_tY_t,t)U_t + U_t^\dagger a(U_tY_t,t)Y_t^\dagger\Big] + \mathbb{E}\Big[\sum_{j=1}^N U_t^\dagger b_j(U_tY_t,t)b_j^\dagger(U_tY_t,t)U_t\Big] \\ \overset{(3.17)}{=}&\ U_t^\dagger\mathbb{E}\left[\mathcal{L}^\dagger\big(U_tY_tY_t^\dagger U_t^\dagger\big)\right]U_t =\ U_t^\dagger\mathcal{L}^\dagger\big(U_t\sigma_tU_t^\dagger\big)U_t. \end{aligned}$$

The time-evolution equation for U_t can be rewritten, in terms of σ_t, as

$$\dot{U}_t = \mathcal{Q}_{U_t}\mathcal{L}^\dagger\big(U_t\sigma_tU_t^\dagger\big)U_t\sigma_t^{-1}.$$

By comparing these two equations with (3.16), one could conclude that the low-rank SDE, after applying the SDLR method, exactly recovers (3.16), i.e., the low-rank unraveling is the unraveling for the low-rank QME in (3.16), and the diagram in Fig. 3.1 indeed commutes in this sense. Also, note that in the above calculation, we have not used any specific choice of unraveling scheme. Thus, the conclusion is independent of the particular unraveling scheme chosen for Lindblad equations. □

Remark. It might not be surprising that the diagram commutes under the above conditions. If we consider $\rho_{LR,t} = \mathbb{E}\big[X_{LR,t}X_{LR,t}^{\dagger}\big]$, by our conditions, $X_{LR,t} = U_t Y_t$, hence $\rho_{LR,t} = U_t\mathbb{E}\big[Y_t Y_t^{\dagger}\big]U_t^{\dagger}$, which is consistent with the ansatz used in [95]. What is interesting is that the commuting diagram result is independent of any unraveling for Lindblad equations, which shows that the low-rank dynamics, derived from the perspective of dynamical low-rank approximation in the space of signed measures, preserves the structure of the Lindblad super-operator $\mathcal{L}^{\dagger}$.

Remark. Recall that in Thm. 3.5, we have studied how the error between the original SDE and the low-rank approximation from the SDLR method propagates with respect to time. In the context of unravelings of Lindblad equations, by Lem. 3.1 and Lem. 3.8, one could show that the growth rate for the error between the unraveling scheme and our corresponding low-rank approximation is bounded by

$$\gamma_t = \|\mathcal{L}^{\dagger}\|_{\mathfrak{A}_0\to\mathfrak{A}_0} := \sup_{H\in\mathfrak{A}_0,\ \|H\|_{HS}=1} \|\mathcal{L}^{\dagger}(H)\|_{HS}. \tag{3.18}$$

3.5.3 Discussion on the trace-preserving restriction

In [95, 96], the trace of low-rank approximated density matrices is required to be one, along the time-evolution. However, we did not consider this condition in Thm. 3.7, nor in the derivation of low-rank dynamics in Thm. 3.2 via the SDLR method. One natural question is that what happens if we impose the trace-preserving condition in the SDLR method. In our set-up and for the Lindblad equation case, the trace-preserving constraint in the SDLR method should be

$$\mathrm{Tr}\big(\rho_{LR,t}\big) \equiv \mathrm{Tr}\left(U_t \mathbb{E}_{\theta_t}\left[yy^\dagger\right] U_t^\dagger\right) = 1, \text{ or equivalently, } \mathrm{Tr}\left(\mathbb{E}_{\theta_t}\left[yy^\dagger\right]\right) = 1,$$

because $U_t^\dagger U_t = \mathbb{I}_r$. From the constraint that $\frac{\mathrm{d}}{\mathrm{d}t}\mathrm{Tr}\left(\mathbb{E}_{\theta_t}\left[yy^\dagger\right]\right) = 0$, one could derive that

$$\mathrm{Tr}\left(\mathbb{E}_{\theta_t}\left[yA(y,t)^\dagger + A(y,t)y^\dagger + \sum_{j=1}^{M} B_j(y,t)B_j^\dagger(y,t)\right]\right) = 0. \tag{3.19}$$

Recall that in the proof of Thm. 3.2, such a constraint does not affect the first-order stationary condition with respect to H_t, hence H_t still has the same form. Therefore, the optimization problem in the derivation is still (cf. (3.13))

$$\min_{A,\ \{B_j\}_{j=1}^M} \left\langle \mathcal{P}_{U_t} C \mathcal{P}_{U_t}, \mathcal{P}_{U_t} C \mathcal{P}_{U_t} \right\rangle_{HS},$$

where

$$
\begin{aligned}
&\mathcal{P}_{U_t} C \mathcal{P}_{U_t} \\
&= \mathcal{P}_{U_t} \mathbb{E}_{\theta_t}\big[\big(U_t A(y,t) - a(U_t y,t)\big) y^\dagger U_t^\dagger\big] + \text{h.c.} \\
&\qquad + \sum_{j=1}^{M} U_t \mathbb{E}_{\theta_t}\big[B_j(y,t) B_j^\dagger(y,t)\big] U_t^\dagger - \sum_{j=1}^{N} \mathcal{P}_{U_t} \mathbb{E}_{\theta_t}\big[b_j(U_t y,t) b_j^\dagger(U_t y,t)\big] \mathcal{P}_{U_t} \\
&\overset{(3.17)}{=} U_t \mathbb{E}_{\theta_t}\left[A(y,t) y^\dagger + y A^\dagger(y,t) + \sum_{j=1}^{M} B_j(y,t) B_j^\dagger(y,t)\right] U_t^\dagger \\
&\qquad\qquad - \mathcal{P}_{U_t} \mathcal{L}^\dagger\left(U_t \mathbb{E}_{\theta_t}\big[y y^\dagger\big] U_t^\dagger\right) \mathcal{P}_{U_t}.
\end{aligned}
$$

If we assume that it is possible to achieve $\mathcal{P}_{U_t} C \mathcal{P}_{U_t} = 0$, then

$$
\mathbb{E}_{\theta_t}\Big[A(y,t) y^\dagger + y A^\dagger(y,t) + \sum_{j=1}^{M} B_j(y,t) B_j^\dagger(y,t)\Big] = U_t^\dagger \mathcal{L}^\dagger\left(U_t \mathbb{E}_{\theta_t}\big[y y^\dagger\big] U_t^\dagger\right) U_t.
$$

The constraint in (3.19) requires that the the trace of the left hand side is zero, while on the right hand side, in general,

$$
\operatorname{Tr}\left(U_t^\dagger \mathcal{L}^\dagger\Big(U_t \mathbb{E}_{\theta_t}\big[y y^\dagger\big] U_t^\dagger\Big) U_t\right) = \operatorname{Tr}\left(\mathcal{P}_{U_t} \mathcal{L}^\dagger\left(U_t \mathbb{E}_{\theta_t}\big[y y^\dagger\big] U_t^\dagger\right)\right) \neq 0,
$$

even though $\operatorname{Tr}\left(\mathcal{L}^\dagger\left(U_t \mathbb{E}_{\theta_t}\big[y y^\dagger\big] U_t^\dagger\right)\right) \equiv 0$. One could conclude that in general, the minimization problem

$$
\min_{A,\, \{B_j\}_{j=1}^M} \langle \mathcal{P}_{U_t} C \mathcal{P}_{U_t}, \mathcal{P}_{U_t} C \mathcal{P}_{U_t} \rangle_{HS} > 0,
$$

under the trace-preserving constraint in (3.19). This could be anticipated since one has to work on a smaller space during minimization. Hence, in general, the low-rank unraveling from the SDLR method with trace-preserving constraint is not optimal in the sense of Thm. 3.2.

In fact, it is not straightforward how to minimize $\langle \mathcal{P}_{U_t} C \mathcal{P}_{U_t}, \mathcal{P}_{U_t} C \mathcal{P}_{U_t} \rangle_{HS}$ under the trace-preserving constraint for functions A and B_j. After all, we need to work on the quadratic variational problem on the functional space with a trace-preserving constraint, not on matrices as in [95].

While it is perhaps desirable to have a trace-preserving dynamical low-rank approximation of density matrices, our choice of not considering trace-preserving constraint can be justified via better approximating measurement outcomes. Recall that the expected measurement outcome for an observable O is $\mathrm{Tr}(O\rho) = \langle O, \rho \rangle_{HS}$. To get an accurate approximation of the measurement outcome, it is sufficient that $\rho_{LR,t}$ is close to ρ_t in the Hilbert-Schmidt norm, without the requirement of a trace-preserving constraint.

3.5.4 Method selection and control variate

From the commuting diagram, one might question the usefulness of the low-rank unraveling in practice. Since in the low-rank unraveling, at each time step, one needs to store $\mathcal{O}(r(n + N_s))$ data for U_t and Y_t, where N_s is the sample size. For the deterministic low-rank dynamics, it only needs to store $\mathcal{O}(r(n+r))$ for $\widetilde{U}_t$ and $\widetilde{\sigma}_t$.

When $r = \mathcal{O}(1)$ is useful to approximate the full dynamics, solving the deterministic low-rank approximation of the Lindblad equation is a better choice, as one does not need to simulate many sample paths to get statistical averages. Moreover, the simulation of deterministic low-rank dynamics has a smaller memory and computational cost.

On the other hand, when rank r requires to be large in order to approximate the system accurately, solving the deterministic low-rank approximation for the Lindblad equation becomes inefficient. It is advantageous to turn to stochastic approximation methods and to use the unraveling of Lindblad equations.

However, one could consider using the control variate method [96] to facilitate the simulation, that is, to use

$$\overline{\rho} = \overline{\rho}_{MC} + \lambda(\rho_{LR} - \overline{\rho}_{\text{LRMC}}),$$

where $\overline{\rho}_{MC}$ is obtained via an unraveling scheme; ρ_{LR} is obtained by solving the deterministic low-rank dynamics of the Lindblad equation; $\overline{\rho}_{\text{LRMC}}$ is obtained by solving a low-rank unraveling scheme. The rank $r = \mathcal{O}(1)$ to ensure that one could simulate both deterministic and stochastic low-rank dynamics. The low-rank dynamics itself may not provide an accurate approximation. However, one could still simulate the low-rank dynamics to achieve the reduction of variance, by choosing the correct parameter λ; see [96] for details.

Generally speaking, for Fokker-Planck equations and Lindblad equations with continuum state space (e.g., probability measures on $\mathbb{R}^d$, or density matrices on $L^2(\mathbb{R}^d)$), one would not want to solve a deterministic low-rank dynamics directly, since that is still a PDE with potentially high dimension when r is not too small (even when $r = 6$, solving such a PDE in 6-dimension is already rather challenging by classical numerical PDE methods). When the low-rank approximation is accurate, one could choose to use low-rank SDEs (derived by the SDLR method) to achieve reduction of complexity in the model; otherwise, one has to simulate the original SDE. Of course, this discussion only involves which model to solve; detailed numerical methods and algorithmic implementations would still make a significant difference to the practical performance.

3.6 Numerical experiments

In this section, we will validate our method by numerical experiments on some high-dimensional SDEs: a high-dimensional geometric Brownian motion, the stochastic Burgers' equation, and unravelings of a quantum damped harmonic oscillator. For the first two examples, we will compare the SDLR method (3.11) with the DO method (3.14). We observe that our SDLR method has comparable performance in approximating the mean when the rank is chosen correctly, compared with the DO method. Moreover, it performs better in approximating the second moment. We will measure the relative error as the indicator of the performance of low-rank approximations:

- the relative error for the mean is defined as

$$\|\mathbb{E}_{\mu_t}[x] - \mathbb{E}_{\mu_{LR,t}}[x]\|/\|\mathbb{E}_{\mu_t}[x]\|\,;$$

- the relative error for the second moment is defined as

$$\|\mathbb{E}_{\mu_t}[xx^\dagger] - \mathbb{E}_{\mu_{LR,t}}[xx^\dagger]\|_{HS}/\|\mathbb{E}_{\mu_t}[xx^\dagger]\|_{HS}\,.$$

Let us comment on some details in numerical implementation. For simplicity, we use the Euler-Maruyama method as the stochastic integrator. An order-one deterministic numerical scheme in [85] is employed to preserve the orthogonality of U_t. The inverse of $\mathbb{E}_{\theta_t}[yy^\dagger]$ in (3.10) or (3.11) could cause numerical instability, when its condition number is significant. This problem also appears in the DO method [127]. [10] suggested using the pseudo-inverse to maintain the algorithmic stability, which is also adopted herein in the numerical simulation. If the chosen rank is representative (not over-estimating the rank), then the pseudo-inverse should be the matrix inverse.

3.6.1 Geometric Brownian motion

Consider the geometric Brownian motion of the form

$$\mathrm{d}X_t = \mathbf{\Lambda} X_t \,\mathrm{d}t + \mathbf{B} X_t \,\mathrm{d}W_t,$$

where $X_t \in \mathbb{C}^{20}$, $\mathbf{\Lambda}$, $\mathbf{B} \in \mathbb{C}^{20\times 20}$.

A rank-5 initial condition is used: $X_0 = x_k$ with probability p_k, where $\{x_k\}_{k=1}^5$ are randomly generated orthogonal vectors in $\mathbb{C}^{20}$, and $p_k \propto \text{Poisson}(k-1, 0.5)$ for $1 \leq k \leq 5$. Poisson(k, λ) is the probability density function of the Poisson distribution with rate λ at value k. $\mathbf{\Lambda}$ is of the form $\mathbf{Q}D\mathbf{Q}^\dagger$ where D is a randomly generated diagonal matrix with diagonal elements uniformly distributed in the interval $[-4.5, -0.5]$, and $\mathbf{Q}$ is a randomly generated orthogonal matrix. We choose $\mathbf{B} = \sqrt{0.05}\,\mathbb{I}_{20}$. It is not difficult to prove that for such a geometric Brownian motion, the second moment $\mathbb{E}_{\mu_t}\left[xx^\dagger\right]$ decays to 0 as $t \to \infty$ (known as the mean-square stability). The result is visualized in Fig. 3.2 and Fig. 3.3 for both SDLR and DO methods. It should be remarked that to faithfully compare the SDLR method with the DO method under the same rank r, e.g., $r = 5$, the matrix U_t in SDLR has dimension 20×5, whereas it has dimension 20×4 for the DO method, since $\bar{X}$ for the DO method should account for one rank and contributes $n = 20$ degrees of freedom.

In our numerical experiment, the sample size is 10^5, and the time step is $\frac{1}{300}$. From

Fig. 3.2, the mean-square stability is observed. In Fig. 3.3, for the SDLR method, when the rank increases, the relative error decreases for both mean and second moment. For the DO method, even for small rank, the relative error of the mean is small; the relative error for the second moment decreases as rank increases as expected. As can be seen, the DO method captures the mean better, and the SDLR method captures the second moment better. This finding is consistent with the theoretical derivation.

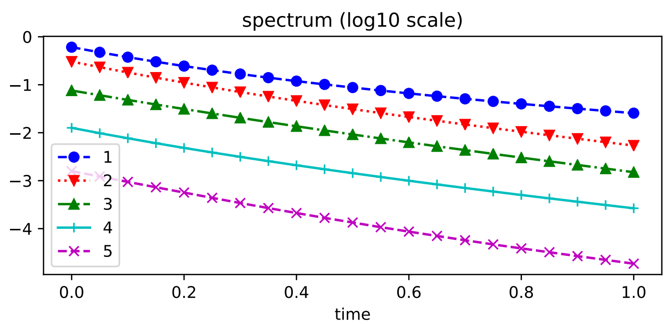


Figure 3.2: Spectrum of $\mathbb{E}_{\mu_t}\left[xx^\dagger\right]$ in $\log 10$ scale (the five largest eigenvalues) for the high-dimensional geometric Brownian motion.

3.6.2 Stochastic Burgers' equation

We consider the stochastic Burgers' equation $h(z,t,\omega)$ of the form

$$\begin{aligned} \mathrm{d}h &= (\nu\partial_z^2 h - h\partial_z h)\,\mathrm{d}t + g(z)\,\mathrm{d}W, \\ h(z,0,\omega) &= h_0(z,\omega), \qquad h(0,t,\omega) = h(1,t,\omega), \end{aligned} \tag{3.20}$$

for $(z,t) \in [0,1] \times [0,T]$. This example is adapted from [117]. Notice that $\mathrm{d}W$ is chosen as a scalar Brownian motion, independent of the spatial coordinate z.

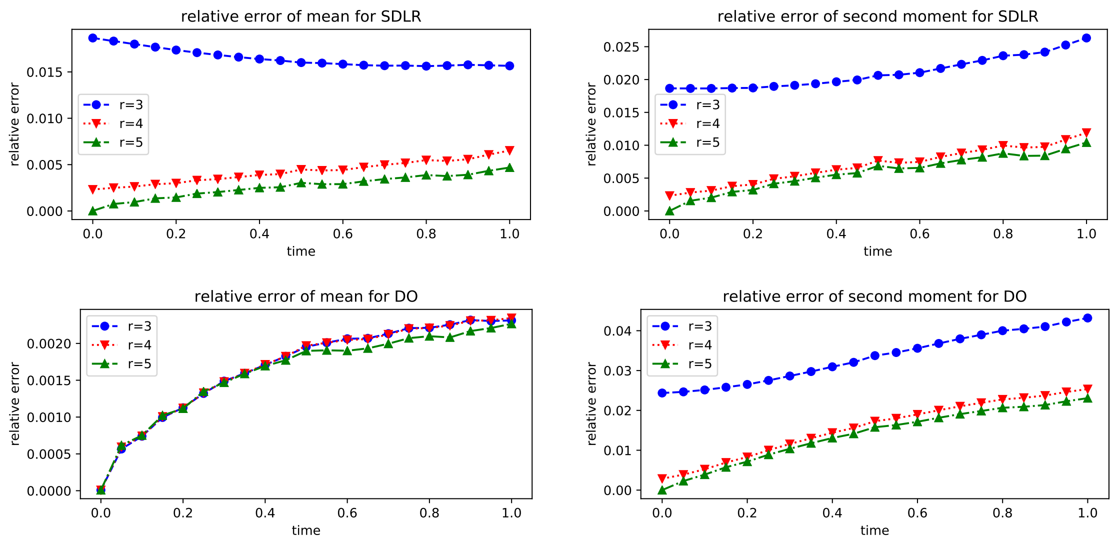


Figure 3.3: Relative errors for SDLR and DO methods in solving the high-dimensional geometric Brownian motion.

Due to the periodic boundary condition, by separating (t, ω) from z, we have

$$h(z, t, \omega) = \sum_{k \in \mathbb{Z}} X_k(t, \omega) e^{2\pi i k z}.$$

Then, the stochastic Burgers' equation could be viewed as an SDE on the Hilbert space $L^2[0, 1]$ with basis functions $\{e^{2\pi i k z}\}_{k \in \mathbb{Z}}$. The SDE has the form

$$\mathrm{d}X_k(t, \omega) = \Big(- (2\pi k)^2 \nu X_k(t, \omega) - \sum_{k'} X_{k-k'}(t, \omega) X_{k'}(t, \omega)(2\pi i k') \Big) \, \mathrm{d}t + g_k \; \mathrm{d}W,$$

$$X_k(0, \omega) = \left\langle e^{2\pi i k z}, h_0(z, \omega) \right\rangle_{L^2([0,1])}, \qquad g_k = \left\langle e^{2\pi i k z}, g(z) \right\rangle_{L^2([0,1])}.$$

Note that $\mathrm{d}W$ is independent of the mode k.

In the direct numerical simulation, we truncate k by letting $X_k(t, \omega) = 0$ for $|k| > \frac{n-1}{2}$ where n is an odd positive integer. Then $X_k(t, \omega)$ can be stored in a $\mathbb{C}^n$ vector. With a careful choice of the initial condition, and let $n \to \infty$, we could expect that the solution of such a truncated SDE converges to the true solution. Adapted from [117, Example 4.1], let $\nu = 0.01$, and $g(z) = \frac{1}{10}\cos(2\pi z) = \frac{1}{20}\left(e^{2\pi i z} + e^{2\pi i(-1)z}\right)$.

As for the initial condition, a rank-5 case is considered

$$h_0(z,\cdot) = \begin{cases} 1, & \text{with probability } p_1; \\ \sqrt{2}\sin(2\pi\lfloor k/2\rfloor z), & \text{with probability } p_k,\ k = 2,4; \\ \sqrt{2}\cos(2\pi\lfloor k/2\rfloor z), & \text{with probability } p_k,\ k = 3,5. \end{cases}$$

We choose $p_k \propto \text{Poisson}(k-1, 0.5)$, just like the last example, and $\lfloor\cdot\rfloor$ is the floor function.

In Fig. 3.4, it could be observed that the numerical result is stable with respect to the dimension of truncation n. Therefore, it is justifiable to simply solve the truncated system with $n = 21$ by SDLR and DO methods.

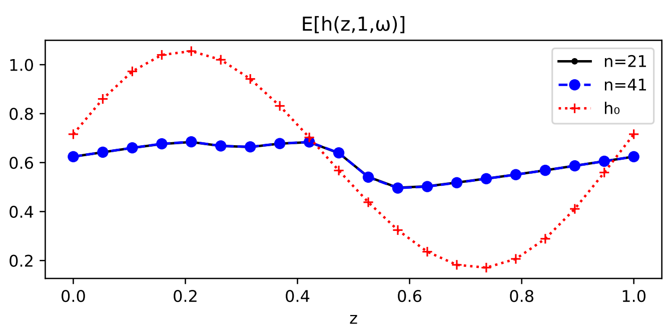


Figure 3.4: This figure shows the expected function $\mathbb{E}[h(z,1,\omega)]$ on $z \in [0,1]$, for different truncated dimension n, in solving the stochastic Burgers' equation.

For $n = 21$, sample size 10^4, and time step $\frac{1}{200}$, the spectrum and relative errors are visualized in Fig. 3.5 and Fig. 3.6, respectively. From the spectrum, the second moment tends to behave like a rank one matrix, since the largest eigenvalue almost keeps constant, while other eigenvalues approximately exponentially decay. In Fig. 3.6, for both methods, when the rank increases, the relative error decreases, as expected. It could be observed that the performance of both SDLR method and DO method is similar and comparable, in calculating both $\mathbb{E}_{\mu_t}[x]$ and $\mathbb{E}_{\mu_t}[xx^\dagger]$. As for more detailed performance comparison for these two methods, Fig. 3.6 is not very informative, especially for the second moment. Therefore, we additionally provide Table 3.1, which gives the ratio between the relative error for the SDLR method and the relative error for the DO method at time $t = 1$, for various ranks. Then, it could be observed that the DO method performs better in approximating the mean, and the SDLR method performs better in approximating the second moment. As a reminder, the result in Table 3.1 can only be interpreted qualitatively, due to the random fluctuation in the simulation.

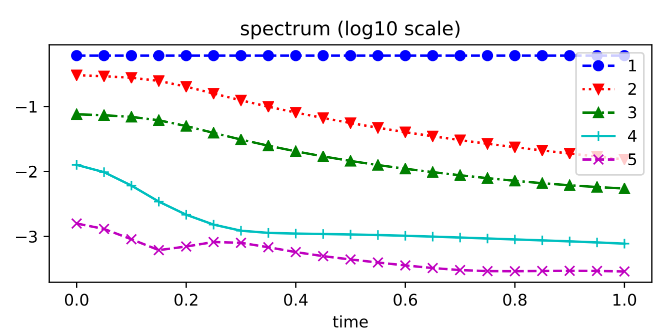


Figure 3.5: Spectrum of $\mathbb{E}_{\mu_t}[xx^\dagger]$ in $\log 10$ scale (the five largest eigenvalues) for the stochastic Burgers' equation.

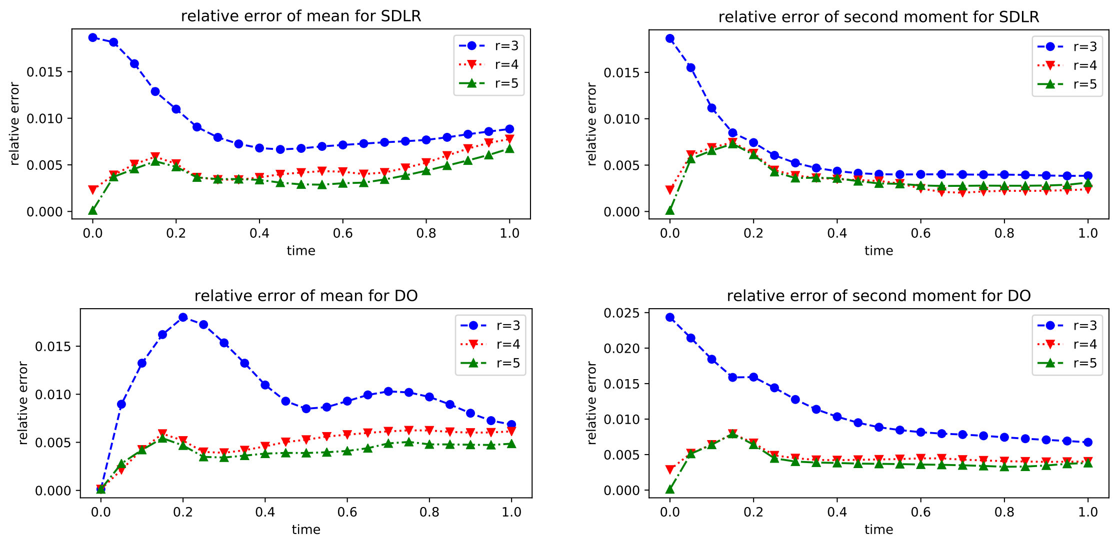


Figure 3.6: Relative errors for SDLR and DO methods in solving the stochastic Burgers' equation.

Table 3.1: This table summarizes the ratio between the relative error for the SDLR method and the relative error for the DO method, at time $t = 1$ for the stochastic Burgers' equation.

r	ratio for relative errors of $\mathbb{E}_{\mu_t}[x]$	ratio for relative errors of $\mathbb{E}_{\mu_t}[xx^\dagger]$
3	1.291	0.572
4	1.264	0.593
5	1.389	0.816

3.6.3 Quantum damped harmonic oscillator

We solve a Lindblad equation of a simple quantum damped harmonic oscillator (see, e.g., [71]), given by

$$\dot{\rho} = -i\left[\mathbf{d}^\dagger \mathbf{d}, \rho\right] + \gamma_1\Big(\mathbf{d}\rho\mathbf{d}^\dagger - \frac{1}{2}\left\{\mathbf{d}^\dagger\mathbf{d}, \rho\right\}\Big) + \gamma_2\Big(\mathbf{d}^\dagger\rho\mathbf{d} - \frac{1}{2}\left\{\mathbf{d}\mathbf{d}^\dagger, \rho\right\}\Big),$$

where $\mathbf{d}^{(\dagger)}$ are annihilation (creation) operators for the harmonic oscillator. Adopting the bra-ket notation, let $\{\ |k\rangle\ \}_{k=0,1,\cdots,n-1}$ be the orthonormal basis of quantum states. Then, the effect of operators $\mathbf{d}^{(\dagger)}$ is $\mathbf{d}^\dagger|k\rangle = \sqrt{k+1}|k+1\rangle$ and $\mathbf{d}|k\rangle = \sqrt{k}|k-1\rangle$. The special state $|0\rangle$ usually refers to the ground state, and $|k\rangle$ ($k \geq 1$) are known as excited states.

We truncate the system by $n = 21$ states. The initial condition is chosen as $|k\rangle$ with probabilities $p_k \propto \text{Poissson}(k, 0.5)$ for $k = 0, 1, \cdots, 4$. Hence, the density matrix ρ_0 is again a rank-5 matrix. This model has been tested for two parameter sets $\gamma_1 = 0.2$, $\gamma_2 = 0$ and $\gamma_1 = 0$, $\gamma_2 = 0.2$. In the first case, the environment acts as an annihilation operator to the system so that the system is moving down to a lower energy state. In the second case, the environment acts as a creation operator so that the system is moving up to a higher energy state. The above Lindblad equation is solved by both QSD and LQSD unraveling schemes, as well as the SDLR method for QSD and LQSD correspondingly. For this example, the solution of the Lindblad equation via a numerical ODE integrator is treated as the exact solution.

The sample size is 3×10^5 and time step is $\frac{1}{500}$. The results are visualized in Fig. 3.7 and Fig. 3.8. The spectrum is consistent with the physical intuition. In Fig. 3.7, the highest eigenvalue roughly indicates the probability at the ground state $|0\rangle$, which steadily increases. This is consistent with the functioning of the annihilation operator, i.e., moving the quantum state to lower energy states. The spectrum in Fig. 3.8 can be explained in a similar way. The solution at rank 5 is comparable with the original unraveling SDE system (i.e., $r = 21$). In the first case $\gamma_1 = 0.2$ and $\gamma_2 = 0$, since the system is lowering down to the ground state, it is expected that the low-rank approximation should perform better.

3.7 Conclusion

In this chapter, we have proposed the stochastic dynamical low-rank approximation method in the space of finite signed measures. Then, by applying the SDLR method, we have derived the low-rank dynamics of SDEs in Thm. 3.2, obtaining an ODE-SDE coupled system as a low-rank approximation of the original high-dimensional SDE in the weak sense. We have also established a commuting diagram for the action of stochastic unraveling and the dynamical low-rank approximation. Furthermore, we have validated this method by error analysis and numerical examples.

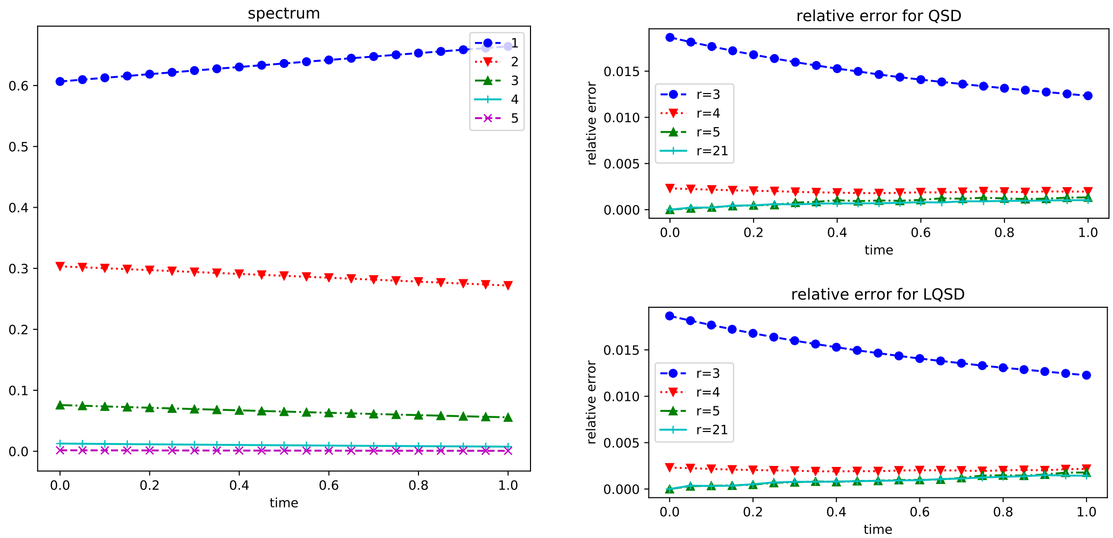


Figure 3.7: Spectrum (the five largest eigenvalues) and relative errors of $\mathbb{E}_{\mu_t}\left[xx^\dagger\right]$, for both QSD and LQSD unraveling schemes in the quantum damped harmonic oscillator with parameters $\gamma_1 = 0.2$, $\gamma_2 = 0$.

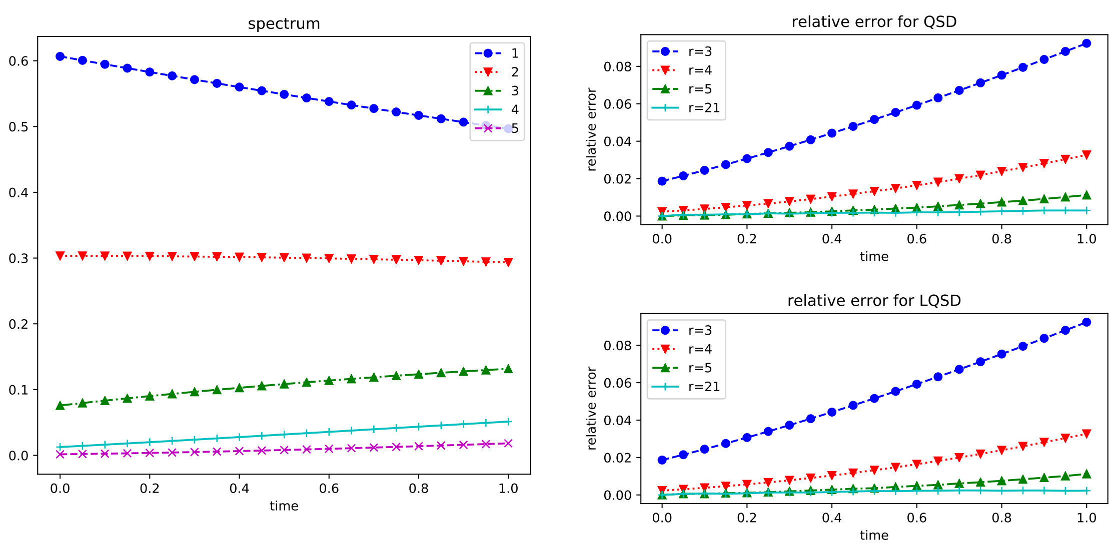


Figure 3.8: Spectrum (the five largest eigenvalues) and relative errors of $\mathbb{E}_{\mu_t}\left[xx^\dagger\right]$, for both QSD and LQSD unraveling schemes in the quantum damped harmonic oscillator with parameters $\gamma_1 = 0$, $\gamma_2 = 0.2$.

Chapter 4

Gradient flow structure and exponential convergence of Lindblad equations with GNS-detailed balance

4.1 Introduction and main results

The characterization of dynamics for open classical and quantum systems is an essential and fundamental topic. One approach to characterize the dynamics for open systems is the entropy production [8, 16, 20, 67, 109, 133], from an information theory perspective. For Markovian dynamics, the entropy production is positive, which relates to the data processing inequality [70, 97, 108] and complies with the second law of thermodynamics [19, 47]. The entropy production also relates to gradient flow structure on well-chosen Riemannian manifolds for those Markov semigroups [35–37, 63, 64, 88, 100, 115]. Furthermore, the validity of log-Sobolev inequalities [11, 53, 82–84, 114, 122] assists in proving the exponential decay of those entropies under respective Markov semigroups.

In this chapter, we focus on the entropy production of sandwiched Rényi divergences, under primitive Lindblad equations with GNS-detailed balance, for finite-dimensional quantum systems. The Lindblad equation with the GNS-detailed balance (see Thm. 4.9 below)

This chapter is prepared based on [30].

has the form

$$\dot{\rho}_t = \mathcal{L}^\dagger(\rho_t) = \sum_{j=1}^{\mathsf{J}} e^{-\omega_j/2}\Big([V_j\rho_t, V_j^\dagger] + [V_j, \rho_t V_j^\dagger]\Big), \tag{4.1}$$

where the positive integer $\mathsf{J} \leq n^2 - 1$; $\omega_j \in \mathbb{R}$; V_j are n-by-n matrices; constraints for ω_j and V_j will be given in Thm. 4.9 below. The notion of GNS-detailed balance will be explained in Sec. 4.2.2 below. The Lindblad equation is called *primitive* (also sometimes referred as *ergodic*, e.g., see [26, 36]) if $\mathrm{Ker}(\mathcal{L})$ is spanned by $\mathbb{I}_n$, which is the n-by-n identity matrix.

Before delving more into Lindblad equations for open quantum systems, we would like to review some previous results of Fokker-Planck equations for open classical systems for comparison. Almost all conclusions in this chapter are motivated by the classical results for Fokker-Planck equations (1.4). Jordan, Kinderlehrer, and Otto [88] showed that such types of Fokker-Planck equations (1.4) could be viewed as Wasserstein gradient flow dynamics of the relative entropy (or the free energy functional). Moreover, based on the method of entropy production and the log-Sobolev inequality [114], one can prove that the solution of the Fokker-Planck equation (1.4) converges to the Gibbs (stationary) distribution exponentially fast in the relative entropy. In our recent work [29], we generalized these results to the classical Rényi divergence of any order $\alpha \in (0, \infty)$: (1) the same Fokker-Planck equation can be formally identified as gradient flow dynamics of the classical Rényi divergence of any order α with respect to a transportation distance defined by a specific metric tensor; (2) the solution of the Fokker-Planck equation converges exponentially fast to the Gibbs distribution in the classical Rényi divergence. Apart from Fokker-Planck equations, several other classical dissipative dynamics can also be viewed as gradient flow dynamics of various transport metrics, for instance, the heat equation [63], finite Markov Chains [100] and porous medium equations [64, 115].

Due to the similar role that Lindblad and Fokker-Planck equations play for open quantum and classical systems, respectively, it is interesting to study Lindblad equations from the perspective of gradient flows. In a previous work by Carlen and Maas, they showed that the Fermionic Fokker-Planck equation could be identified as gradient flow dynamics of the quantum relative entropy [35]. In their more recent work, they extended results from [35], and they showed that for finite-dimensional quantum systems, primitive Lindblad equations with GNS-detailed balance could indeed be identified as gradient flow dynamics of the quantum relative entropy [36]. They also established a unified picture of gradient flows, including the previous two cases and Markov Chains in the setting of C^*-algebras [37]. The generalization of their results from finite-dimensional matrices to infinite-dimensional noncommutative algebras has been recently considered in [142]. In this chapter, we focus on the finite-dimensional setting, and aim to characterize the gradient flow structure of Lindblad equations in terms of a family of quantum divergences — sandwiched Rényi divergences [111, 141], including the quantum relative entropy as a specific instance. Moreover, we are also interested in quantifying the decaying property of the solution ρ_t of the Lindblad equation (4.1) in terms of the sandwiched Rényi divergence.

The sandwiched Rényi divergence is defined as follows.

Definition 4.1 (Sandwiched Rényi divergence [111])**.** Given two positive semi-definite matrices ρ, σ such that $\rho \ll \sigma$ and $\rho \neq 0$, the sandwiched Rényi divergence $D_\alpha\left(\rho||\sigma\right)$ of order $\alpha \in (0, \infty)$ is defined by

$$D_\alpha\left(\rho||\sigma\right) := \begin{cases} \dfrac{1}{\alpha - 1} \log\left(\operatorname{Tr}\left[\left(\sigma^{\frac{1-\alpha}{2\alpha}} \rho \sigma^{\frac{1-\alpha}{2\alpha}}\right)^{\alpha}\right] \right), & \alpha \in (0,1) \cup (1, \infty); \\ \operatorname{Tr}\left(\rho \log(\rho) - \rho \log(\sigma)\right), & \alpha = 1. \end{cases} \tag{4.2}$$

The notation $\rho \ll \sigma$ means that the kernel of σ is a subspace of the kernel of ρ. The term $\operatorname{Tr}(\rho)$ in [111, Def. 2] is simply set as one, since we only consider ρ to be a density matrix herein. σ is assumed to be a full-rank density matrix, so $\rho \ll \sigma$ is always satisfied in our context. For any fixed ρ and full-rank σ, the sandwiched Rényi divergence is continuous with respect to the order α, in particular, $D\left(\rho||\sigma\right) = \lim_{\alpha \to 1} D_\alpha\left(\rho||\sigma\right)$.

The sandwiched Rényi divergence is a quantum analog of the classical Rényi divergence and has recently received much attention due to its applications in quantum information theory [105, 141].

The sandwiched Rényi divergence unifies many entropy measures: when $\alpha = 1$, it reduces to the quantum relative entropy; when $\alpha \to \infty$, it converges to the quantum relative max-entropy; when $\alpha = 1/2$ and $\alpha = 2$, it is closely connected to the quantum fidelity and χ^2 divergence respectively [134].

The sandwiched Rényi divergence satisfies the data processing inequality for any order $\alpha \geq \frac{1}{2}$ and any completely positive trace-preserving (CPTP) map [13, 70, 111]. For fixed density matrices ρ and σ, it is monotonically increasing with respect to the order $\alpha \in (0, \infty)$ [111, Thm. 7]. The sandwiched Rényi divergence has also been used to study the quantum second laws of thermodynamics [19], which is a major motivation of our work. The global convergence rates and mixing times of sandwiched Rényi divergences under quantum Markov semigroups have been studied in [107].

We remark that there is another well-studied family of quantum Rényi divergences known as the *Petz-Rényi* divergence [121], defined by

$$\widetilde{D}_\alpha(\rho||\sigma) := \frac{1}{\alpha - 1} \log\left(\mathrm{Tr}\left(\rho^\alpha \sigma^{1-\alpha}\right)\right).$$

When ρ and σ commute, the sandwiched Rényi divergence is identical to the Petz-Rényi divergence. We also remark that [9] has studied more generalized (α, z)-*quantum Rényi divergences*, which include sandwiched Rényi divergences and Petz-Rényi divergences as special instances. The discussion on the data processing inequality of (α, z)-quantum Rényi divergences can be found in a recent review paper [34] and the references therein.

The first main result of this chapter is the following theorem.

Theorem 4.2. *Consider a primitive Lindblad equation with GNS-detailed balance* (4.1)*, and suppose its unique stationary state σ is full-rank. For any order $\alpha \in (0, \infty)$, consider the energy functional $D_\alpha(\rho||\sigma)$ with respect to the quantum state ρ. In the space of strictly positive density matrices $\mathfrak{D}^+$, equipped with the metric tensor defined in Def. 4.23, the gradient flow dynamics of the sandwiched Rényi divergence $D_\alpha(\rho||\sigma)$ is exactly the Lindblad equation* (4.1).

This theorem immediately implies the monotonicity of the sandwiched Rényi divergence under the evolution of such Lindblad equations, summarized in Cor. 4.3 below. For $\alpha \geq \frac{1}{2}$, this corollary can be proved simply by applying the data processing inequality [70]. For primitive Lindblad equations with GNS-detailed balance, the monotonicity also holds when $\alpha \in (0, \frac{1}{2})$. However, the data processing inequality does not generally hold for sandwiched Rényi divergences when $\alpha < \frac{1}{2}$ [17, Sec. 4].

Corollary 4.3. *Under the same conditions as in Thm. 4.2, suppose $\rho_\cdot : [0, \infty) \to \mathfrak{D}^+$ is the solution of the Lindblad equation* (4.1)*. For any $\alpha \in (0, \infty)$, the sandwiched Rényi divergence $D_\alpha(\rho_t||\sigma)$ is non-increasing in time t, i.e., $\partial_t D_\alpha(\rho_t||\sigma) \leq 0$.*

Thm. 4.2 only provides a sufficient detailed balance condition for Lindblad equations to be gradient flow dynamics of sandwiched Rényi divergences. Later in Sec. 4.4, we will discuss the necessary condition for Lindblad equations to be possibly expressed as gradient flow dynamics of sandwiched Rényi divergences (see Thm. 4.24), by adapting and extending the analysis in [37, Thm. 2.9]. Loosely speaking, the family of Lindblad equations that can be possibly written as gradient flow dynamics of sandwiched Rényi divergences of any order $\alpha \in (0, \infty)$ is not substantially larger than the class of Lindblad equations with GNS-detailed balance, which explains the reason why we restrict our attention to Lindblad equations with GNS-detailed balance in Thm. 4.2. However, identifying a sufficient and necessary condition remains open.

Next, we would like to study the exponential decay of sandwiched Rényi divergences. For Fokker-Planck equations in classical systems, the exponential decay of the relative entropy could be proved via the (classical) log-Sobolev inequality, i.e., the relative entropy is bounded above by the relative Fisher information. This approach has been generalized in our previous work [29] to obtain the exponential decay of the classical Rényi divergence by introducing the *relative α-Fisher information*. For quantum systems, in the same flavor, let us first define the *quantum relative α-Fisher information* as follows.

Definition 4.4. Suppose $\mathcal{L}^\dagger$ is the generator of a primitive Lindblad equation with GNS-detailed balance and the unique stationary state σ is full-rank. The *quantum relative α-*

Fisher information between $\rho \in \mathfrak{D}^+$ and σ is defined as

$$\mathscr{I}_\alpha(\rho||\sigma) := \begin{cases} -\left\langle \frac{\delta D_\alpha(\rho||\sigma)}{\delta \rho}, \mathcal{L}^\dagger(\rho) \right\rangle_{HS}, & \text{if } \alpha \neq 1; \\ -\left\langle \log(\rho) - \log(\sigma), \mathcal{L}^\dagger(\rho) \right\rangle_{HS}, & \text{if } \alpha = 1. \end{cases} \tag{4.3}$$

When $\alpha = 1$, we denote $\mathscr{I}(\rho||\sigma) \equiv \mathscr{I}_1(\rho||\sigma)$ and call it *quantum relative Fisher information* for simplicity. Notice that in the definition above, the quantum relative α-Fisher information depends on the Lindblad super-operator $\mathcal{L}^\dagger$, which is different from the classical case where the relative α-Fisher information only depends on the Gibbs stationary distribution (see [29]). Also, observe that if ρ_t solves the Lindblad equation (4.1), then $\mathscr{I}_\alpha(\rho_t||\sigma) = -\partial_t D_\alpha(\rho_t||\sigma)$. Since $\partial_t D_\alpha(\rho_t||\sigma) \leq 0$ by Cor. 4.3, the quantum relative α-Fisher information $\mathscr{I}_\alpha(\rho_t||\sigma)$ is non-negative.

Definition 4.5. Suppose the density matrix σ is the unique stationary state of the Lindblad equation (4.1). For a fixed order $\alpha \in (0, \infty)$, σ is said to satisfy the *quantum α-log Sobolev inequality* (or quantum α-LSI in short), if there exists $K_\alpha > 0$ such that

$$D_\alpha(\rho||\sigma) \leq \frac{1}{2K_\alpha} \mathscr{I}_\alpha(\rho||\sigma), \qquad \forall \rho \in \mathfrak{D}, \tag{4.4}$$

where $\mathfrak{D}$ is the space of density matrices. K_α is called the *α-log Sobolev constant*. When $\alpha = 1$, we call (4.4) the *quantum log-Sobolev inequality* and denote $K \equiv K_1$ for simplicity.

The α-log Sobolev constant K_α has been considered in [107, Def. 3.1]. When $\alpha = 1$, the above form reduces to the quantum log-Sobolev inequality in [36, Eq. (8.17)].

Remark. Another definition of the quantum α-LSI was given previously under the framework of noncommutative $\mathbb{L}_\alpha$ spaces (see, e.g., [113, Def. 3.5], [89, Def. 11] and [14, Sec. 2.5]). We will briefly revisit this concept in Sec. 4.2.5, and we shall denote the log-Sobolev constant in this set-up by κ_α. The essential difference is that for the definition of κ_α, the entropy function uses the quantum relative entropy for a density matrix raising to the power of α, instead of using sandwiched Rényi divergence (see (4.19) for more details).

For Lindblad equations with GNS-detailed balance, the primitivity of Lindblad equations is equivalent to the validity of the quantum LSI, summarized in Prop. 4.6. It is important to note that by Lem. 4.12 below, if the Lindblad equation satisfies GNS-detailed balance, then $-\mathcal{L}$ is a positive semi-definite operator, with respect to the inner product $\langle \cdot, \cdot \rangle_{1/2}$, which is defined by $\langle A, B \rangle_{1/2} := \operatorname{Tr}\left(A^\dagger \sigma^{1/2} B \sigma^{1/2}\right)$ for all matrices A, B. As a consequence, primitivity of $\mathcal{L}$ implies that $-\mathcal{L}$ has a positive spectral gap, denoted by $\lambda_{\mathcal{L}}$. The proposition below provides two-side estimates of K in terms of $\lambda_{\mathcal{L}}$.

Proposition 4.6. *Consider the Lindblad equation with GNS-detailed balance* (4.1) *and suppose a full-rank density matrix σ is its stationary state. Then, σ satisfies the quantum LSI iff the Lindblad equation is primitive. More quantitatively, if the Lindblad equation is* *primitive, we have*

$$\frac{1}{1-\log(\sqrt{\lambda_{\min}(\sigma)})}\lambda_{\mathcal{L}} \leq K \leq \lambda_{\mathcal{L}},$$

where $\lambda_{\min}(\sigma)$ is the smallest eigenvalue of σ.

Prop. 4.6 follows mainly from [107, Thm. 5.3], [89, Thm. 16] (or see [107, Thm. 4.1]), and the quantum Stroock-Varopoulos inequality [14, Cor. 16]. The proof is postponed to Sec. 4.5.3. We remark that the lower bound above is not sharp, and it remains an interesting question to find a tighter lower bound.

It is not hard to see that if the quantum α-LSI (4.4) is valid, then it follows immediately from Grönwall inequality that (see also [107, Thm. 3.2])

$$D_{\alpha}\left(\rho_t||\sigma\right) \leq D_{\alpha}\left(\rho_0||\sigma\right) e^{-2K_{\alpha}t}, \qquad \forall t \geq 0. \tag{4.5}$$

The validity of (4.4) with a positive K_{α} is, in general, difficult to establish. For the depolarizing semigroup, K_2 has been computed in [107, Thm. 3.3], and K has been studied in [110]. For the depolarizing semigroup with stationary state σ being the maximally mixed state, both upper and lower bounds of K_{α} with $\alpha \geq 2$, have been provided in [107, Thm. 3.4]. Moreover, it has been proved in [107, Thm. 4.1] that the α-log Sobolev constant cannot exceed the spectral gap, namely for any $\alpha \geq 1$, $K_{\alpha} \leq \lambda_{\mathcal{L}}$. In this chapter, instead of pursuing the global decay of the sandwiched Rényi divergence $D_\alpha(\rho_t||\sigma)$ as in (4.5), we focus on the exponential decay of $D_\alpha(\rho_t||\sigma)$ in the large time regime. More specifically, our second main result (Thm. 4.7) shows that the sandwiched Rényi divergence $D_\alpha(\rho_t||\sigma)$ decays exponentially with the sharp rate $2\lambda_{\mathcal{L}}$ when t is large.

Theorem 4.7. *Under the same conditions as in Thm. 4.2, for any given solution* $\rho_\cdot : [0,\infty) \to \mathfrak{D}^+$ *of the Lindblad equation and any* $\epsilon \in \left(0, \frac{\lambda_{\min}(\sigma)^2}{2}\right)$, *there exist* $C_{\alpha,\epsilon}$ *and* $\tau_{\alpha,\epsilon}$ *such that*

$$D_\alpha(\rho_t||\sigma) \leq C_{\alpha,\epsilon} D_\alpha(\rho_0||\sigma)\, e^{-2\lambda_{\mathcal{L}} t}, \qquad \forall t \geq \tau_{\alpha,\epsilon}, \tag{4.6}$$

where

$$C_{\alpha,\epsilon} = \frac{e^{D_2(\rho_0||\sigma)} - 1}{D_\alpha(\rho_0||\sigma)} \exp\left(\Theta(\alpha-2) 2\lambda_{\mathcal{L}} T_{\alpha,\epsilon}\right),$$

and

$$\tau_{\alpha,\epsilon} = \begin{cases} 0, & \alpha \leq 2; \\ T_{\alpha,\epsilon} + \max\left(0, \dfrac{1}{2K}\log(D(\rho_0||\sigma)/\epsilon)\right), & \alpha > 2. \end{cases}$$

In the above, $\lambda_{\min}(\sigma)$ *is the smallest eigenvalue of* σ; $T_{\alpha,\epsilon} = \frac{1}{2K\eta}\log(\alpha-1)$ *is a function depending on* α *and* ϵ; Θ *is the Heaviside function;*

$$\begin{cases} \eta = \min\left(\dfrac{1}{2}, \min\limits_{j=1}^{\mathsf{J}} \left\{\dfrac{2\sqrt{e^{\omega_j}\frac{1}{\Lambda}}}{1+e^{\omega_j}\Lambda}\right\}\right), \\ \Lambda = \dfrac{\lambda_{\max}(\sigma)}{\lambda_{\min}(\sigma)} \exp\left(2\sqrt{2\epsilon}\dfrac{2\lambda_{\min}(\sigma) - \sqrt{2\epsilon}}{\lambda_{\min}(\sigma)(\lambda_{\min}(\sigma) - \sqrt{2\epsilon})}\right), \end{cases}$$

where ω_j are Bohr frequencies determined by the stationary state σ (see Sec. 4.2.2), and $\lambda_{\max}(\sigma)$ is the largest eigenvalue of σ.

The proof of Thm. 4.7 can be found in Sec. 4.5. We sketch the strategies used in the proof herein. First, we prove this theorem for the case $\alpha = 2$ by using a uniform lower bound of $\mathscr{I}_2(\rho||\sigma)$ (see (4.35)). Next, we prove that the quantum LSI (4.4) holds (see Prop. 4.6). Based on the quantum LSI (see (4.4) for $\alpha = 1$), we prove a quantum comparison theorem (see Prop. 4.32), which implies the exponential decay of the sandwiched Rényi divergence of any order $\alpha \in (0, \infty)$ (see Prop. 4.33).

Some remarks of this theorem are in order:

Remark. (i) The upper bounds for $\tau_{\alpha,\epsilon}$ and $C_{\alpha,\epsilon}$ may not be sharp. A better bound might be crucial if one wants to study the mixing time for quantum decoherence, which, however, will not be pursued here. For fixed $\alpha > 2$, when ϵ becomes smaller, $T_{\alpha,\epsilon}$ decreases, and we have a smaller prefactor $C_{\alpha,\epsilon}$, at the expense of waiting longer time $\tau_{\alpha,\epsilon}$ (note that the term $D(\rho_0||\sigma)/\epsilon$ blows up as $\epsilon \downarrow 0^+$).

(ii) The artificial ϵ and the complicated expression of $T_{\alpha,\epsilon}$ come from Prop. 4.32. There is a different way to quantify η under the assumption of the quantum 2-LSI in the sense of noncommutative $\mathbb{L}_\alpha$ spaces; please see Sec. 4.2.5 for background and the discussion in Sec. 4.5.6 for more details. We prefer this version because the log-

Sobolev inequality for $\alpha = 1$ is also used for the classical result in [29, Thm. 1.2]. One could remove the dependence of the parameter ϵ by simply choosing, e.g., $\epsilon = \lambda_{\min}(\sigma)^2/8$ in this theorem. Then,

$$\Lambda = \frac{\lambda_{\max}(\sigma)}{\lambda_{\min}(\sigma)} e^3, \qquad \eta = \min\left(\frac{1}{2}, \min_{j=1}^{J}\left\{\frac{2e^{-3/2}\sqrt{e^{\omega_j}\frac{\lambda_{\min}(\sigma)}{\lambda_{\max}(\sigma)}}}{1+e^{\omega_j}\frac{\lambda_{\max}(\sigma)}{\lambda_{\min}(\sigma)}e^3}\right\}\right).$$

If we further assume that σ is the maximally mixed state, i.e., $\sigma = \frac{1}{n}\mathbb{I}_n$, then

$$\Lambda = e^3, \qquad \eta = \frac{2e^{-3/2}}{1+e^3},$$

are independent of the dimension n.

In order to better illustrate the introduction of waiting time $\tau_{\alpha,\epsilon}$ and the sharpness of our decay rate $2\lambda_{\mathcal{L}}$ compared to $2K_\alpha$, we consider a special Lindblad equation below.

Example 4.8. Consider the following Lindblad equation

$$\dot{\rho}_t = \mathcal{L}^\dagger(\rho_t) = 2(\sigma_X \rho_t \sigma_X + \sigma_Z \rho_t \sigma_Z - 2\rho_t),$$

whose unique stationary state $\sigma = \mathbb{I}_2/2$. It could be verified that this Lindblad equation is primitive and satisfies GNS-detailed balance (see Sec. 4.2.2 below). As a remark, σ_X and σ_Z herein are Pauli matrices.

From Fig. 4.1, we observe the followings.

- $D_\alpha\left(\rho_t||\sigma\right)$ has the same asymptotic decay rate for various α, which verifies (4.6).

- It requires certain waiting time to reach the asymptotic decay rate, which justifies the introduction of $\tau_{\alpha,\epsilon}$ in Thm. 4.7.

- For $\alpha > 1$, $D_\alpha\left(\rho_t||\sigma\right)$ may decay slower initially, compared to its asymptotic behavior. Thus, in general, $K_\alpha < \lambda_{\mathcal{L}}$ is justified (cf. [107, Thm. 4.1]). Hence, log-Sobolev constants K_α could not capture the long-time behavior of sandwiched Rényi divergences, under the evolution of Lindblad equations.

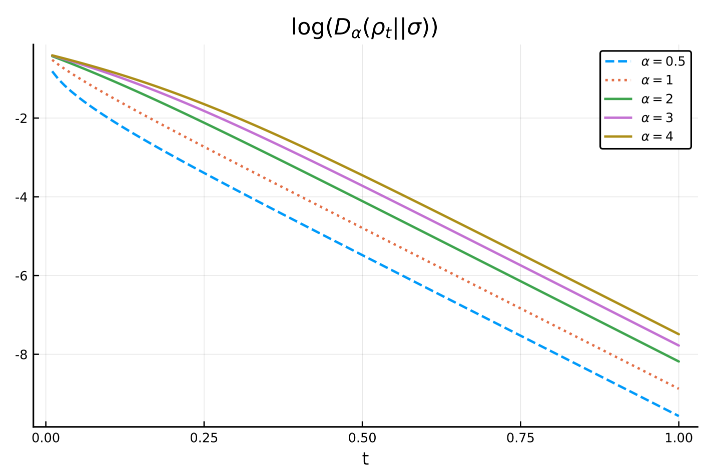


Figure 4.1: $\log(D_\alpha\left(\rho_t||\sigma\right))$ for various α with the same fixed initial condition ρ_0.

Our contribution

We summarize here contributions of this chapter.

Firstly, we extend the work of Carlen and Maas in [36] by proving that primitive Lindblad equations with GNS-detailed balance can be identified as gradient flow dynamics of the sandwiched Rényi divergence of any order $\alpha \in (0, \infty)$; see Thm. 4.2. Besides, we follow a recent work of Carlen and Maas [37] to study the necessary condition for Lindblad equations to be possibly written as gradient flow dynamics of sandwiched Rényi divergences; see Thm. 4.24.

Secondly, we prove that the sandwiched Rényi divergence decays exponentially fast under the evolution of primitive Lindblad equations with GNS-detailed balance; see Thm. 4.7. Our result (4.6) shows that the sandwiched Rényi divergence decays with an asymptotic rate $2\lambda_{\mathcal{L}}$, which is sharp according to [107, Thm. 4.1].

To prove Thm. 4.7, we first prove a quantum comparison theorem (see Prop. 4.32), which has its own interest. It is essentially a hypercontractivity-type estimation of the solution of the Lindblad equation, enabling us to bound $D_{\alpha_1}(\rho_T||\sigma)$ by $D_{\alpha_0}(\rho_0||\sigma)$ with some $T > 0$ when $\alpha_1 > \alpha_0$. As a direct consequence of the comparison theorem, Prop. 4.33 shows that the sandwiched Rényi divergence exponentially decays for all orders $\alpha \in (0, \infty)$. Such a comparison result for classical Rényi divergences was proved for the Fokker-Planck equation in [29, Thm. 1.2]. For quantum dynamics, a similar comparison result was also obtained in [14, Cor. 17]. However, a primary difference between our results and theirs is that we use K in the estimation of the waiting time, while they use κ_2 (see (4.17)). See a detailed discussion in Sec. 4.5.6.

The rest of this chapter is organized as follows. In Sec. 4.2, we will recall several preliminary results on, e.g., gradient flows, Lindblad equations with detailed balance, and noncommutative $\mathbb{L}_\alpha$ spaces. Sec. 4.3 will be devoted to the proof of Thm. 4.2 and Cor. 4.3. In Sec. 4.4, we will study the necessary detailed balance condition for Lindblad equations to be gradient flow dynamics of the sandwiched Rényi divergence of order $\alpha \in (0, \infty)$; see Thm. 4.24. Prop. 4.6 and Thm. 4.7 will be proved in Sec. 4.5. Finally, we will provide a concise summary in Sec. 4.6.

4.2 Preliminaries

In this section, we recall some technical notions: gradient flow dynamics on the Riemannian manifold, Lindblad equations with GNS-detailed balance, the quantum analogies of gradient and divergence operators, chain rule identity, and quantum noncommutative $\mathbb{L}_\alpha$ spaces. Most results from Sec. 4.2.2 to Sec. 4.2.4 follow from [36]; Sec. 4.2.5 is mostly based on [14].

Notations: We will use the following notations for this chapter here and below. We extend the inner product $\langle \cdot, \cdot \rangle$ to vector fields over $\mathfrak{M}$ by $\left\langle \vec{A}, \vec{B} \right\rangle := \sum_{j=1}^{\mathsf{J}} \langle A_j, B_j \rangle$, where two vector fields $\vec{A} = \left(A_1 \; A_2 \; \cdots A_{\mathsf{J}} \right)$, $\vec{B} = \left(B_1 \; B_2 \; \cdots B_{\mathsf{J}} \right)$, and $A_j, B_j \in \mathfrak{M}$ for all $1 \leq j \leq \mathsf{J}$; here J is a fixed positive integer. The space of such vector fields is denoted by $\mathfrak{M}^{\mathsf{J}}$. Such a convention is also applied analogously to all inner products on $\mathfrak{M}$.

Denote $\lambda_{\min}(\rho)$ as the smallest eigenvalue of a density matrix ρ and $\lambda_{\max}(\rho)$ as the largest one. To simplify the notation, a *weight operator* $\Upsilon_\sigma : \mathfrak{M} \to \mathfrak{M}$ is defined by

$$\Upsilon_\sigma(A) := \sigma^{\frac{1}{2}} A \sigma^{\frac{1}{2}},$$

for any matrix $A \in \mathfrak{M}$. Hence, for any $\gamma \in \mathbb{R}$,

$$\Upsilon_\sigma^\gamma(A) = \sigma^{\gamma/2} A \sigma^{\gamma/2}.$$

4.2.1 Gradient flow dynamics

Consider a Riemannian manifold $(\mathcal{M}, g_\rho)$: $\mathcal{M}$ is a smooth manifold, and $g_\rho(\cdot,\cdot)$ is an inner product on the tangent space $\mathcal{T}_\rho\mathcal{M}$ for any $\rho \in \mathcal{M}$. Given any (energy) functional E on $\mathcal{M}$, the gradient, denoted by $\mathrm{grad}E\big|_\rho$ at any state $\rho \in \mathcal{M}$, is defined as the element in the tangent space $\mathcal{T}_\rho\mathcal{M}$ such that

$$g_\rho\Big(\mathrm{grad}E\big|_\rho, \nu\Big) = \frac{\mathrm{d}}{\mathrm{d}\epsilon} E(\rho + \epsilon\nu)\Big|_{\epsilon=0}, \qquad \forall \nu \in \mathcal{T}_\rho\mathcal{M}. \tag{4.7}$$

Then, gradient flow dynamics refers to following autonomous differential equation

$$\dot{\rho}_t = -\mathrm{grad}E\big|_{\rho_t}. \tag{4.8}$$

This gradient flow dynamics is essentially a generalization of steepest descend dynamics in the Euclidean space. In this chapter, we will take $\mathcal{M}$ to be the space of strictly positive density matrices $\mathfrak{D}^+$. The choice of g_ρ is more subtle and technical, and we will explain it below in Sec. 4.3.

4.2.2 Quantum Markov semigroup with GNS-detailed balance

Recall the notion of quantum Markov semigroup (QMS) $\mathcal{P}_t$ from Chp. 1. To explain the quantum detailed balance, let us define inner products $\langle \cdot, \cdot \rangle_s$ parameterized by $s \in [0,1]$, for a given full-rank density matrix σ, in the following way: for any $A, B \in \mathfrak{M}$,

$$\langle A, B \rangle_s := \mathrm{Tr}\left(\sigma^s A^\dagger \sigma^{1-s} B\right).$$

Essentially, $\langle \cdot, \cdot \rangle_s$ is a σ-weighted Hilbert-Schmidt inner product. The QMS is said to satisfy *GNS-detailed balance* if $\mathcal{P}_t$ is self-adjoint with respect to the inner product $\langle \cdot, \cdot \rangle_1$ for all $t \geq 0$, that is to say, $\langle \mathcal{P}_t(A), B \rangle_1 = \langle A, \mathcal{P}_t(B) \rangle_1$ for all $A, B \in \mathfrak{M}$ and $t \geq 0$. In some literature, the case $s = 1/2$ is considered instead, and the QMS is said to satisfy the *KMS-detailed balance* if $\mathcal{P}_t$ is self-adjoint with respect to the inner product $\langle \cdot, \cdot \rangle_{1/2}$. It is not hard to verify that the QMS satisfies the KMS-detailed balance iff

$$\Upsilon_\sigma^{-1} \circ \mathcal{L}^\dagger \circ \Upsilon_\sigma = \mathcal{L}. \tag{4.9}$$

A notable relation is that the GNS-detailed balance is a more restrictive situation, and it implies the KMS-detailed balance; reversely, it is not valid. More discussion on the relation between GNS and KMS detailed balances could be found in [36], in particular, [36, Thm. 2.9]. Another important consequence is that if $\mathcal{P}_t$ satisfies detailed balance (either KMS or GNS), then σ is invariant under the dynamical evolution $\mathcal{P}_t^\dagger$, that is to say, $\mathcal{P}_t^\dagger(\sigma) = \sigma$. We add the prefix GNS or KMS to explicitly distinguish these two conditions. More results on the quantum detailed balance can be found in, e.g., [2, 68, 69, 134].

The generic form of quantum Markov semigroups with GNS-detailed balance is given by the theorem below.

Theorem 4.9 ([36, Thm. 3.1])**.** *Suppose that the quantum Markov semigroup $\mathcal{P}_t$ satisfies GNS-detailed balance, then*

$$\begin{aligned}\mathcal{L}(A) &= \sum_{j=1}^{\mathsf{J}} e^{-\omega_j/2} V_j^\dagger \left[A, V_j\right] + e^{\omega_j/2} \left[V_j, A\right] V_j^\dagger \\ &= \sum_{j=1}^{\mathsf{J}} e^{-\omega_j/2} \left(V_j^\dagger \left[A, V_j\right] + \left[V_j^\dagger, A\right] V_j\right),\end{aligned} \tag{4.10}$$

where $\omega_j \in \mathbb{R}$ and $\mathsf{J} \leq n^2 - 1$. Moreover $\{V_j\}_{j=1}^{\mathsf{J}}$ satisfy the following:

(i) $\mathrm{Tr}(V_j) = 0$ *and* $\langle V_j, V_k\rangle_{HS} = \delta_{j,k} c_j$ *for all* $1 \leq j, k \leq \mathsf{J}$*; constants* $c_j > 0$ *(for all* $1 \leq j \leq \mathsf{J}$*) come from the normalization;*

(ii) for each j, there exists $1 \leq j' \leq \mathsf{J}$ such that $V_j^\dagger = V_{j'}$;

(iii) $\Delta_\sigma(V_j) = e^{-\omega_j} V_j$*, where the modular operator* $\Delta_\sigma : \mathfrak{M} \to \mathfrak{M}$ *is defined as*

$$\Delta_\sigma(A) := \sigma A \sigma^{-1}, \qquad \textit{for } A \in \mathfrak{M};$$

(iv) if $V_j^\dagger = V_{j'}$ *then* $c_j = c_{j'}$ *and* $\omega_j = -\omega_{j'}$*.*

Then it could be easily derived that (see also [36, Eq. (3.7)])

$$\begin{aligned} \mathcal{L}^\dagger(A) &= \sum_{j=1}^{\mathsf{J}} \left(e^{-\omega_j/2} [V_j A, V_j^\dagger] + e^{\omega_j/2} [V_j^\dagger, A V_j] \right) \\ &= \sum_{j=1}^{\mathsf{J}} e^{-\omega_j/2} \left([V_j A, V_j^\dagger] + [V_j, A V_j^\dagger] \right), \end{aligned} \tag{4.11}$$

which is well known as the Lindblad super-operator; see also (1.1).

Example 4.10. Consider the depolarizing semigroup with the generator

$$\mathcal{L}_{\text{depol}}^\dagger(A) = \gamma(\sigma \operatorname{Tr}(A) - A), \qquad \gamma > 0.$$

It is not hard to verify that $\mathcal{L}_{\text{depol}}(A) = \gamma \left(\operatorname{Tr}(\sigma A) \mathbb{I}_n - A \right)$ and $\mathcal{L}_{\text{depol}}$ is self-adjoint with respect to the inner product $\langle \cdot, \cdot \rangle_1$. Thus, this depolarizing semigroup satisfies the GNS-detailed balance.

The following lemma is a special case of [36, Lem. 2.5] and is enough for our purpose.

Lemma 4.11. *Suppose QMS $\mathcal{P}_t$ satisfies the GNS-detailed balance, then*

$$[\mathcal{P}_t, \Delta_\sigma] = 0, \quad \textit{and equivalently} \quad [\mathcal{L}, \Delta_\sigma] = 0.$$

4.2.3 Noncommutative analog of gradient and divergence

Given those V_j from (4.10), define the *noncommutative partial derivative* ∂_j by

$$\partial_j(A) := [V_j, A],$$

for any matrix $A \in \mathfrak{M}$. As a result, its adjoint operator $\partial_j^\dagger$ has the form

$$\partial_j^\dagger(A) = [V_j^\dagger, A].$$

Let $\mathfrak{M}^{\mathsf{J}}$ be the vector fields over $\mathfrak{M}$, i.e., for $\vec{A} \in \mathfrak{M}^{\mathsf{J}}$, we have $\vec{A} = \left(\begin{smallmatrix} A_1 & A_2 & \cdots & A_{\mathsf{J}} \end{smallmatrix} \right)$, where $A_j \in \mathfrak{M}$ for all $1 \leq j \leq \mathsf{J}$. The *noncommutative gradient* $\nabla : \mathfrak{M} \to \mathfrak{M}^{\mathsf{J}}$ is defined by

$$\nabla A := \begin{pmatrix} \partial_1 A & \partial_2 A & \cdots & \partial_{\mathsf{J}} A \end{pmatrix} \in \mathfrak{M}^{\mathsf{J}}.$$

The *noncommutative divergence*, acting on $\mathfrak{M}^{\mathsf{J}}$, is defined by

$$\operatorname{div}(\vec{A}) := -\sum_{j=1}^{\mathsf{J}} \partial_j^\dagger(A_j) \equiv \sum_{j=1}^{\mathsf{J}} [A_j, V_j^\dagger], \tag{4.12}$$

for any $\vec{A} = \begin{pmatrix} A_1 & A_2 & \cdots A_{\mathsf{J}} \end{pmatrix} \in \mathfrak{M}^{\mathsf{J}}$. The divergence operator is defined in this way so that it is the adjoint of the gradient operator, with a multiplicative factor -1, analogous to the classical case.

Lemma 4.12 ([36, Eq. (5.3)])**.** *For a quantum Markov semigroup with GNS-detailed balance, we have*

$$\langle \nabla A, \nabla A \rangle_{1/2} = \langle A, -\mathcal{L}(A) \rangle_{1/2} . \tag{4.13}$$

Thus, $-\mathcal{L}$ is a positive semi-definite operator on the Hilbert space $(\mathfrak{M}, \langle \cdot, \cdot \rangle_{1/2})$.

Lemma 4.13 ([36, Thm. 5.3])**.** *Suppose the Lindblad equation satisfies GNS-detailed balance. Then the kernel of $\mathcal{L}$ equals to the commutant of $\{V_j\}_{j=1}^{\mathsf{J}}$. In other words, Ker$(\mathcal{L})$ = Ker(∇).*

Lem. 4.14 below gives a sufficient condition for Lindblad equations with GNS-detailed balance to be primitive, though the analysis in the rest of this chapter does not rely on this result. However, the condition $\mathsf{J} = n^2 - 1$ is not necessary for Lindblad equations with GNS-detailed balance to be primitive; see, e.g., Example 4.8 above.

Lemma 4.14. *If the Lindblad equation satisfies the GNS-detailed balance and $\mathsf{J} = n^2 - 1$, then the Lindblad equation is primitive.*

Proof. Note that $\{\mathbb{I}_n, V_1, V_2, \cdots V_{\mathsf{J}}\}$ form a (non-normalized) basis of $\mathfrak{M}$. Hence, for any matrix $A \in \mathfrak{M}$, one can decompose $A = \frac{\mathrm{Tr}(A)}{n}\mathbb{I}_n + \widetilde{A}$ where $\mathrm{Tr}(\widetilde{A}) = 0$. By the definition of primitive Lindblad equation and Lem. 4.13, it is equivalent to show that the commutant of $\{V_j\}_{j=1}^{\mathsf{J}}$ is spanned by $\mathbb{I}_n$, that is, if A is in the commutant of $\{V_j\}_{j=1}^{\mathsf{J}}$, then $\widetilde{A} = 0$.

If A is in the commutant, then $[\widetilde{A}, V_j] = 0$ for all j. Since $\{V_j\}_{j=1}^{\mathsf{J}}$ spans the space of traceless matrices, $0 = \left[\widetilde{A}, |j\rangle \langle k|\right]$ for all $j \neq k$, that is, $\widetilde{A} |j\rangle \langle k| = |j\rangle \langle k| \widetilde{A}$. Considering the matrix element $\langle a| \cdot |b\rangle$, we have $\widetilde{A}_{a,j}\delta_{k,b} = \widetilde{A}_{k,b}\delta_{a,j}$. Case (1): take $a = b = k$, then we know that $\widetilde{A}_{k,j} = 0$, and hence off-diagonal terms of $\widetilde{A}$ are all zero. Case (2): take $a = j$ and $b = k$, then we know that $\widetilde{A}_{j,j} = \widetilde{A}_{k,k}$. Since $0 = \mathrm{Tr}(\widetilde{A}) = \sum_j \widetilde{A}_{j,j}$, then $\widetilde{A}_{j,j} = 0$ for all j. In summary, $\widetilde{A} = 0$, thus completes the proof. □

4.2.4 Chain rule identity

One of the critical steps in connecting the Lindblad equation and the gradient flow dynamics of the quantum relative entropy is the following chain rule identity. In Sec. 4.3 below, this chain rule identity is also the key to proving Lem. 4.19.

Lemma 4.15 (Chain rule identity [36, Lem. 5.5])**.** *For any $V \in \mathfrak{M}$, $X \in \mathfrak{P}^+$ and $\omega \in \mathbb{R}$,*

$$[X]_\omega \left(V \log(e^{-\omega/2}X) - \log(e^{\omega/2}X)V\right) = e^{-\omega/2}VX - e^{\omega/2}XV, \tag{4.14}$$

where the operator $[X]_\omega : \mathfrak{M} \to \mathfrak{M}$ is defined by

$$[X]_\omega (A) := \int_0^1 e^{\omega(s-1/2)} X^s A X^{1-s} \, \mathrm{d}s. \tag{4.15}$$

This operator $[X]_\omega$ is a noncommutative multiplication of the operator X. For convenience, when $\omega = 0$, let $[X] \equiv [X]_0$. The operator $[X]_\omega$ can be extended and defined for vector fields $\mathfrak{M}^{\mathsf{J}}$. For $\vec{A} = \left(\begin{smallmatrix} A_1 & A_2 & \cdots & A_{\mathsf{J}} \end{smallmatrix} \right) \in \mathfrak{M}^{\mathsf{J}}$ and $\vec{\omega} = \left(\begin{smallmatrix} \omega_1 & \omega_2 & \cdots & \omega_{\mathsf{J}} \end{smallmatrix} \right) \in \mathbb{R}^{\mathsf{J}}$, define the operator $[X]_{\vec{\omega}} : \mathfrak{M}^{\mathsf{J}} \to \mathfrak{M}^{\mathsf{J}}$ by

$$[X]_{\vec{\omega}} (\vec{A}) := \begin{pmatrix} [X]_{\omega_1} (A_1) & [X]_{\omega_2} (A_2) & \cdots & [X]_{\omega_{\mathsf{J}}} (A_{\mathsf{J}}) \end{pmatrix}.$$

As a reminder, these ω_j with $1 \leq j \leq \mathsf{J}$ come from the spectrum of the modular operator Δ_σ (see Thm. 4.9 above).

Lemma 4.16. *If $X \in \mathfrak{P}^+$, then for any $\omega \in \mathbb{R}$, $[X]_\omega$ is a strictly positive operator on $\mathfrak{M}$. As a consequence, $[X]_\omega^{-1}$ is a well-defined strictly positive operator.*

Proof. Notice that

$$\begin{aligned} \langle A, [X]_\omega A \rangle &= \int_0^1 e^{\omega(s-1/2)} \operatorname{Tr}(A^\dagger X^s A X^{1-s}) \, \mathrm{d}s \\ &= \int_0^1 e^{\omega(s-1/2)} \left\langle X^{s/2} A X^{(1-s)/2}, X^{s/2} A X^{(1-s)/2} \right\rangle_{HS} \mathrm{d}s \geq 0. \end{aligned}$$

Moreover, $\langle A, [X]_\omega A \rangle_{HS} = 0$ iff $A = 0$. Thus, $[X]_\omega$ is a strictly positive operator on $\mathfrak{M}$ for any $\omega \in \mathbb{R}$. □

More explicitly, the inverse of $[X]_\omega$ is (see [36, Lem. 5.8])

$$\begin{aligned}[X]_\omega^{-1}(A) &= \int_0^\infty \frac{1}{t+e^{\omega/2}X} A \frac{1}{t+e^{-\omega/2}X}\ \mathrm{d}t \\ &\equiv \int_0^1 \frac{1}{1-s+se^{\omega/2}X} A \frac{1}{1-s+se^{-\omega/2}X}\ \mathrm{d}s,\end{aligned} \tag{4.16}$$

where we change the variable $s = \frac{1}{t+1}$ from the first to the second line above.

From (4.15) and (4.16), it is straightforward to verify the following Lemma.

Lemma 4.17 ([36, Lem. 5.8])**.** *For any $\omega \in \mathbb{R}$ and $A \in \mathfrak{M}$,*

$$([X]_\omega(A))^\dagger = [X]_{-\omega}(A^\dagger), \qquad ([X]_\omega^{-1}(A))^\dagger = [X]_{-\omega}^{-1}(A^\dagger).$$

4.2.5 Noncommutative $\mathbb{L}_\alpha$ spaces

In this subsection, we briefly recall some concepts from noncommutative $\mathbb{L}_\alpha$ spaces and log-Sobolev constants in that set-up [14, 89, 113]. We also recall an implication of the quantum Stroock-Varopoulos inequality [14], which compares the magnitudes of various log-Sobolev constants.

Let us first introduce the *weighted* $\mathbb{L}_\alpha$ *norm*, defined as

$$\|A\|_{\alpha,\sigma} := \Big(\operatorname{Tr}\left(|\Upsilon_\sigma^{1/\alpha}(A)|^\alpha\right) \Big)^{1/\alpha}.$$

Then, we need to introduce the *power operator*, for any $\alpha, \beta \in \mathbb{R}$, $A \in \mathfrak{M}$,

$$I_{\beta,\alpha}(A) := \Upsilon_\sigma^{-\frac{1}{\beta}}\left(|\Upsilon_\sigma^{\frac{1}{\alpha}}(A)|^{\frac{\alpha}{\beta}}\right).$$

When A is positive and commutes with σ, $I_{\beta,\alpha}(A) = A^{\alpha/\beta}$.

Two primary ingredients in noncommutative $\mathbb{L}_\alpha$ spaces are *entropy function* and *Dirichlet form*. The α-*Entropy function* for $X \in \mathfrak{P}^+$ is defined as

$$\begin{aligned}\mathrm{Ent}_{\alpha,\sigma}(X) :=& \operatorname{Tr}\left[\left(\Upsilon_\sigma^{1/\alpha}(X)\right)^\alpha \log\left(\left(\Upsilon_\sigma^{1/\alpha}(X)\right)^\alpha\right)\right] \\ &- \operatorname{Tr}\left[\left(\Upsilon_\sigma^{1/\alpha}(X)\right)^\alpha \log(\sigma)\right] - \|X\|_{\alpha,\sigma}^\alpha \log\left(\|X\|_{\alpha,\sigma}^\alpha\right),\end{aligned}$$

and the α-*Dirichlet form*, for $\alpha \in (0,1) \cup (1,\infty)$, is defined as

$$\mathcal{E}_{\alpha,\mathcal{L}}(X) := \frac{\alpha\widetilde{\alpha}}{4}\left\langle I_{\widetilde{\alpha},\alpha}(X), -\mathcal{L}(X)\right\rangle_{1/2},$$

where $1/\widetilde{\alpha} + 1/\alpha = 1$. When $\alpha = 1$, the Dirichlet form is defined by taking the limit $\alpha \to 1$, i.e.,

$$\mathcal{E}_{1,\mathcal{L}}(X) := \lim_{\alpha\to 1}\mathcal{E}_{\alpha,\mathcal{L}}(X) = \frac{1}{4}\left\langle \log(\Upsilon_\sigma(X)) - \log(\sigma), -\mathcal{L}(X)\right\rangle_{1/2}.$$

The α-*log Sobolev constant* in this set-up is defined as

$$\kappa_\alpha \equiv \kappa_\alpha(\mathcal{L}) := \inf_{X>0}\frac{\mathcal{E}_{\alpha,\mathcal{L}}(X)}{\mathrm{Ent}_{\alpha,\sigma}(X)}, \tag{4.17}$$

and thus the *quantum α-LSI* refers to the following

$$\kappa_\alpha \mathrm{Ent}_{\alpha,\sigma}(X) \leq \mathcal{E}_{\alpha,\mathcal{L}}(X), \qquad \forall X > 0. \tag{4.18}$$

In order to compare (4.4) and (4.18), for any matrix $X > 0$, let us introduce $\rho := \Upsilon_\sigma(X)/\operatorname{Tr}(\Upsilon_\sigma(X))$, which is a strictly positive density matrix. By inverting the operator Υ_σ, we have $X = \operatorname{Tr}(\Upsilon_\sigma(X))\,\Upsilon_\sigma^{-1}(\rho)$. By the fact that both $\mathcal{E}_{\alpha,\mathcal{L}}(X)$ and $\mathrm{Ent}_{\alpha,\sigma}(X)$ are homogeneous with respect to X (see, e.g., [14, Prop. 4 (ii) and Prop. 8 (ii)]), we have

$$\begin{aligned}\kappa_\alpha &= \inf_{\rho\in\mathfrak{D}^+} \frac{\mathcal{E}_{\alpha,\mathcal{L}}(\Upsilon_\sigma^{-1}(\rho))}{\mathrm{Ent}_{\alpha,\sigma}(\Upsilon_\sigma^{-1}(\rho))} \\ &= \inf_{\rho\in\mathfrak{D}^+} \frac{\frac{\alpha^2}{4(\alpha-1)}\left\langle \Upsilon_\sigma^{\frac{1-\alpha}{\alpha}}\left(\left(\Upsilon_\sigma^{\frac{1-\alpha}{\alpha}}(\rho)\right)^{\alpha-1}\right), -\mathcal{L}(\Upsilon_\sigma^{-1}(\rho))\right\rangle_{1/2}}{ZD\left(\left(\Upsilon_\sigma^{(1-\alpha)/\alpha}(\rho)\right)^\alpha/Z||\sigma\right)},\end{aligned}$$

where $Z = \operatorname{Tr}\left[\left(\Upsilon_\sigma^{(1-\alpha)/\alpha}(\rho)\right)^\alpha\right]$. Note that Z turns out to be the exponent in the sandwiched Rényi divergence (4.2). Provided that Lindblad equations satisfy KMS-detailed balance, by (4.9) and Lem. 4.19 (see below),

$$\begin{aligned}\kappa_\alpha &= \inf_{\rho\in\mathfrak{D}^+} \frac{\frac{\alpha^2}{4(\alpha-1)}\frac{1}{Z}\left\langle \Upsilon_\sigma^{\frac{1-\alpha}{\alpha}}\left(\left(\Upsilon_\sigma^{\frac{1-\alpha}{\alpha}}(\rho)\right)^{\alpha-1}\right), -\mathcal{L}^\dagger(\rho)\right\rangle_{HS}}{D\left(\left(\Upsilon_\sigma^{(1-\alpha)/\alpha}(\rho)\right)^\alpha/Z||\sigma\right)} \\ &= \inf_{\rho\in\mathfrak{D}^+} \frac{\frac{\alpha}{4}\left\langle \frac{\delta D_\alpha(\rho||\sigma)}{\delta\rho}, -\mathcal{L}^\dagger(\rho)\right\rangle_{HS}}{D\left(\left(\Upsilon_\sigma^{(1-\alpha)/\alpha}(\rho)\right)^\alpha/Z||\sigma\right)} = \inf_{\rho\in\mathfrak{D}^+} \frac{\frac{\alpha}{4}\mathscr{I}_\alpha(\rho||\sigma)}{D\left(\left(\Upsilon_\sigma^{(1-\alpha)/\alpha}(\rho)\right)^\alpha/Z||\sigma\right)}.\end{aligned} \tag{4.19}$$

As one might observe, the α-Dirichlet form becomes the quantum relative α-Fisher information up to a multiplicative constant, after the above transformation. However, the denominator in the above equation is different from the sandwiched Rényi divergence, which makes two definitions of the quantum α-LSI in (4.4) and (4.18) different. When $\alpha = 1$, this log-Sobolev constant κ_1 becomes

$$\kappa_1 = \frac{1}{4} \inf_{\rho \in \mathfrak{D}^+} \frac{\mathscr{I}(\rho||\sigma)}{D(\rho||\sigma)} = \frac{1}{4} 2K = \frac{K}{2}. \tag{4.20}$$

Therefore, both definitions (4.4) and (4.18) coincide when $\alpha = 1$, up to a multiplicative constant $1/2$.

The following implication of the quantum Stroock-Varopoulos inequality (see [14, Thm. 14]) characterizes the relationship between various κ_α.

Theorem 4.18 ([14, Cor. 16]). *Let us consider primitive Lindblad equations with GNS-detailed balance. Then κ_α is non-increasing with respect to $\alpha \in [0, 2]$. In particular,*

$$\kappa_1 \geq \kappa_2. \tag{4.21}$$

4.3 Lindblad equations as gradient flow dynamics of sandwiched Rényi divergences

This section devotes to the proof of Thm. 4.2. We will focus on the case $\alpha \neq 1$, since the case $\alpha = 1$ has been treated previously in [36]. Alternatively, one may easily adapt the proof below to the case $\alpha = 1$. The proof follows from a sequence of lemmas whose proofs are postponed to Sec. 4.3.2.

4.3.1 Proof of Thm. 4.2

To simplify notations, for any density matrix $\rho \in \mathfrak{D}$, define

$$\rho_\sigma := \Upsilon_\sigma^{\frac{1-\alpha}{\alpha}}(\rho) \equiv \sigma^{\frac{1-\alpha}{2\alpha}} \rho \sigma^{\frac{1-\alpha}{2\alpha}}.$$

Then the sandwiched Rényi divergence for $\alpha \neq 1$ can then be rewritten as

$$D_\alpha(\rho||\sigma) = \frac{1}{\alpha - 1} \log\Big(\operatorname{Tr}(\rho_\sigma^\alpha)\Big).$$

The following lemma, connecting the functional derivative of sandwiched Rényi divergence and the Lindblad equation, plays an essential role in the proof of Thm. 4.2.

Lemma 4.19 (Functional derivative)**.** *The functional derivative of the sandwiched Rényi divergence in the space $\mathfrak{D}$ is*

$$\frac{\delta D_\alpha(\rho||\sigma)}{\delta \rho} = \frac{\alpha}{\alpha - 1} \frac{\Upsilon_\sigma^{\frac{1-\alpha}{\alpha}}(\rho_\sigma^{\alpha-1})}{\operatorname{Tr}(\rho_\sigma^\alpha)}. \tag{4.22}$$

Moreover, for $\rho \in \mathfrak{D}^+$,

$$[\rho]_{\alpha,\omega_j} \partial_j\Big(\frac{\delta D_\alpha(\rho||\sigma)}{\delta\rho}\Big) = e^{-\omega_j/2} V_j \rho - e^{\omega_j/2} \rho V_j, \tag{4.23}$$

where the operator $[X]_{\alpha,\omega} : \mathfrak{M} \to \mathfrak{M}$, for positive $X \in \mathfrak{P}^+$ and $\omega \in \mathbb{R}$, is defined by

$$\begin{aligned}[X]_{\alpha,\omega} := \frac{\operatorname{Tr}\left(\left(\Upsilon_\sigma^{\frac{1-\alpha}{\alpha}}(X)\right)^\alpha\right)}{\alpha} \Upsilon_\sigma^{\frac{\alpha-1}{\alpha}} &\circ \left[\Upsilon_\sigma^{\frac{1-\alpha}{\alpha}}(X)\right]_{\omega/\alpha} \\ &\circ \left[\left(\Upsilon_\sigma^{\frac{1-\alpha}{\alpha}}(X)\right)^{\alpha-1}\right]^{-1}_{(\alpha-1)\omega/\alpha} \circ \Upsilon_\sigma^{\frac{\alpha-1}{\alpha}}.\end{aligned} \tag{4.24}$$

Remark. (i) When $\alpha = 1$, $\Upsilon_\sigma^{\frac{1-\alpha}{\alpha}} = \mathrm{Id}$ and $\rho_\sigma = \rho$. Then $[\rho]_{\alpha,\omega}$ reduces to $[\rho]_\omega$. Hence, $[\rho]_{\alpha,\omega}$ is the generalization of $[\rho]_\omega$ to the case of sandwiched Rényi divergences.

(ii) When $\alpha = 2$, the above operator $[\rho]_{2,\omega}$ has a simple form

$$[\rho]_{2,\omega} = \frac{\operatorname{Tr}(\rho_\sigma^2)}{2}\Upsilon_\sigma = \frac{\operatorname{Tr}(\sigma^{-1/2}\rho\sigma^{-1/2}\rho)}{2}\Upsilon_\sigma.$$

Thus, the mapping

$$(A,\rho) \to \left\langle A, [\rho]_{2,\omega}(A)\right\rangle_{HS} = \frac{\left\langle \rho, \sigma^{-1/2}\rho\sigma^{-1/2}\right\rangle_{HS}\left\langle A, \sigma^{1/2}A\sigma^{1/2}\right\rangle_{HS}}{2}$$

is the multiplication of quadratic terms with respect to both A and ρ.

The properties of the operator $[X]_\omega$ stated in Lem. 4.16 and Lem. 4.17 generalize to the operator $[X]_{\alpha,\omega}$, as we summarize in the next lemma.

Lemma 4.20. *If $X \in \mathfrak{P}^+$ and $\omega \in \mathbb{R}$, then*

(i) $[X]_{\alpha,\omega} : \mathfrak{M} \to \mathfrak{M}$ *is strictly positive;*

(ii) For any matrix A,

$$\Big([X]_{\alpha,\omega}(A)\Big)^\dagger = [X]_{\alpha,-\omega}(A^\dagger).$$

The operator $[X]_{\alpha,\omega}$ can be similarly extended to the space $\mathfrak{M}^{\mathsf{J}}$. For vector fields $\vec{A} = \left(\begin{smallmatrix} A_1 & A_2 & \cdots & A_{\mathsf{J}} \end{smallmatrix}\right) \in \mathfrak{M}^{\mathsf{J}}$ and $\vec{\omega} = \left(\begin{smallmatrix} \omega_1 & \omega_2 & \cdots & \omega_{\mathsf{J}} \end{smallmatrix}\right) \in \mathbb{R}^{\mathsf{J}}$, we define

$$[X]_{\alpha,\vec{\omega}}(\vec{A}) := \Big([X]_{\alpha,\omega_1}(A_1) \quad [X]_{\alpha,\omega_2}(A_2) \quad \cdots \quad [X]_{\alpha,\omega_{\mathsf{J}}}(A_{\mathsf{J}})\Big),$$

and analogously,

$$[X]^{-1}_{\alpha,\vec{\omega}}(\vec{A}) := \Big([X]^{-1}_{\alpha,\omega_1}(A_1) \quad [X]^{-1}_{\alpha,\omega_2}(A_2) \quad \cdots \quad [X]^{-1}_{\alpha,\omega_{\mathsf{J}}}(A_{\mathsf{J}})\Big).$$

The following lemma will be useful to characterize the tangent space of $\mathfrak{D}^+$.

Lemma 4.21. *Assume that the Lindblad equation* (4.1) *is primitive and $X \in \mathfrak{P}^+$. Denote the space of traceless matrices by $\mathfrak{M}_0$, and denote the space of traceless Hermitian matrices by $\mathfrak{A}_0$, that is, $\mathfrak{A}_0 = \mathfrak{M}_0 \cap \mathfrak{A}$.*

(i) Given $\vec{\omega} \in \mathbb{R}^{\mathsf{J}}$, define the operator $\mathcal{D}_{\alpha,X,\vec{\omega}} : \mathfrak{M}_0 \to \mathfrak{M}_0$ by

$$\mathcal{D}_{\alpha,X,\vec{\omega}}(A) := -\operatorname{div}\left([X]_{\alpha,\vec{\omega}}(\nabla A)\right).$$

This operator $\mathcal{D}_{\alpha,X,\vec{\omega}}$ is strictly positive, and thus invertible.

(ii) For any $A \in \mathfrak{M}_0$,

$$(\mathcal{D}_{\alpha,X,\vec{\omega}}(A))^\dagger = \mathcal{D}_{\alpha,X,\vec{\omega}}(A^\dagger),$$

and the restriction of $\mathcal{D}_{\alpha,X,\vec{\omega}}$ on $\mathfrak{A}_0$ is still strictly positive and invertible.

In the case $\alpha = 1$, this lemma has been proved in [51, Lem. 3]. Before defining the metric tensor, we need to first introduce an inner product, weighted by $[\rho]_{\alpha,\vec{\omega}}$.

Definition 4.22 (Inner product)**.** Given $\alpha \in (0,\infty)$, $\rho \in \mathfrak{D}^+$, for any two vector fields $\vec{A} = \begin{pmatrix} A_1 & A_2 & \cdots & A_{\mathsf{J}} \end{pmatrix}$ and $\vec{B} = \begin{pmatrix} B_1 & B_2 & \cdots & B_{\mathsf{J}} \end{pmatrix}$, an inner product on vector fields is defined as

$$\langle \vec{A}, \vec{B} \rangle_{\alpha,\rho,\mathcal{L}} := \sum_{j=1}^{\mathsf{J}} \left\langle A_j, [\rho]_{\alpha,\omega_j}(B_j) \right\rangle_{HS}.$$

Given $\rho \in \mathfrak{D}^+$, the tangent space is $\mathfrak{A}_0$. For ν_1 and $\nu_2 \in \mathfrak{A}_0$, by Lem. 4.21, there is a unique $U_k \in \mathfrak{A}_0$, such that

$$\nu_k = \mathcal{D}_{\alpha,\rho,\vec{\omega}}(U_k) \equiv -\operatorname{div}\left([\rho]_{\alpha,\vec{\omega}}(\nabla U_k)\right), \qquad k = 1, 2. \tag{4.25}$$

Finally, we are ready to define the metric tensor.

Definition 4.23 (Metric tensor)**.** Given $\rho \in \mathfrak{D}^+$, for $\nu_k \in \mathcal{T}_\rho \mathfrak{D}^+ \equiv \mathfrak{A}_0$, define a metric tensor $g_{\alpha,\rho,\mathcal{L}}$ by

$$g_{\alpha,\rho,\mathcal{L}}(\nu_1, \nu_2) := \langle \nabla U_1, \nabla U_2 \rangle_{\alpha,\rho,\mathcal{L}} \equiv \sum_{j=1}^{\mathsf{J}} \Big\langle \partial_j U_1, [\rho]_{\alpha,\omega_j} (\partial_j U_2) \Big\rangle_{HS}, \tag{4.26}$$

where U_k and ν_k are linked via (4.25).

Proof of Thm. 4.2. Up to now, the Riemannian structure $(\mathfrak{D}^+, g_{\alpha,\rho,\mathcal{L}})$ has been specified, as well as the energy functional $E(\rho) \equiv D_\alpha(\rho||\sigma)$. Then, we will verify that the gradient flow dynamics defined in (4.8) is exactly the Lindblad equation in (4.1). To achieve that, we need to compute the gradient at any $\rho_0 \in \mathfrak{D}^+$.

Recall the notion of gradient in the Riemannian manifold from (4.7): the gradient $\mathrm{grad}E\big|_{\rho_0}$ is defined as the traceless Hermitian matrix such that for any differentiable trajectory $\rho_t \in \mathfrak{D}^+$ with ρ_0 at time $t = 0$,

$$g_{\alpha,\rho_0,\mathcal{L}}\Big(\mathrm{grad}E\big|_{\rho_0}, \dot{\rho}_t\big|_{t=0}\Big) \equiv \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0} E(\rho_t) \equiv \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0} D_\alpha(\rho_t||\sigma).$$

By direct computation,

$$
\begin{aligned}
\left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} D_\alpha\left(\rho_t||\sigma\right) &= \left\langle \left.\frac{\delta D_\alpha\left(\rho||\sigma\right)}{\delta\rho}\right|_{\rho=\rho_0}, \dot{\rho}_t\big|_{t=0}\right\rangle_{HS} \\
&= \left\langle \left.\frac{\delta D_\alpha\left(\rho||\sigma\right)}{\delta\rho}\right|_{\rho=\rho_0}, -\operatorname{div}\Big(\left[\rho_0\right]_{\alpha,\vec{\omega}}\left(\nabla U_0\right)\Big)\right\rangle_{HS} \\
&= \left\langle \nabla\left.\frac{\delta D_\alpha\left(\rho||\sigma\right)}{\delta\rho}\right|_{\rho=\rho_0}, \left[\rho_0\right]_{\alpha,\vec{\omega}}\left(\nabla U_0\right)\right\rangle_{HS} \\
&= \left\langle \nabla\left.\frac{\delta D_\alpha\left(\rho||\sigma\right)}{\delta\rho}\right|_{\rho=\rho_0}, \nabla U_0\right\rangle_{\alpha,\rho_0,\mathcal{L}},
\end{aligned}
$$

where we used the following fact that follows from Lem. 4.21: there exists a unique family of traceless Hermitian matrices $\{U_t\}_{t\geq 0}$ such that $\dot{\rho}_t = -\operatorname{div}(\left[\rho_t\right]_{\alpha,\vec{\omega}}\nabla U_t)$. After combing the last two equations, we have

$$
g_{\alpha,\rho_0,\mathcal{L}}\left(\mathrm{grad}E\big|_{\rho_0}, \dot{\rho}_t\big|_{t=0}\right) = \left\langle \nabla\left.\frac{\delta D_\alpha\left(\rho||\sigma\right)}{\delta\rho}\right|_{\rho=\rho_0}, \nabla U_0\right\rangle_{\alpha,\rho_0,\mathcal{L}}.
$$

By the definition of the metric tensor $g_{\alpha,\rho,\mathcal{L}}(\cdot,\cdot)$ in (4.26),

$$
\begin{aligned}
\mathrm{grad}E\big|_{\rho_0} &= \mathcal{D}_{\alpha,\rho_0,\vec{\omega}}\Big(\left.\frac{\delta D_\alpha\left(\rho||\sigma\right)}{\delta\rho}\right|_{\rho=\rho_0}\Big) \\
&= -\operatorname{div}\left(\left[\rho_0\right]_{\alpha,\vec{\omega}}\left(\nabla\left.\frac{\delta D_\alpha\left(\rho||\sigma\right)}{\delta\rho}\right|_{\rho=\rho_0}\right)\right).
\end{aligned}
$$

Consequently, the gradient flow dynamics (4.8) is

$$
\dot{\rho}_t = -\mathrm{grad}E\big|_{\rho_t} = \operatorname{div}\left(\left[\rho_t\right]_{\alpha,\vec{\omega}}\left(\nabla\left.\frac{\delta D_\alpha\left(\rho||\sigma\right)}{\delta\rho}\right|_{\rho=\rho_t}\right)\right).
$$

Let us rewrite the term on the right hand side

$$
\begin{aligned}
&\operatorname{div}\left([\rho_t]_{\alpha,\vec{\omega}} \left(\nabla \frac{\delta D_\alpha(\rho||\sigma)}{\delta\rho}\bigg|_{\rho=\rho_t} \right)\right) \\
&\overset{(4.12)}{=} \sum_{j=1}^{\mathsf{J}} \left[[\rho_t]_{\alpha,\omega_j} \left(\partial_j \frac{\delta D_\alpha(\rho||\sigma)}{\delta\rho}\bigg|_{\rho=\rho_t} \right), V_j^\dagger \right] \\
&\overset{(4.23)}{=} \sum_{j=1}^{\mathsf{J}} \left[e^{-\omega_j/2} V_j \rho_t - e^{\omega_j/2} \rho_t V_j, V_j^\dagger \right] \\
&\overset{(4.11)}{=} \mathcal{L}^\dagger(\rho_t).
\end{aligned}
\tag{4.27}
$$

Thus, the gradient flow dynamics is exactly the Lindblad equation in (4.1). □

Proof of Cor. 4.3. If ρ_t is the solution of the Lindblad equation, then we know that $U_0 = -\frac{\delta D_\alpha(\rho||\sigma)}{\delta\rho}\Big|_{\rho=\rho_0}$. Hence,

$$
\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0} D_\alpha(\rho_t||\sigma) &= \left\langle \nabla \frac{\delta D_\alpha(\rho||\sigma)}{\delta\rho}\bigg|_{\rho=\rho_0}, \nabla U_0 \right\rangle_{\alpha,\rho_0,\mathcal{L}} \\
&= -\left\langle \nabla \frac{\delta D_\alpha(\rho||\sigma)}{\delta\rho}\bigg|_{\rho=\rho_0}, \nabla \frac{\delta D_\alpha(\rho||\sigma)}{\delta\rho}\bigg|_{\rho=\rho_0} \right\rangle_{\alpha,\rho_0,\mathcal{L}} \leq 0.
\end{aligned}
$$

□

4.3.2 Proof of Lemmas

Next, we present the proofs of lemmas used in the previous subsection.

Proof of Lem. 4.19. For any differentiable curve $\rho_t \in \mathfrak{D}$ passing through ρ at time $t = 0$, that is, $\rho_0 = \rho$, we have

$$\left\langle \frac{\delta D_\alpha(\rho||\sigma)}{\delta \rho}, \dot{\rho}_0 \right\rangle_{HS} \equiv \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} D_\alpha(\rho_t||\sigma) = \frac{\alpha}{\alpha-1} \frac{\mathrm{Tr}\left(\rho_\sigma^{\alpha-1} \sigma^{\frac{1-\alpha}{2\alpha}} \dot{\rho}_0 \, \sigma^{\frac{1-\alpha}{2\alpha}}\right)}{\mathrm{Tr}(\rho_\sigma^\alpha)}$$

$$= \frac{\alpha}{\alpha-1} \frac{\mathrm{Tr}\left(\sigma^{\frac{1-\alpha}{2\alpha}} \rho_\sigma^{\alpha-1} \sigma^{\frac{1-\alpha}{2\alpha}} \dot{\rho}_0\right)}{\mathrm{Tr}(\rho_\sigma^\alpha)}.$$

Then it is straightforward to obtain (4.22).

Next, we compute the partial derivative of $\frac{\delta D_\alpha(\rho||\sigma)}{\delta\rho}$. In fact,

$$\begin{aligned}
&\partial_j\Big(\frac{\delta D_\alpha(\rho||\sigma)}{\delta \rho}\Big) = \frac{\alpha}{(\alpha-1)\,\mathrm{Tr}(\rho_\sigma^\alpha)} \left[V_j, \sigma^{\frac{1-\alpha}{2\alpha}} \rho_\sigma^{\alpha-1} \sigma^{\frac{1-\alpha}{2\alpha}}\right] \\
&= \frac{\alpha}{(\alpha-1)\,\mathrm{Tr}(\rho_\sigma^\alpha)} \sigma^{\frac{1-\alpha}{2\alpha}} \left(\sigma^{\frac{\alpha-1}{2\alpha}} V_j \sigma^{\frac{1-\alpha}{2\alpha}} \rho_\sigma^{\alpha-1} - \rho_\sigma^{\alpha-1} \sigma^{\frac{1-\alpha}{2\alpha}} V_j \sigma^{\frac{\alpha-1}{2\alpha}}\right) \sigma^{\frac{1-\alpha}{2\alpha}} \\
&= \frac{\alpha}{(\alpha-1)\,\mathrm{Tr}(\rho_\sigma^\alpha)} \Upsilon_\sigma^{\frac{1-\alpha}{\alpha}} \left(e^{-\frac{\alpha-1}{2\alpha}\omega_j} V_j \rho_\sigma^{\alpha-1} - e^{\frac{\alpha-1}{2\alpha}\omega_j} \rho_\sigma^{\alpha-1} V_j\right) \\
&\overset{(4.14)}{=} \frac{\alpha}{(\alpha-1)\,\mathrm{Tr}(\rho_\sigma^\alpha)} \Upsilon_\sigma^{\frac{1-\alpha}{\alpha}} \Big(\left[\rho_\sigma^{\alpha-1}\right]_{\frac{\alpha-1}{\alpha}\omega_j} \left(V_j \log(e^{-\frac{\alpha-1}{2\alpha}\omega_j} \rho_\sigma^{\alpha-1}) - \log(e^{\frac{\alpha-1}{2\alpha}\omega_j} \rho_\sigma^{\alpha-1}) V_j\right)\Big) \\
&= \frac{\alpha}{\mathrm{Tr}(\rho_\sigma^\alpha)} \left(\Upsilon_\sigma^{\frac{1-\alpha}{\alpha}} \circ \left[\rho_\sigma^{\alpha-1}\right]_{\frac{\alpha-1}{\alpha}\omega_j}\right) \left(V_j \log(e^{-\frac{\omega_j}{2\alpha}} \rho_\sigma) - \log(e^{\frac{\omega_j}{2\alpha}} \rho_\sigma) V_j\right).
\end{aligned} \tag{4.28}$$

To get the fifth line, we have used (4.14) with $X = \rho_\sigma^{\alpha-1}$, $\omega = \frac{\alpha-1}{\alpha}\omega_j$, and $V = V_j$. Note that since $\rho, \sigma \in \mathfrak{D}^+$, both $\rho_\sigma, \rho_\sigma^{\alpha-1} \in \mathfrak{P}^+$.

Furthermore, applying (4.14) again with $X = \rho_\sigma$, $\omega = \omega_j/\alpha$ and $V = V_j$,

$$
\begin{aligned}
&[\rho_\sigma]_{\omega_j/\alpha} \left(V_j \log\!\left(e^{-\omega_j/(2\alpha)}\rho_\sigma\right) - \log\!\left(e^{\omega_j/(2\alpha)}\rho_\sigma\right) V_j\right) \\
&= e^{-\omega_j/(2\alpha)} V_j \rho_\sigma - e^{\omega_j/(2\alpha)} \rho_\sigma V_j \\
&= e^{-\omega_j/(2\alpha)} \sigma^{\frac{1-\alpha}{2\alpha}} \left(\sigma^{\frac{\alpha-1}{2\alpha}} V_j \sigma^{\frac{1-\alpha}{2\alpha}}\right) \rho \sigma^{\frac{1-\alpha}{2\alpha}} - e^{\omega_j/(2\alpha)} \sigma^{\frac{1-\alpha}{2\alpha}} \rho \left(\sigma^{\frac{1-\alpha}{2\alpha}} V_j \sigma^{\frac{\alpha-1}{2\alpha}}\right) \sigma^{\frac{1-\alpha}{2\alpha}} \\
&= \Upsilon_\sigma^{\frac{1-\alpha}{\alpha}} \left(e^{-\omega_j/2} V_j \rho - e^{\omega_j/2} \rho V_j\right).
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
e^{-\omega_j/2} V_j \rho - & e^{\omega_j/2} \rho V_j \\
&= \left(\Upsilon_\sigma^{\frac{\alpha-1}{\alpha}} \circ [\rho_\sigma]_{\omega_j/\alpha}\right) \left(V_j \log\!\left(e^{-\omega_j/(2\alpha)}\rho_\sigma\right) - \log\!\left(e^{\omega_j/(2\alpha)}\rho_\sigma\right) V_j\right).
\end{aligned}
$$

After plugging it back into (4.28), one could straightforwardly obtain (4.23), after arranging a few terms and by

$$
[\rho]_{\alpha,\omega_j}^{-1} := \frac{\alpha}{\operatorname{Tr}(\rho_\sigma^\alpha)} \left(\Upsilon_\sigma^{\frac{1-\alpha}{\alpha}} \circ \left[\rho_\sigma^{\alpha-1}\right]_{\frac{\alpha-1}{\alpha}\omega_j}\right) \circ \left(\Upsilon_\sigma^{\frac{\alpha-1}{\alpha}} \circ [\rho_\sigma]_{\omega_j/\alpha}\right)^{-1}.
$$

This becomes (4.24) after we invert this operator. $\square$

Proof of Lem. 4.20. (i) Since $X \in \mathfrak{P}^+$, $\sigma \in \mathfrak{D}^+$, then $X_\sigma := \Upsilon_\sigma^{\frac{1-\alpha}{\alpha}}(X)$ is strictly positive. For any $A \in \mathfrak{M}$, let $\widetilde{A} := \Upsilon_\sigma^{\frac{\alpha-1}{\alpha}}(A)$. Then we could rewrite the quadratic term

$\langle A, [X]_{\alpha,\omega}(A)\rangle_{HS}$ by

$$\begin{aligned}
&\langle A, [X]_{\alpha,\omega}(A)\rangle_{HS} \\
&\overset{(4.24)}{=} \frac{\mathrm{Tr}(X_\sigma^\alpha)}{\alpha} \left\langle A, \Upsilon_\sigma^{\frac{\alpha-1}{\alpha}} \circ [X_\sigma]_{\omega/\alpha} \circ \left[X_\sigma^{\alpha-1}\right]^{-1}_{(\alpha-1)\omega/\alpha} \circ \Upsilon_\sigma^{\frac{\alpha-1}{\alpha}}(A) \right\rangle_{HS} \\
&= \frac{\mathrm{Tr}(X_\sigma^\alpha)}{\alpha} \left\langle \widetilde{A}, [X_\sigma]_{\omega/\alpha} \circ \left[X_\sigma^{\alpha-1}\right]^{-1}_{(\alpha-1)\omega/\alpha}(\widetilde{A}) \right\rangle_{HS}.
\end{aligned}$$

Then by (4.15) and (4.16), we have

$$\begin{aligned}
&\langle A, [X]_{\alpha,\omega}(A)\rangle_{HS} \\
&\overset{(4.15)}{=} \frac{\mathrm{Tr}(X_\sigma^\alpha)}{\alpha} \int_0^1 e^{\frac{\omega}{\alpha}(s-\frac{1}{2})}\, \mathrm{Tr}\left(\widetilde{A}^\dagger X_\sigma^s \left(\left[X_\sigma^{\alpha-1}\right]^{-1}_{(\alpha-1)\omega/\alpha}(\widetilde{A})\right) X_\sigma^{1-s}\right) \mathrm{d}s \\
&= \frac{\mathrm{Tr}(X_\sigma^\alpha)}{\alpha} \int_0^1 e^{\frac{\omega}{\alpha}(s-\frac{1}{2})}\, \mathrm{Tr}\left(X_\sigma^{\frac{1-s}{2}} \widetilde{A}^\dagger X_\sigma^{\frac{s}{2}} X_\sigma^{\frac{s}{2}} \left(\left[X_\sigma^{\alpha-1}\right]^{-1}_{(\alpha-1)\omega/\alpha}(\widetilde{A})\right) X_\sigma^{\frac{1-s}{2}}\right) \mathrm{d}s \\
&\overset{(4.16)}{=} \frac{\mathrm{Tr}(X_\sigma^\alpha)}{\alpha} \int_0^1 e^{\frac{\omega}{\alpha}(s-\frac{1}{2})}\, \mathrm{Tr}\left(X_\sigma^{\frac{1-s}{2}} \widetilde{A}^\dagger X_\sigma^{\frac{s}{2}} \left(\left[X_\sigma^{\alpha-1}\right]^{-1}_{(\alpha-1)\omega/\alpha}(X_\sigma^{\frac{s}{2}} \widetilde{A} X_\sigma^{\frac{1-s}{2}})\right)\right) \mathrm{d}s \\
&= \frac{\mathrm{Tr}(X_\sigma^\alpha)}{\alpha} \int_0^1 e^{\frac{\omega}{\alpha}(s-\frac{1}{2})} \left\langle X_\sigma^{\frac{s}{2}} \widetilde{A} X_\sigma^{\frac{1-s}{2}}, \left[X_\sigma^{\alpha-1}\right]^{-1}_{(\alpha-1)\omega/\alpha}(X_\sigma^{\frac{s}{2}} \widetilde{A} X_\sigma^{\frac{1-s}{2}}) \right\rangle_{HS} \mathrm{d}s \geq 0,
\end{aligned}$$

since $[X_\sigma^{\alpha-1}]^{-1}_{(\alpha-1)\omega/\alpha}$ is a strictly positive operator by Lem. 4.16. Additionally,

$$\langle A, [X]_{\alpha,\omega}(A)\rangle_{HS} = 0 \iff 0 = X_\sigma^{\frac{s}{2}} \widetilde{A} X_\sigma^{\frac{1-s}{2}} \equiv X_\sigma^{\frac{s}{2}} \left(\Upsilon_\sigma^{\frac{\alpha-1}{\alpha}}(A)\right) X_\sigma^{\frac{1-s}{2}},$$

which implies that $A = 0$.

(ii) For any matrix A, by Lem. 4.17,

$$\begin{aligned}
\left([X]_{\alpha,\omega}(A)\right)^{\dagger} &= \frac{\operatorname{Tr}(X_\sigma^\alpha)}{\alpha}\Upsilon_\sigma^{\frac{\alpha-1}{\alpha}}\left([X_\sigma]_{\omega/\alpha}\circ\left[X_\sigma^{\alpha-1}\right]^{-1}_{(\alpha-1)\omega/\alpha}\circ\Upsilon_\sigma^{\frac{\alpha-1}{\alpha}}(A)\right)^{\dagger}\\
&= \frac{\operatorname{Tr}(X_\sigma^\alpha)}{\alpha}\Upsilon_\sigma^{\frac{\alpha-1}{\alpha}}\circ[X_\sigma]_{-\omega/\alpha}\left(\left[X_\sigma^{\alpha-1}\right]^{-1}_{(\alpha-1)\omega/\alpha}\circ\Upsilon_\sigma^{\frac{\alpha-1}{\alpha}}(A)\right)^{\dagger}\\
&= \frac{\operatorname{Tr}(X_\sigma^\alpha)}{\alpha}\Upsilon_\sigma^{\frac{\alpha-1}{\alpha}}\circ[X_\sigma]_{-\omega/\alpha}\circ\left[X_\sigma^{\alpha-1}\right]^{-1}_{-(\alpha-1)\omega/\alpha}\left(\Upsilon_\sigma^{\frac{\alpha-1}{\alpha}}(A)\right)^{\dagger}\\
&= \frac{\operatorname{Tr}(X_\sigma^\alpha)}{\alpha}\Upsilon_\sigma^{\frac{\alpha-1}{\alpha}}\circ[X_\sigma]_{-\omega/\alpha}\circ\left[X_\sigma^{\alpha-1}\right]^{-1}_{-(\alpha-1)\omega/\alpha}\circ\Upsilon_\sigma^{\frac{\alpha-1}{\alpha}}(A^{\dagger})\\
&= [X]_{\alpha,-\omega}(A^{\dagger}).
\end{aligned}$$

□

Proof of Lem. 4.21. (i) Consider a traceless matrix $A \in \mathfrak{M}_0$. It is easy to verify that $\operatorname{Tr}(\mathcal{D}_{\alpha,X,\vec{\omega}}(A)) = 0$. From the definition of the divergence operator (4.12), we have

$$\langle A, \mathcal{D}_{\alpha,X,\vec{\omega}}(A)\rangle_{HS} = \sum_{j=1}^{\mathsf{J}}\left\langle \partial_j A, [X]_{\alpha,\omega_j}(\partial_j A)\right\rangle_{HS} \geq 0.$$

Moreover, since $[X]_{\alpha,\omega_j}$ is strictly positive by Lem. 4.20, we know that

$$\langle A, \mathcal{D}_{\alpha,X,\vec{\omega}}(A)\rangle_{HS} = 0 \text{ iff } \partial_j A = 0, \forall 1 \leq j \leq \mathsf{J}.$$

Thanks to Lem. 4.13 and the assumption that the Lindblad equation (4.1) is primitive, we have $A = 0$ (note that A is assumed to be traceless). Therefore, $\mathcal{D}_{\alpha,X,\vec{\omega}}$ is a strictly positive operator on the space of traceless matrices $\mathfrak{M}_0$.

(ii) We only need to prove that $(\mathcal{D}_{\alpha,X,\vec{\omega}}(A))^\dagger = \mathcal{D}_{\alpha,X,\vec{\omega}}(A^\dagger)$. The rest simply follows from the first part of the lemma. Note that

$$\begin{aligned}\mathcal{D}_{\alpha,X,\vec{\omega}}(A) &= -\operatorname{div}\left([X]_{\alpha,\vec{\omega}}(\nabla A)\right) \overset{(4.12)}{=} \sum_j \left[V_j^\dagger, [X]_{\alpha,\omega_j}(\partial_j A)\right] \\ &= \sum_j V_j^\dagger \left([X]_{\alpha,\omega_j}(V_j A - A V_j)\right) - \left([X]_{\alpha,\omega_j}(V_j A - A V_j)\right) V_j^\dagger.\end{aligned}$$

By Thm. 4.9 and Lem. 4.20, we have

$$\begin{aligned}&(\mathcal{D}_{\alpha,X,\vec{\omega}}(A))^\dagger \\ &= \sum_j \left([X]_{\alpha,-\omega_j}(A^\dagger V_j^\dagger - V_j^\dagger A^\dagger)\right) V_j - V_j \left([X]_{\alpha,-\omega_j}(A^\dagger V_j^\dagger - V_j^\dagger A^\dagger)\right) \\ &= \sum_{j'} \left([X]_{\alpha,\omega_{j'}}(A^\dagger V_{j'} - V_{j'} A^\dagger)\right) V_{j'}^\dagger - V_{j'}^\dagger \left([X]_{\alpha,\omega_{j'}}(A^\dagger V_{j'} - V_{j'} A^\dagger)\right) \\ &= \mathcal{D}_{\alpha,X,\vec{\omega}}(A^\dagger).\end{aligned}$$

□

4.4 The necessary condition for Lindblad equations to be gradient flow dynamics of sandwiched Rényi divergences

Recall from Thm. 4.2 that Lindblad equations with GNS-detailed balance can be regarded as gradient flow dynamics of sandwiched Rényi divergences, including the quantum relative entropy. A natural and immediate follow-up question is the extent that one can generalize Thm. 4.2 in the direction of considering a larger family of Lindblad equations. In a recent paper [37], such an issue has been briefly addressed. However, currently, there is still a gap between the class of Lindblad equations (i.e., primitive Lindblad equations with GNS-detailed balance) that are known to be gradient flow dynamics of the quantum relative entropy and the necessary condition (i.e., BKM-detailed balance condition) for a primitive Lindblad equation to be possibly expressed as gradient flow dynamics of the quantum relative entropy. This gap will be revisited and explained in more detail below.

In this section, we shall explore the necessary condition for Lindblad equations to be gradient flow dynamics of sandwiched Rényi divergences, summarized in Thm. 4.24, which adapts the argument from [37, Thm 2.9]. Next, in Sec. 4.4.2, we will discuss the detailed balance conditions arising from Thm. 4.24. Finally, in Sec. 4.4.3, we will discuss relations between various detailed balance conditions.

4.4.1 Necessary condition

Below is the main result of this section. Recall the notation $[X]_\omega$ (4.15) and recall that $[X] \equiv [X]_0$.

Theorem 4.24. *Suppose that the dual quantum Markov semigroup $\mathcal{P}_t^\dagger = e^{t\mathcal{L}^\dagger}$ is primitive, and its unique stationary state is $\sigma \in \mathfrak{D}^+$. If there exists a continuously differentiable metric tensor $g_\rho(\cdot,\cdot) : \mathfrak{A}_0 \times \mathfrak{A}_0 \to \mathbb{R}$ for any $\rho \in \mathfrak{D}^+$ such that the Lindblad equation $\dot{\rho}_t = \mathcal{L}^\dagger(\rho_t)$ is gradient flow dynamics of the sandwiched Rényi divergence $D_\alpha(\rho||\sigma)$, then $\mathcal{L}$ is self-adjoint with respect to the inner product weighted by an operator $\mathscr{W}_{\sigma,\alpha} : \mathfrak{M} \to \mathfrak{M}$ defined by*

$$\begin{aligned}\mathscr{W}_{\sigma,\alpha}(A) &:= \left[\sigma^{\frac{1}{\alpha}}\right] \circ \left[\sigma^{\frac{\alpha-1}{\alpha}}\right]^{-1} \circ \Upsilon_\sigma^{\frac{2(\alpha-1)}{\alpha}}(A) \\ &\equiv \int_0^1 \int_0^1 \Big(\frac{\left(\sigma^{\frac{1}{\alpha}}\right)^{1-s} \sigma^{\frac{\alpha-1}{\alpha}}}{(1-r)\mathbb{I}_n + r\sigma^{\frac{\alpha-1}{\alpha}}}\Big) A \Big(\frac{\sigma^{\frac{\alpha-1}{\alpha}} \left(\sigma^{\frac{1}{\alpha}}\right)^s}{(1-r)\mathbb{I}_n + r\sigma^{\frac{\alpha-1}{\alpha}}}\Big)\, \mathrm{d}r\, \mathrm{d}s,\end{aligned} \tag{4.29}$$

for all $A \in \mathfrak{M}$. More specifically, we have for all $A, B \in \mathfrak{M}$,

$$\langle \mathcal{L}(A), B\rangle_{\mathscr{W}_{\sigma,\alpha}} = \langle A, \mathcal{L}(B)\rangle_{\mathscr{W}_{\sigma,\alpha}}, \tag{4.30}$$

where $\langle A, B\rangle_{\mathscr{W}_{\sigma,\alpha}} := \langle A, \mathscr{W}_{\sigma,\alpha}(B)\rangle_{HS}$.

Remark. The operator $\mathscr{W}_{\sigma,\alpha}$ is a noncommutative way to multiply σ and the condition (4.30) is equivalent to

$$\mathscr{W}_{\sigma,\alpha} \circ \mathcal{L} \circ \mathscr{W}_{\sigma,\alpha}^{-1} = \mathcal{L}^\dagger. \tag{4.31}$$

Motivated by the definition of GNS and KMS detailed balance conditions, we shall define a detailed balance condition based on sandwiched Rényi divergences, for the convenience of the discussion below.

Definition 4.25 (Sandwiched Rényi divergence (SRD) detailed balance condition)**.** For a primitive Lindblad equation with a unique stationary state $\sigma \in \mathfrak{D}^+$, it is said to satisfy the *sandwiched Rényi divergence (SRD) detailed balance condition* if the generator $\mathcal{L}$ is self-adjoint with respect to the inner product $\langle \cdot, \cdot \rangle_{\mathscr{W}_{\sigma,\alpha}}$ for *any* $\alpha \in (0, \infty)$.

Proof of Thm. 4.24. We first need to introduce some notations following [37, Sec. 2.3]. Define the operator $\mathscr{G}_\rho : \mathfrak{A}_0 \to \mathfrak{A}_0$ by $\langle A, \mathscr{G}_\rho(B) \rangle_{HS} = g_\rho(A, B)$. Since g_ρ is a $\mathbb{R}$-valued inner product, $\mathscr{G}_\rho$ must be invertible and self-adjoint. The inverse of $\mathscr{G}_\rho$ is denoted by $\mathscr{K}_\rho$, which is also self-adjoint. From (4.7), one could easily derive that the gradient flow dynamics can be expressed as

$$\dot{\rho}_t = -\mathscr{K}_{\rho_t}\Big(\frac{\delta D_\alpha(\rho||\sigma)}{\delta \rho}\Big|_{\rho=\rho_t}\Big). \tag{4.32}$$

For convenience, notice that the domain of $\mathscr{K}_\rho$ can be extended from $\mathfrak{A}_0$ to $\mathfrak{A}$ by setting $\mathscr{K}_\rho(\mathbb{I}_n) := 0$.

Under the assumption that $\dot{\rho}_t = \mathcal{L}^\dagger(\rho_t)$ is gradient flow dynamics of the sandwiched Rényi divergence $D_\alpha(\rho||\sigma)$ and by (4.32),

$$\mathcal{L}^\dagger(\rho) = -\mathscr{K}_\rho\left(\frac{\delta D_\alpha(\rho||\sigma)}{\delta\rho}\right).$$

Choose $\rho_\epsilon = \sigma + \epsilon A$ with $A \in \mathfrak{A}_0$ and let $\epsilon > 0$ small enough such that $\rho_\epsilon \in \mathfrak{D}^+$. Then

$$\begin{aligned}
\mathcal{L}^\dagger(A) &= \frac{\mathcal{L}^\dagger(\rho_\epsilon) - \mathcal{L}^\dagger(\sigma)}{\epsilon} = \lim_{\epsilon\downarrow 0^+}\frac{\mathrm{d}}{\mathrm{d}\epsilon}\mathcal{L}^\dagger(\rho_\epsilon) \\
&= -\lim_{\epsilon\downarrow 0^+}\frac{\mathrm{d}}{\mathrm{d}\epsilon}\mathscr{K}_{\rho_\epsilon}\Big(\frac{\delta D_\alpha(\rho||\sigma)}{\delta\rho}\Big|_{\rho=\rho_\epsilon}\Big) \\
&\overset{(4.22)}{=} -\frac{\alpha}{\alpha-1}\lim_{\epsilon\downarrow 0^+}\frac{\mathrm{d}}{\mathrm{d}\epsilon}\left(\Big(\mathscr{K}_\sigma + \epsilon(\delta\mathscr{K}_\sigma) + \mathcal{O}(\epsilon^2)\Big)\Big(\frac{\Upsilon_\sigma^{\frac{1-\alpha}{\alpha}}(\rho_\sigma^{\alpha-1})}{\mathrm{Tr}\,(\rho_\sigma^\alpha)}\Big|_{\rho=\rho_\epsilon}\Big)\right) \\
&= -\frac{\alpha}{\alpha-1}\lim_{\epsilon\downarrow 0^+}\frac{\mathrm{d}}{\mathrm{d}\epsilon}\left(\Big(\mathscr{K}_\sigma + \epsilon(\delta\mathscr{K}_\sigma) + \mathcal{O}(\epsilon^2)\Big)\Big(\frac{\mathbb{I}_n + \epsilon B + \mathcal{O}(\epsilon^2)}{1+\mathcal{O}(\epsilon)}\Big)\right) \\
&= -\frac{\alpha}{\alpha-1}\mathscr{K}_\sigma(B),
\end{aligned}$$

where the term B in the expansion is (see, e.g., [35, Prop. 7.2])

$$\begin{aligned}
B &= \frac{\mathrm{d}}{\mathrm{d}\epsilon}\Big|_{\epsilon=0}\Upsilon_\sigma^{\frac{1-\alpha}{\alpha}}\left(\left(\Upsilon_\sigma^{\frac{1-\alpha}{\alpha}}(\rho_\epsilon)\right)^{\alpha-1}\right) \\
&= \Upsilon_\sigma^{\frac{1-\alpha}{\alpha}}\left(\int_0^1\int_0^{\alpha-1}\frac{\sigma^{\frac{\alpha-1-\beta}{\alpha}}}{(1-s)\mathbb{I}_n + s\sigma^{\frac{1}{\alpha}}}\Upsilon_\sigma^{\frac{1-\alpha}{\alpha}}(A)\frac{\sigma^{\frac{\beta}{\alpha}}}{(1-s)\mathbb{I}_n + s\sigma^{\frac{1}{\alpha}}}\,\mathrm{d}\beta\,\mathrm{d}s\right) \\
&= (\alpha-1)\int_0^1\int_0^1\frac{\left(\sigma^{\frac{\alpha-1}{\alpha}}\right)^{1-\beta}}{(1-s)\mathbb{I}_n + s\sigma^{\frac{1}{\alpha}}}\Upsilon_\sigma^{\frac{2(1-\alpha)}{\alpha}}(A)\frac{\left(\sigma^{\frac{\alpha-1}{\alpha}}\right)^{\beta}}{(1-s)\mathbb{I}_n + s\sigma^{\frac{1}{\alpha}}}\,\mathrm{d}\beta\,\mathrm{d}s \\
&= (\alpha-1)\left[\sigma^{\frac{\alpha-1}{\alpha}}\right]\circ\left[\sigma^{\frac{1}{\alpha}}\right]^{-1}\circ\Upsilon_\sigma^{\frac{2(1-\alpha)}{\alpha}}(A) = (\alpha-1)\mathscr{W}_{\sigma,\alpha}^{-1}(A).
\end{aligned}$$

To get the third line, we change the variable β by $(\alpha-1)\beta$; in the fourth line, we use expressions (4.15) and (4.16). Note that all three operators in the fourth line above pairwisely commute. By combining the last two equations,

$$\mathcal{L}^\dagger(A) = -\alpha \mathscr{K}_\sigma\left(\mathscr{W}_{\sigma,\alpha}^{-1}(A)\right).$$

Therefore, for any $C, D \in \mathfrak{A}_0$,

$$\langle \mathcal{L}(C), D\rangle_{\mathscr{W}_{\sigma,\alpha}} = \left\langle C, \mathcal{L}^\dagger \circ \mathscr{W}_{\sigma,\alpha}(D)\right\rangle_{HS} = -\alpha \left\langle C, \mathscr{K}_\sigma(D)\right\rangle_{HS},$$

and similarly,

$$\langle C, \mathcal{L}(D)\rangle_{\mathscr{W}_{\sigma,\alpha}} = (-\alpha)\left\langle \mathscr{K}_\sigma(C), D\right\rangle_{HS}.$$

As mentioned earlier at the beginning of this proof, the operator $\mathscr{K}_\sigma$ is self-adjoint. Thus, $\mathcal{L}$ is self-adjoint with respect to the inner product $\langle\cdot,\cdot\rangle_{\mathscr{W}_{\sigma,\alpha}}$. $\square$

4.4.2 Decomposition of the operator $\mathscr{W}_{\sigma,\alpha}$

To understand better the operator $\mathscr{W}_{\sigma,\alpha}$, we study its decomposition and various special cases. Suppose the eigenvalue decomposition of σ is $\sigma = \sum_k \lambda_k |\psi_k\rangle\langle\psi_k|$ where $\lambda_k > 0$ for all k. When $\alpha \neq 1$, the expansion of (4.29) in terms of eigenbasis of σ leads into

$$\mathscr{W}_{\sigma,\alpha}(\cdot) = \sum_{k,j} f_{k,j}^\alpha |\psi_k\rangle\langle\psi_k| \cdot |\psi_j\rangle\langle\psi_j|,$$

where the coefficients $f^{\alpha}_{k,j}$ are given by

$$f^{\alpha}_{k,j} = \begin{cases} \lambda_k, & \text{when } \lambda_k = \lambda_j; \\ (\alpha - 1)\dfrac{\lambda_k^{\frac{1}{\alpha}} - \lambda_j^{\frac{1}{\alpha}}}{\lambda_j^{\frac{1-\alpha}{\alpha}} - \lambda_k^{\frac{1-\alpha}{\alpha}}}, & \text{when } \lambda_k \neq \lambda_j. \end{cases}$$

Let us consider a few special weight operators $\mathscr{W}_{\sigma,\alpha}$:

- ($\alpha = 1$). The operator $\mathscr{W}_{\sigma,\alpha}$ reduces to $[\sigma](\cdot) = \int_0^1 \sigma^{1-s} \cdot \sigma^s \,\mathrm{d}s$, and the inner product $\langle \cdot, \cdot \rangle_{\mathscr{W}_{\sigma,\alpha}}$ is known as the *BKM inner product*. As a remark, this case has been shown in [37, Thm. 2.9].

- ($\alpha = 2$). The operator $\mathscr{W}_{\sigma,\alpha}$ reduces to Υ_σ and the inner product $\langle \cdot, \cdot \rangle_{\mathscr{W}_{\sigma,\alpha}}$ is the KMS inner product. This immediately implies that in the context of gradient flow of sandwiched Rényi divergences, one cannot work on a class of Lindblad equations larger than the one with KMS-detailed balance condition.

Even though $\alpha \in (0, \infty)$ in Def. 4.25, we can still apply the limiting argument to define $\mathscr{W}_{\sigma,\infty}$ and $\mathscr{W}_{\sigma,0}$.

- ($\alpha = \infty$). As $\alpha \to \infty$, $\sigma^{\frac{1}{\alpha}} \to \mathbb{I}_n$ and $\sigma^{\frac{\alpha-1}{\alpha}} \to \sigma$. As a result, the operator $\mathscr{W}_{\sigma,\alpha}$ converges to $\mathscr{W}_{\sigma,\infty} := [\sigma]^{-1} \circ \Upsilon^2_\sigma$.

- ($\alpha = 0$). This case is slightly more subtle, since $\sigma^{\frac{\alpha-1}{\alpha}}$ will blow up. When $\lambda_k \neq \lambda_j$, as $\alpha \downarrow 0$, one could show that $f^{\alpha}_{k,j} \to \max(\lambda_k, \lambda_j)$. Therefore,

$$\mathscr{W}_{\sigma,0}(\cdot) = \sum_{k,j} \max(\lambda_k, \lambda_j) \left|\psi_k\right\rangle \left\langle\psi_k\right| \cdot \left|\psi_j\right\rangle \left\langle\psi_j\right| .$$

4.4.3 Relations between GNS, SRD, KMS detailed balance conditions

Proposition 4.26. *Consider any primitive Lindblad equation whose unique stationary state $\sigma \in \mathfrak{D}^+$:*

(i) GNS-detailed balance condition implies SRD-detailed balance condition.

(ii) SRD-detailed balance condition implies KMS-detailed balance condition, while conversely it is generally not true.

Remark. (i) This suggests that the class of Lindblad equations with GNS-detailed balance considered in Thm. 4.2 seems to be general enough, because it is impossible to generalize Thm. 4.2 to the class of Lindblad equations with KMS-detailed balance by Prop. 4.26. More specifically, as we shall show later in the proof of Prop. 4.26, there exists at least one primitive Lindblad equation that satisfies KMS-detailed balance condition, but does not satisfy SRD-detailed balance condition (in particular, $\mathscr{W}_{\sigma,\alpha} \circ \mathcal{L} \circ \mathscr{W}^{-1}_{\sigma,\alpha} \neq \mathcal{L}^{\dagger}$ for any $\alpha \neq 2$). Because SRD-detailed balance condition is necessary for having a gradient flow structure, this particular Lindblad equation cannot be expressed as gradient flow dynamics of $D_{\alpha}\left(\rho||\sigma\right)$ for any $\alpha \neq 2$.

(ii) It is still unknown whether the SRD-detailed balance condition is equivalent to the GNS-detailed balance condition or not. Thus, it might be interesting to characterize Lindblad equations with SRD-detailed balance, though this task seems to be somewhat technical.

Proof of Prop. 4.26. The first statement comes immediately by combining both Thm. 4.2 and Thm. 4.24. As for the second statement, since the SRD-detailed balance condition contains KMS-detailed balance condition as a particular instance when $\alpha = 2$, obviously SRD implies KMS. In the following, we provide a primitive two-level Lindblad equation with KMS-detailed balance from [36, App. B] to demonstrate that KMS-detailed balance condition does not imply SRD-detailed balance condition.

Let us consider

$$K_1 = |\psi\rangle \langle 0| , \qquad |\psi\rangle = \frac{1}{\sqrt{2}} (|0\rangle + |1\rangle) ;$$
$$K_2 = |\phi\rangle \langle 1| , \qquad |\phi\rangle = \frac{1}{\sqrt{5}} (|0\rangle + 2\,|1\rangle) .$$

- Define an operator $\mathscr{K}$ by $\mathscr{K}(A) := K_1^\dagger A K_1 + K_2^\dagger A K_2$. Thus, $\mathscr{K}$ is completely positive, and $\mathscr{K}(\mathbb{I}_2) = \mathbb{I}_2$.

- The adjoint operator

$$\mathscr{K}^\dagger(A) = K_1 A K_1^\dagger + K_2 A K_2^\dagger = |\psi\rangle \langle 0| A |0\rangle \langle \psi| + |\phi\rangle \langle 1| A |1\rangle \langle \phi|$$

is also completely positive. Note that

$$\mathscr{K}^\dagger(|0\rangle\langle 1|) = \mathscr{K}^\dagger(|1\rangle\langle 0|) = 0,$$

$$\mathscr{K}^\dagger(|0\rangle\langle 0|) = |\psi\rangle\langle\psi|,$$

$$\mathscr{K}^\dagger(|1\rangle\langle 1|) = |\phi\rangle\langle\phi|.$$

It can be readily verified that there is one and only one eigenstate for $\mathscr{K}^\dagger$ associated with eigenvalue 1, which is

$$\sigma = \frac{1}{7}\begin{pmatrix} 2 & 3 \\ 3 & 5 \end{pmatrix}.$$

The other non-zero eigenvalue of $\mathscr{K}^\dagger$ is $3/10$.

- Define $\widetilde{K}_j = \sigma^{1/2} K_j^\dagger \sigma^{-1/2}$ for $j \in \{1, 2\}$, and define a completely positive operator $\widetilde{\mathscr{K}}$ via $\widetilde{\mathscr{K}}(A) = \widetilde{K}_1^\dagger A \widetilde{K}_1 + \widetilde{K}_2^\dagger A \widetilde{K}_2$. Then $\widetilde{\mathscr{K}}(\mathbb{I}_2) = \mathbb{I}_2$ and $\widetilde{\mathscr{K}}^\dagger(\sigma) = \sigma$. Furthermore, we can verify that

$$\left\langle \widetilde{\mathscr{K}} B, A \right\rangle_{1/2} = \langle B, \mathscr{K} A \rangle_{1/2}.$$

Hence, the QMS $\mathcal{P}_t = e^{t\mathcal{L}}$ with $\mathcal{L} := \widetilde{\mathscr{K}}\mathscr{K} - \mathrm{Id}_2$ can be readily verified to satisfy the KMS-detailed balance condition. Moreover, it is easy to show $\mathcal{L}^\dagger(\sigma) = 0$, and one could verify that it is the only eigenvector of $\mathcal{L}^\dagger$ with eigenvalue 0, i.e., the Lindblad equation is primitive.

Next, we numerically show that (4.31) does not hold. Fig. 4.2 plots the trace-norm of $\mathscr{W}_{\sigma,\alpha} \circ \mathcal{L} \circ \mathscr{W}_{\sigma,\alpha}^{-1} - \mathcal{L}^\dagger$ for various α. From this figure, it is clear that (4.31) holds only when $\alpha = 2$ (i.e., KMS-detailed balance condition). $\square$

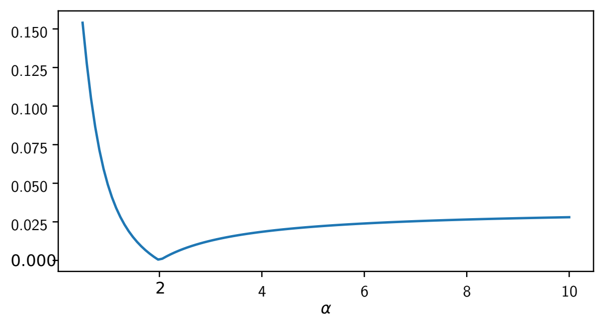


Figure 4.2: $\|\mathscr{W}_{\sigma,\alpha} \circ \mathcal{L} \circ \mathscr{W}_{\sigma,\alpha}^{-1} - \mathcal{L}^\dagger\|_{\mathrm{Tr}}$ with respect to various α

4.5 Exponential decay of sandwiched Rényi divergences

This section is devoted to proving Thm. 4.7, i.e., the exponential decay of sandwiched Rényi divergences for primitive Lindblad equations with GNS-detailed balance. We start by recalling the definition of a spectral gap, and then we prove a Poincaré inequality (see Prop. 4.27). With this Poincaré inequality, we derive a uniform lower bound of the quantum relative 2-Fisher information, which immediately implies Thm. 4.7 in the case $\alpha = 2$. Next, we prove the Prop. 4.6, which shows that for primitive Lindblad equations with GNS-detailed balance, the quantum LSI (4.4) holds for $\alpha = 1$ (i.e., there exists $K > 0$). Then we prove a quantum comparison theorem (see Prop. 4.32), which implies that the exponential decay rates of sandwiched Rényi divergences along the Lindblad equation are the same for all $\alpha > 1$. The quantum comparison theorem, together with the monotonicity of sandwiched Rényi divergences with respect to the order α, concludes the proof of Thm. 4.7.

4.5.1 Spectral gap and Poincaré inequality

Let us recall from Sec. 4.2 that if the Lindblad equation satisfies GNS-detailed balance, then it also satisfies KMS-detailed balance (i.e., $\mathcal{L}$ is self-adjoint with respect to the inner product $\langle \cdot, \cdot \rangle_{1/2}$). Moreover, by Lem. 4.12, $-\mathcal{L}$ is a positive semi-definite operator with respect to the inner product $\langle \cdot, \cdot \rangle_{1/2}$. By the additional assumption that the Lindblad equation is primitive, we know $\mathbb{I}_n$ is the only eigenvector of $-\mathcal{L}$ with respect to the eigenvalue zero. Hence, spectral theory shows that there exists an orthonormal basis $\{L_1, L_2, \cdots, L_{n^2-1}, \mathbb{I}_n\}$ such that

$$-\mathcal{L}(L_j) = \theta_j L_j, \qquad \theta_1 \geq \theta_2 \geq \cdots \geq \theta_{n^2-1} > 0; \qquad -\mathcal{L}(\mathbb{I}_n) = 0. \tag{4.33}$$

The spectral gap of the operator $-\mathcal{L}$ is $\lambda_{\mathcal{L}} := \theta_{n^2-1}$. It is worth remarking that the set $\{L_1, L_2, \cdots, L_{n^2-1}, \mathbb{I}_n\}$ is an orthonormal basis in the Hilbert space $\mathfrak{M}$ equipped with the inner product $\langle \cdot, \cdot \rangle_{1/2}$, but not necessarily with other inner products. The following Poincaré inequality follows directly from the definition of the spectral gap.

Proposition 4.27 (Poincaré inequality)**.** *Assume that the primitive Lindblad equation satisfies GNS-detailed balance* (4.1)*. For any $A \in \mathfrak{M}$ such that $\langle \mathbb{I}_n, A\rangle_{1/2} = 0$ (or equivalently, $tr(\sigma A) = 0$),*

$$\langle \nabla A, \nabla A\rangle_{1/2} = \langle A, -\mathcal{L}(A)\rangle_{1/2} \geq \lambda_{\mathcal{L}} \langle A, A\rangle_{1/2} . \tag{4.34}$$

Moreover, the equality can be achieved when $A = L_{n^2-1}$.

Proof. For any $A \in \mathfrak{M}$ such that $\langle \mathbb{I}_n, A\rangle_{1/2} = 0$, we know that $A \in \text{span}\{L_j\}_{j=1}^{n^2-1}$. By the definition of the spectral gap, $\langle A, -\mathcal{L}(A)\rangle_{1/2} \geq \lambda_{\mathcal{L}} \langle A, A\rangle_{1/2}$. Moreover, when $A = V_{n^2-1}$, the equality holds. Then (4.34) follows immediately from (4.13). $\square$

Remark. In fact, the above Poincaré inequality is a special case of a whole family of Poincaré inequalities (see [123, Eq. (51)]): given any function $\varphi : (0, \infty) \to (0, \infty)$, we could consider

$$\begin{aligned} C_\varphi \langle A, \widetilde{\varphi}(\Delta_\sigma)(A)\rangle_{1/2} &\equiv C_\varphi \langle A, \varphi(\Delta_\sigma)(A)\rangle_1 \\ &\leq \langle A, \varphi(\Delta_\sigma) \circ (-\mathcal{L})(A)\rangle_1 \equiv \langle A, \widetilde{\varphi}(\Delta_\sigma) \circ (-\mathcal{L})(A)\rangle_{1/2} , \end{aligned}$$

for all $A \in \mathfrak{M}$ such that $\mathrm{Tr}(\sigma A) = 0$, where $\widetilde{\varphi}(x) := x^{-1/2}\varphi(x)$. The inequality in (4.34) simply refers to the case $\varphi(x) = x^{1/2}$. Because $[-\mathcal{L}, \Delta_\sigma] = 0$ (see Lem. 4.11), $-\mathcal{L}$ and Δ_σ are simultaneously diagonalizable. Thus, we might as well choose L_j (4.33) in a way that all L_j are eigenvectors of both $-\mathcal{L}$ and Δ_σ. Also note that both $-\mathcal{L}$ and Δ_σ

are positive semi-definite operators, so they have non-negative spectrum. By expanding $A = \sum_{j=1}^{n^2-1} a_j L_j$, it is not hard to show that $C_\varphi = \lambda_{\mathcal{L}}$ in our situation. Therefore, all such Poincaré inequalities have the same prefactor C_φ for primitive Lindblad equations with GNS-detailed balance.

4.5.2 Proof of Thm. 4.7 for the case $\alpha = 2$

Proposition 4.28 (Lower bound of $\mathscr{I}_2\left(\rho||\sigma\right)$)**.** *Assume that the primitive Lindblad equation* (4.1) *satisfies GNS-detailed balance. Then*

$$\mathscr{I}_2\left(\rho||\sigma\right) \geq 2\lambda_{\mathcal{L}}\left(1 - e^{-D_2(\rho||\sigma)}\right), \qquad \forall \rho \in \mathfrak{D}. \tag{4.35}$$

Proof. The case $\alpha = 2$ is easier to deal with, as all quantities become more explicit:

$$\frac{\delta D_2\left(\rho||\sigma\right)}{\delta \rho} = \frac{2\sigma^{-1/2}\rho\sigma^{-1/2}}{\langle \rho, \sigma^{-1/2}\rho\sigma^{-1/2}\rangle_{HS}} = \frac{2\Upsilon_\sigma^{-1}(\rho)}{\langle \Upsilon_\sigma^{-1}(\rho), \Upsilon_\sigma^{-1}(\rho)\rangle_{1/2}},$$

$$D_2\left(\rho||\sigma\right) = \log \operatorname{Tr}\left(\sigma^{-1/2}\rho\sigma^{-1/2}\rho\right) = \log\left(\left\langle \Upsilon_\sigma^{-1}(\rho), \Upsilon_\sigma^{-1}(\rho)\right\rangle_{1/2}\right).$$

By the definition of the quantum relative α-Fisher information (4.3),

$$\begin{aligned}\mathscr{I}_2\left(\rho||\sigma\right) &= -\Big\langle \frac{2\Upsilon_\sigma^{-1}(\rho)}{\langle \Upsilon_\sigma^{-1}(\rho), \Upsilon_\sigma^{-1}(\rho)\rangle_{1/2}}, \mathcal{L}^\dagger(\rho)\Big\rangle_{HS} \\ &= \frac{-2}{\langle \Upsilon_\sigma^{-1}(\rho), \Upsilon_\sigma^{-1}(\rho)\rangle_{1/2}}\left\langle \Upsilon_\sigma^{-1}(\rho), \Upsilon_\sigma^{-1} \circ \mathcal{L}^\dagger \circ \Upsilon_\sigma(\Upsilon_\sigma^{-1}(\rho))\right\rangle_{1/2}.\end{aligned}$$

By the relation (4.9) and the Poincaré inequality (4.34), we have

$$\begin{aligned}
\mathscr{I}_2(\rho||\sigma) &\overset{(4.9)}{=} \frac{-2}{\langle \Upsilon_\sigma^{-1}(\rho), \Upsilon_\sigma^{-1}(\rho)\rangle_{1/2}} \left\langle \Upsilon_\sigma^{-1}(\rho), \mathcal{L}(\Upsilon_\sigma^{-1}(\rho))\right\rangle_{1/2} \\
&= \frac{-2}{\langle \Upsilon_\sigma^{-1}(\rho), \Upsilon_\sigma^{-1}(\rho)\rangle_{1/2}} \left\langle \Upsilon_\sigma^{-1}(\rho-\sigma), \mathcal{L}(\Upsilon_\sigma^{-1}(\rho-\sigma))\right\rangle_{1/2} \\
&\overset{(4.34)}{\geq} \frac{2}{\langle \Upsilon_\sigma^{-1}(\rho), \Upsilon_\sigma^{-1}(\rho)\rangle_{1/2}} \lambda_{\mathcal{L}} \left\langle \Upsilon_\sigma^{-1}(\rho-\sigma), \Upsilon_\sigma^{-1}(\rho-\sigma)\right\rangle_{1/2}.
\end{aligned}$$

It is not hard to verify that

$$\left\langle \Upsilon_\sigma^{-1}(\rho), \Upsilon_\sigma^{-1}(\rho)\right\rangle_{1/2} = 1 + \left\langle \Upsilon_\sigma^{-1}(\rho-\sigma), \Upsilon_\sigma^{-1}(\rho-\sigma)\right\rangle_{1/2}.$$

It is interesting to note that the term $\langle \Upsilon_\sigma^{-1}(\rho-\sigma), \Upsilon_\sigma^{-1}(\rho-\sigma)\rangle_{1/2}$ turns out to be the quantum χ^2-divergence $\chi^2_{1/2}(\rho,\sigma)$ studied in [134]. Then

$$\begin{aligned}
\mathscr{I}_2(\rho||\sigma) &\geq \frac{2\lambda_{\mathcal{L}}}{\langle \Upsilon_\sigma^{-1}(\rho), \Upsilon_\sigma^{-1}(\rho)\rangle_{1/2}} \Big(\left\langle \Upsilon_\sigma^{-1}(\rho), \Upsilon_\sigma^{-1}(\rho)\right\rangle_{1/2} - 1 \Big) \\
&= \frac{2\lambda_{\mathcal{L}}}{e^{D_2(\rho||\sigma)}} \left(e^{D_2(\rho||\sigma)} - 1 \right),
\end{aligned}$$

which yields (4.35). □

With Prop. 4.28, we can immediately show the exponential decay of the sandwiched Rényi divergence of order $\alpha = 2$.

Proof of Thm. 4.7 for the case $\alpha = 2$. From the definition of the quantum relative 2-Fisher information in (4.3) and by Prop. 4.28,

$$\frac{\mathrm{d}}{\mathrm{d}t} D_2\left(\rho_t||\sigma\right) = -\mathscr{I}_2\left(\rho_t||\sigma\right) \leq -2\lambda_{\mathcal{L}}\left(1 - e^{-D_2(\rho_t||\sigma)}\right).$$

Then $\frac{\mathrm{d}}{\mathrm{d}t}\log\left(e^{D_2(\rho_t||\sigma)} - 1\right) \leq -2\lambda_{\mathcal{L}}$. After integrating it from time 0 to t, one could find, after some straightforward simplification, that

$$D_2\left(\rho_t||\sigma\right) \leq \log\left(1 + (e^{D_2(\rho_0||\sigma)} - 1)e^{-2\lambda_{\mathcal{L}}t}\right) \leq (e^{D_2(\rho_0||\sigma)} - 1)e^{-2\lambda_{\mathcal{L}}t}.$$

Thus, $D_2\left(\rho_t||\sigma\right)$ decays exponentially fast with rate $2\lambda_{\mathcal{L}}$. Apparently, the prefactor $C_{2,\epsilon}$ and the waiting time $\tau_{2,\epsilon}$ could be chosen as

$$C_{2,\epsilon} = \frac{1}{D_2\left(\rho_0||\sigma\right)}\left(e^{D_2(\rho_0||\sigma)} - 1\right), \qquad \tau_{2,\epsilon} = 0.$$

□

The lower bound of $\mathscr{I}_2\left(\rho||\sigma\right)$ in Prop. 4.28 immediately leads into a lower bound of the quantum 2-log Sobolev constant K_2 in (4.4).

Corollary 4.29 ([107, Thm. 4.2])**.** *For primitive Lindblad equations with GNS-detailed balance, the quantum 2-log Sobolev constant K_2 is bounded below by the spectral gap. More specifically,*

$$K_2 \geq \frac{1 - \lambda_{\min}(\sigma)}{\log\left(\lambda_{\min}(\sigma)^{-1}\right)}\lambda_{\mathcal{L}}.$$

Proof. From Prop. 4.28, we know that

$$\frac{\mathscr{I}_2\left(\rho||\sigma\right)}{D_2\left(\rho||\sigma\right)} \geq 2\lambda_{\mathcal{L}} \frac{1-e^{-D_2(\rho||\sigma)}}{D_2\left(\rho||\sigma\right)} \geq 2\lambda_{\mathcal{L}} \frac{1-\lambda_{\min}(\sigma)}{\log(\lambda_{\min}(\sigma)^{-1})},$$

because $x \to \frac{1-e^{-x}}{x}$ is monotonically decreasing on $(0,\infty)$ and the following fact from [107, Lem. 2.1] that $0 \leq D_2\left(\rho||\sigma\right) \leq \log(\lambda_{\min}(\sigma)^{-1})$. The above result follows immediately, by minimizing over ρ for both sides. □

4.5.3 Proof of the quantum LSI

In this subsection, we prove Prop. 4.6, which shows the equivalence of the primitivity and the validity of the quantum LSI (4.4), for Lindblad equations with GNS-detailed balance.

One key ingredient is the l_2 mixing time defined as follows: for any $\epsilon > 0$,

$$\begin{aligned} t_2(\epsilon) &:= \inf\Big\{t \geq 0 : \|e^{t\mathcal{L}}(A) - \mathrm{Tr}(\sigma A)\mathbb{I}_n\|_{2,\sigma} \leq \epsilon,\ \forall A \in \mathfrak{M} \text{ s.t. } \|A\|_{1,\sigma} = 1\Big\} \\ &\equiv \inf\Big\{t \geq 0 : \|e^{t\mathcal{L}}(\cdot) - \mathrm{Tr}(\sigma \cdot)\mathbb{I}_n\|_{1\to 2,\sigma} \leq \epsilon\Big\}. \end{aligned} \tag{4.36}$$

We need the following two results.

Lemma 4.30. *Under the same assumption as in Prop. 4.6, if the Lindblad equation is primitive with the spectral gap $\lambda_{\mathcal{L}}$, then*

$$t_2(\epsilon) \leq \max\Big(0, \frac{1}{2\lambda_{\mathcal{L}}} \log\big(\frac{1}{\lambda_{\min}(\sigma)\epsilon^2}\big)\Big). \tag{4.37}$$

Its proof will be given slightly later in this subsection. As a remark, when $\epsilon \geq \sqrt{\frac{1}{\lambda_{\min}(\sigma)}}$, we have $t_2(\epsilon) = 0$. This is intuitively reasonable, since when ϵ is large enough, the inequality in (4.36) holds trivially.

Theorem 4.31 ([107, Thm. 5.3])**.** *Under the same assumption as in Prop. 4.6, if the Lindblad equation is primitive, then*

$$t_2(e^{-1})\kappa_2 \geq \frac{1}{2}. \tag{4.38}$$

Proof of Prop. 4.6. First, recall that [89, Thm. 16] has proved $K \leq \lambda_{\mathcal{L}}$, by linearizing the quantum LSI. Thus, $K > 0$ implies $\lambda_{\mathcal{L}} > 0$.

As for the other direction, we assume that the Lindblad equation is primitive. Suppose $\rho_t = e^{t\mathcal{L}^\dagger}(\rho_0)$ is the solution of the Lindblad equation. By combining the estimates above, we have

$$K \overset{(4.20)}{=} 2\kappa_1 \overset{(4.21)}{\geq} 2\kappa_2 \overset{(4.38)}{\geq} \frac{1}{t_2(e^{-1})} \overset{(4.37)}{\geq} \frac{2\lambda_{\mathcal{L}}}{2 - \log(\lambda_{\min}(\sigma))} = \frac{1}{1 - \log(\sqrt{\lambda_{\min}(\sigma)})}\lambda_{\mathcal{L}}.$$

$\square$

Remark. In the proof above, we use the mixing time $t_2(e^{-1})$ as a bridge to connect κ_2 and $\lambda_{\mathcal{L}}$, and use the quantum Stroock-Varopoulos inequality to connect K and κ_2. To the best of our knowledge, a direct proof of the quantum LSI (4.4) has not appeared in the literature. For classical systems, Bakry-Émery [12] condition is an important criterion for classical

LSI to hold for Fokker-Planck equations. It is an interesting open question to see whether there is a quantum analog.

Proof of Lem. 4.30. Let us introduce $B = e^{t\mathcal{L}}(A) - \mathrm{Tr}(\sigma A)\mathbb{I}_n$ and decompose $A = a_0\mathbb{I}_n + \sum_{j=1}^{n^2-1} a_j L_j$, where L_j are eigenvectors of the operator $\mathcal{L}$. Recall (4.33) for notations. Then we easily know that $B = \sum_{j=1}^{n^2-1} a_j e^{-t\theta_j} L_j$, and that

$$\begin{aligned}\|e^{t\mathcal{L}}(A) - \mathrm{Tr}(\sigma A)\mathbb{I}_n\|_{2,\sigma}^2 &= \mathrm{Tr}\left(\sigma^{\frac{1}{4}} B^\dagger \sigma^{\frac{1}{4}} \sigma^{\frac{1}{4}} B \sigma^{\frac{1}{4}}\right) = \langle B, B\rangle_{1/2} \\ &= \sum_{j=1}^{n^2-1} e^{-2t\theta_j} |a_j|^2 \le e^{-2t\lambda_{\mathcal{L}}} \sum_{j=1}^{n^2-1} |a_j|^2.\end{aligned}$$

Let $\widetilde{A} := |\sigma^{\frac{1}{2}} A \sigma^{\frac{1}{2}}|$. From the condition that $\|A\|_{1,\sigma} = 1$ in (4.36), we have $\mathrm{Tr}(\widetilde{A}) = 1$. Then, we have

$$\begin{aligned} 1 \ge \mathrm{Tr}(\widetilde{A}^2) &= \mathrm{Tr}\left(\sigma^{\frac{1}{2}} A^\dagger \sigma^{\frac{1}{2}} \sigma^{\frac{1}{2}} A \sigma^{\frac{1}{2}}\right) = \left\langle \sigma^{\frac{1}{4}} \left(\sigma^{\frac{1}{4}} A \sigma^{\frac{1}{4}}\right) \sigma^{\frac{1}{4}}, \sigma^{\frac{1}{4}} \left(\sigma^{\frac{1}{4}} A \sigma^{\frac{1}{4}}\right) \sigma^{\frac{1}{4}} \right\rangle_{HS} \\ &\ge \lambda_{\min}(\sigma) \left\langle \sigma^{\frac{1}{4}} A \sigma^{\frac{1}{4}}, \sigma^{\frac{1}{4}} A \sigma^{\frac{1}{4}} \right\rangle_{HS} = \lambda_{\min}(\sigma) \left\langle A, A\right\rangle_{1/2} \ge \lambda_{\min}(\sigma) \Big(\sum_{j=1}^{n^2-1} |a_j|^2 \Big).\end{aligned}$$

Therefore, we know that

$$\|e^{t\mathcal{L}}(A) - \mathrm{Tr}(\sigma A)\mathbb{I}_n\|_{2,\sigma}^2 \le e^{-2t\lambda_{\mathcal{L}}} \frac{1}{\lambda_{\min}(\sigma)}.$$

Easily, we know that whenever $t \ge \frac{1}{2\lambda_{\mathcal{L}}} \log\left(\frac{1}{\lambda_{\min}(\sigma)\epsilon^2}\right)$, we have $\|e^{t\mathcal{L}}(A) - \mathrm{Tr}(\sigma A)\mathbb{I}_n\|_{2,\sigma} \le \epsilon$. Then, (4.37) follows immediately. ∎

4.5.4 Proof of Thm. 4.7 for $\alpha \in (0, \infty)$ via a quantum comparison theorem

As a quantum analog of [29, Lem. 3.4] for Fokker-Planck equations and [122, Thm. 3.2.3] for the Ornstein-Uhlenbeck process, in Prop. 4.32, we will show that under primitive Lindblad equations with GNS-detailed balance (4.1), the sandwiched Rényi divergence of higher-order α_1 can be bounded above by the sandwiched Rényi divergence of lower-order α_0 ($1 < \alpha_0 < \alpha_1$), at the expense of some time delay. This result helps to prove the exponential decay for sandwiched Rényi divergences under Lindblad equations, summarized in Prop. 4.33.

Proposition 4.32 (Quantum comparison theorem). *Let ρ_t be the solution of the Lindblad equation with GNS-detailed balance* (4.1)*. Assume that the quantum LSI* (4.4) *holds with constant K. Assume also that*

$$D\left(\rho_0||\sigma\right) \leq \epsilon \text{ for some fixed } \epsilon \in \left(0, \frac{\lambda_{\min}(\sigma)^2}{2}\right). \tag{4.39}$$

Then for any two orders $1 < \alpha_0 \leq \alpha_1$, we have

$$D_{\alpha_1}\left(\rho_T||\sigma\right) \leq D_{\alpha_0}\left(\rho_0||\sigma\right), \tag{4.40}$$

with

$$T = \frac{1}{2K\eta} \log\left(\frac{\alpha_1 - 1}{\alpha_0 - 1}\right), \tag{4.41}$$

where

$$\eta = \min\left(\frac{1}{2}, \min_{j=1}^{J}\left\{\frac{2\sqrt{e^{\omega_j}\frac{1}{\Lambda}}}{1+e^{\omega_j}\Lambda}\right\}\right),$$

$$\Lambda = \frac{\lambda_{\max}(\sigma)}{\lambda_{\min}(\sigma)} \exp\left(\alpha_0\sqrt{2\epsilon}\frac{2\lambda_{\min}(\sigma) - \sqrt{2\epsilon}}{\lambda_{\min}(\sigma)(\lambda_{\min}(\sigma) - \sqrt{2\epsilon})}\right).$$

The proof is postponed to Sec. 4.5.5. Later in Sec. 4.5.6, we will discuss this quantum comparison theorem in detail. As a reminder, we have implicitly assumed that Lindblad equations under consideration are primitive, by Prop. 4.6.

Remark. (i) The assumption (4.39) on the initial condition ρ_0 is merely a technical assumption allowing us to obtain a neat expression for the delay time T. This assumption (4.39) is not essential for the inequality (4.40) to hold. One could remove the assumption at the expense of longer delay time: the validity of the quantum LSI (4.4) leads to the exponential decay of the quantum relative entropy so that $D(\rho_{t_0}||\sigma) \le \epsilon$ holds for some $t_0 \ge 0$. Taking ρ_{t_0} in place of ρ_0, one sees that this would imply (4.40) with a larger delay time T.

(ii) Another reason for imposing the assumption (4.39) is to avoid technicalities. Thanks to the quantum Pinsker's inequality

$$D\left(\rho||\sigma\right) \geq \frac{1}{2}\left(\mathrm{Tr}|\rho-\sigma|\right)^2, \tag{4.42}$$

the assumption that $D\left(\rho_0||\sigma\right) \leq \epsilon$ with $0 < \epsilon < \frac{\lambda_{\min}(\sigma)^2}{2}$ implies that ρ_0 is a strictly positive density matrix. We can even show that $\rho_t \in \mathfrak{D}^+$, for any $t \geq 0$. This assumption helps to avoid many technical issues arising from degenerate density matrices, and it validates the usage of results in the previous two sections, e.g., properties of the operator $[X]_\omega$ (which require X to be strictly positive).

Proposition 4.33. *Let ρ_t be the solution of the Lindblad equation with GNS-detailed balance* (4.1)*. Assume that the quantum LSI* (4.4) *holds with constant K. If the sandwiched Rényi divergence decays exponentially fast with rate R for some order $\alpha_0 > 1$, then it decays exponentially fast for any $\alpha \in (0,\infty)$ with rate at least R.*

Proof of Prop. 4.33. The exponential decay of sandwiched Rényi divergences means there exists some time t_0 and constant C_0 such that

$$D_{\alpha_0}\left(\rho_t||\sigma\right) \leq C_0 D_{\alpha_0}\left(\rho_0||\sigma\right) e^{-Rt}, \qquad \forall t \geq t_0.$$

Case (I): $\alpha \leq \alpha_0$. By the monotonicity of sandwiched Rényi divergences with respect to the order α [111, Thm. 7],

$$D_\alpha\left(\rho_t||\sigma\right) \le D_{\alpha_0}\left(\rho_t||\sigma\right) \le C_0 D_{\alpha_0}\left(\rho_0||\sigma\right)e^{-Rt} = C_\alpha D_\alpha\left(\rho_0||\sigma\right)e^{-Rt},$$

for all $t \ge t_0$, where $C_\alpha = \frac{C_0 D_{\alpha_0}(\rho_0||\sigma)}{D_\alpha(\rho_0||\sigma)}$.

Case (II): $\alpha > \alpha_0$. Pick an arbitrary $0 < \epsilon < \frac{\lambda_{\min}(\sigma)^2}{2}$. Notice that the quantum LSI implies the exponential decay of the quantum relative entropy, namely,

$$D\left(\rho_t||\sigma\right) \le D\left(\rho_0||\sigma\right)e^{-2Kt}.$$

Hence, when $t \ge \frac{1}{2K}\log\left(\frac{D(\rho_0||\sigma)}{\epsilon}\right)$, one has $D\left(\rho_t||\sigma\right) \le \epsilon$. Moreover, if

$$t \ge T + \max\left(t_0, \frac{1}{2K}\log\left(\frac{D\left(\rho_0||\sigma\right)}{\epsilon}\right)\right),$$

where T is the delay time given in Prop. 4.32 with the choice $\alpha_1 = \alpha$, then one obtains from Prop. 4.32 that

$$D_\alpha\left(\rho_t||\sigma\right) \le D_{\alpha_0}\left(\rho_{t-T}||\sigma\right) \le C_0 D_{\alpha_0}\left(\rho_0||\sigma\right)e^{-R(t-T)} = C_\alpha D_\alpha\left(\rho_0||\sigma\right)e^{-Rt},$$

where $C_\alpha = \frac{C_0 D_{\alpha_0}(\rho_0||\sigma)e^{RT}}{D_\alpha(\rho_0||\sigma)}$.

Combining both cases above finishes the proof of Prop. 4.33. $\square$

Proof of Thm. 4.7. Set $\alpha_0 = 2$ and $R = 2\lambda_{\mathcal{L}}$ in Prop. 4.33. Thm. 4.7 follows immediately from Prop. 4.33 and the case $\alpha = 2$ proved in Sec. 4.5.2. The validity of the quantum LSI has been shown in Prop. 4.6. The expression of $T_{\alpha,\epsilon}$ follows immediately from Prop. 4.32.

□

4.5.5 Proof of the quantum comparison theorem

We now turn to the proof of Prop. 4.32. Consider the density matrix

$$\varrho = \frac{\left(\Upsilon_{\sigma}^{\frac{1-\alpha}{\alpha}}(\rho)\right)^{\alpha}}{Z} \equiv \frac{\rho_{\sigma}^{\alpha}}{Z},$$

where the normalization constant $Z = \operatorname{Tr}(\rho_{\sigma}^{\alpha})$ and $\rho \in \mathfrak{D}^{+}$. Then apparently, $\varrho \in \mathfrak{D}^{+}$. The log-Sobolev inequality (4.4), with ρ in (4.4) chosen as ϱ above, becomes

$$\begin{aligned} &\operatorname{Tr}\left(\rho_{\sigma}^{\alpha}\log\left(\rho_{\sigma}^{\alpha}\right)\right) - Z\log(Z) - \operatorname{Tr}\left(\rho_{\sigma}^{\alpha}\log(\sigma)\right) \\ &\qquad\qquad \leq \frac{1}{2K}\operatorname{Tr}\Big(-\mathcal{L}^{\dagger}(\rho_{\sigma}^{\alpha})\big(\log(\rho_{\sigma}^{\alpha}) - \log(\sigma)\big)\Big). \end{aligned} \tag{4.43}$$

This inequality can be regarded as a variant of the log-Sobolev inequality, and it will be used later. Let us define the operator $G_{X,\omega}(s) : \mathfrak{M} \to \mathfrak{M}$ by

$$G_{X,\omega}(s)(A) := e^{\omega s} X^{s} A X^{-s} + e^{\omega(1-s)} X^{1-s} A X^{-(1-s)}. \tag{4.44}$$

The lemma below collects some properties of $G_{X,\omega}(s)$, which will be useful in the proof of Prop. 4.32.

Lemma 4.34. *Assume that $X \in \mathfrak{P}^+$ and $\omega \in \mathbb{R}$. Then the operator $G_{X,\omega}(s)$ satisfies:*

(i) For $0 \leq s_1 < s_2 \leq \frac{1}{2}$,

$$G_{X,\omega}(s_1) \geq G_{X,\omega}(s_2).$$

Moreover, the equality is obtained iff $\omega = 0$ and $X = c\mathbb{I}_n$ for some constant $c > 0$.

(ii) For $s \in [0, \frac{1}{2}]$, a bound for the spectrum of $G_{X,\omega}(s)$ is

$$2\sqrt{e^{\omega}\frac{\lambda_{\min}(X)}{\lambda_{\max}(X)}}Id \leq G_{X,\omega}(s) \leq \Big(1 + e^{\omega}\frac{\lambda_{\max}(X)}{\lambda_{\min}(X)}\Big)Id.$$

As a consequence,

$$\xi G_{X,\omega}(s_1) \leq G_{X,\omega}(s_2), \qquad \forall s_1, s_2 \in \big[0, \frac{1}{2}\big],$$

with

$$\xi = 2\sqrt{e^{\omega}\frac{\lambda_{\min}(X)}{\lambda_{\max}(X)}}\Big/\Big(1 + e^{\omega}\frac{\lambda_{\max}(X)}{\lambda_{\min}(X)}\Big).$$

Proof of Lem. 4.34. (i) For any positive number b, notice that $b^s + b^{1-s}$ is convex and non-increasing with respect to s on the interval $[0, \frac{1}{2}]$. Moreover,

- the range of the function $b^s + b^{1-s}$ on this interval is $[2\sqrt{b}, 1 + b]$;
- $b^{s_1} + b^{1-s_1} = b^{s_2} + b^{1-s_2}$ for some $0 \leq s_1 < s_2 \leq \frac{1}{2}$ iff $b = 1$.

Let us write the spectral decomposition of X as $X = \sum_{k=1}^{n} \lambda_k |k\rangle \langle k|$ where $\lambda_k > 0$

by assumption. Then for any matrix A,

$$\langle A, G_{X,\omega}(s)A\rangle_{HS} = \sum_{k,l=1}^{n}\left(\left(e^{\omega}\frac{\lambda_k}{\lambda_l}\right)^{s} + \left(e^{\omega}\frac{\lambda_k}{\lambda_l}\right)^{1-s}\right)|A_{k,l}|^2.$$

Since $b^s + b^{1-s}$ is non-increasing with respect to s, so is $G_{X,\omega}(s)$. To achieve equality, we need $e^{\omega}\frac{\lambda_k}{\lambda_l} = 1$ for all k, l. Hence, $\omega = 0$ and there exists some $c = \lambda_k$ for all $1 \leq k \leq n$ (i.e., $X = c\mathbb{I}_n$).

(ii) From the last equation, by using the range of $b^s + b^{1-s}$, we immediately have

$$\langle A, G_{X,\omega}(s)A\rangle_{HS} \leq \sum_{k,l}(1 + e^{\omega}\frac{\lambda_k}{\lambda_l})|A_{k,l}|^2 \leq \left(1 + e^{\omega}\frac{\lambda_{\max}(X)}{\lambda_{\min}(X)}\right)\langle A, A\rangle_{HS},$$

$$\langle A, G_{X,\omega}(s)A\rangle_{HS} \geq \sum_{k,l} 2\sqrt{e^{\omega}\frac{\lambda_k}{\lambda_l}}|A_{k,l}|^2 \geq 2\sqrt{e^{\omega}\frac{\lambda_{\min}(X)}{\lambda_{\max}(X)}}\langle A, A\rangle_{HS},$$

which proves part (ii).

□

Proof of Prop. 4.32. Let us first introduce some useful notations. Suppose ρ_t is evolving according to the Lindblad equation (4.1), and $\beta_t = 1 + (\alpha_0 - 1)e^{\eta 2Kt}$ is also changing with respect to time for some positive constant η, independent of any specific initial condition ρ_0. The parameter $\eta \in (0, 1]$ is yet to be determined later. We are interested in the time interval $[0, T]$ such that $\beta_t|_{t=0} = \alpha_0, \beta_t|_{t=T} = \alpha_1$. More specifically, $T = \frac{1}{2K\eta}\log\left(\frac{\alpha_1 - 1}{\alpha_0 - 1}\right)$.

Let us define an important quantity $\rho_{\sigma,t}$ by

$$\rho_{\sigma,t} := \Upsilon_{\sigma}^{\frac{1-\beta_t}{\beta_t}}(\rho_t).$$

We also define

$$Z_t := \mathrm{Tr}\left(\left(\Upsilon_{\sigma}^{\frac{1-\beta_t}{\beta_t}}(\rho_t)\right)^{\beta_t}\right) \equiv \mathrm{Tr}\left(\rho_{\sigma,t}^{\beta_t}\right), \qquad F_t := \frac{1}{\beta_t}\log Z_t.$$

After taking the derivative of F_t with respect to time t and arranging terms,

$$\beta_t^2 Z_t \dot{F}_t = \beta_t \dot{Z}_t - \dot{\beta}_t Z_t \log(Z_t).$$

We claim that to prove the inequality (4.40), it suffices to show that F_t is monotonically decreasing, i.e., $\dot{F}_t \leq 0$ on the interval $[0, T]$. In fact, from the definition of F and the sandwiched Rényi divergence (4.2), we immediately obtain from $F_T \leq F_0$ that

$$D_{\alpha_1}(\rho_T||\sigma) \leq \frac{\alpha_1}{\alpha_1 - 1}\frac{\alpha_0 - 1}{\alpha_0} D_{\alpha_0}(\rho_0||\sigma) \leq D_{\alpha_0}(\rho_0||\sigma).$$

The rest devotes to the proof of $\dot{F}_t \leq 0$, which is done in the following steps.

Step (I): Simplification of $\dot{F}_t$.

Let us first compute $\dot{Z}_t$.

$$
\begin{aligned}
\dot{Z}_t &\equiv \frac{\mathrm{d}}{\mathrm{d}t}\operatorname{Tr}\Big(\big(\sigma^{\frac{1-\beta_t}{2\beta_t}}\rho_t\sigma^{\frac{1-\beta_t}{2\beta_t}}\big)^{\beta_t}\Big)\\
&= \beta_t \operatorname{Tr}\left(\rho_{\sigma,t}^{\beta_t-1}\Big(\sigma^{\frac{1-\beta_t}{2\beta_t}}\mathcal{L}^\dagger(\rho_t)\sigma^{\frac{1-\beta_t}{2\beta_t}} - \frac{\dot{\beta}_t}{2\beta_t^2}\sigma^{\frac{1-\beta_t}{2\beta_t}}\left\{\log(\sigma),\rho_t\right\}\sigma^{\frac{1-\beta_t}{2\beta_t}}\Big)\right)\\
&\qquad + \dot{\beta}_t \operatorname{Tr}\Big(\rho_{\sigma,t}^{\beta_t}\log(\rho_{\sigma,t})\Big)\\
&= \beta_t \operatorname{Tr}\Big(\sigma^{\frac{1-\beta_t}{2\beta_t}}\rho_{\sigma,t}^{\beta_t-1}\sigma^{\frac{1-\beta_t}{2\beta_t}}\mathcal{L}^\dagger(\rho_t)\Big) - \frac{\dot{\beta}_t}{\beta_t}\operatorname{Tr}\Big(\rho_{\sigma,t}^{\beta_t}\log(\sigma)\Big) + \frac{\dot{\beta}_t}{\beta_t}\operatorname{Tr}\Big(\rho_{\sigma,t}^{\beta_t}\log(\rho_{\sigma,t}^{\beta_t})\Big).
\end{aligned}
$$

Hence, by (4.43) and the fact that $\dot{\beta}_t > 0$,

$$
\begin{aligned}
&\beta_t^2 Z_t \dot{F}_t\\
&= \beta_t^2 \operatorname{Tr}\Big(\sigma^{\frac{1-\beta_t}{2\beta_t}}\rho_{\sigma,t}^{\beta_t-1}\sigma^{\frac{1-\beta_t}{2\beta_t}}\mathcal{L}^\dagger(\rho_t)\Big) - \dot{\beta}_t \operatorname{Tr}\Big(\rho_{\sigma,t}^{\beta_t}\log(\sigma)\Big)\\
&\qquad + \dot{\beta}_t \operatorname{Tr}\Big(\rho_{\sigma,t}^{\beta_t}\log(\rho_{\sigma,t}^{\beta_t})\Big) - \dot{\beta}_t Z_t \log(Z_t) \qquad (4.45)\\
&\le \beta_t^2 \operatorname{Tr}\Big(\sigma^{\frac{1-\beta_t}{2\beta_t}}\rho_{\sigma,t}^{\beta_t-1}\sigma^{\frac{1-\beta_t}{2\beta_t}}\mathcal{L}^\dagger(\rho_t)\Big) + \frac{\dot{\beta}_t}{2K}\operatorname{Tr}\Big(-\mathcal{L}^\dagger(\rho_{\sigma,t}^{\beta_t})\big(\log(\rho_{\sigma,t}^{\beta_t}) - \log(\sigma)\big)\Big)\\
&=: T_1 + \frac{\dot{\beta}_t}{2K}T_2.
\end{aligned}
$$

In the last line, we introduce T_1 and T_2 as short-hand notations.

Next, we further simplify T_1 and T_2. For T_1,

$$
\begin{aligned}
T_1 &= \beta_t^2 \operatorname{Tr}\left(\sigma^{\frac{1-\beta_t}{2\beta_t}} \rho_{\sigma,t}^{\beta_t-1} \sigma^{\frac{1-\beta_t}{2\beta_t}} \mathcal{L}^\dagger(\rho_t)\right) \\
&\overset{(4.22)}{=} \beta_t(\beta_t-1) Z_t \operatorname{Tr}\left(\frac{\delta D_{\beta_t}(\rho||\sigma)}{\delta\rho}\bigg|_{\rho=\rho_t} \mathcal{L}^\dagger(\rho_t)\right) \\
&\overset{(4.27)}{=} -\beta_t(\beta_t-1) Z_t \Big\langle \nabla \frac{\delta D_{\beta_t}(\rho||\sigma)}{\delta\rho}\bigg|_{\rho=\rho_t}, [\rho_t]_{\beta_t,\vec{\omega}} \left(\nabla \frac{\delta D_{\beta_t}(\rho||\sigma)}{\delta\rho}\bigg|_{\rho=\rho_t}\right)\Big\rangle_{HS} \\
&\overset{(4.24)}{=} -(\beta_t-1) Z_t^2 \sum_j \left\langle A_j, \left[\rho_{\sigma,t}^{\beta_t-1}\right]_{\frac{(\beta_t-1)\omega_j}{\beta_t}} \circ [\rho_{\sigma,t}]_{\frac{\omega_j}{\beta_t}} (A_j)\right\rangle_{HS},
\end{aligned}
$$

where $A_j = \left[\rho_{\sigma,t}^{\beta_t-1}\right]^{-1}_{\frac{(\beta_t-1)\omega_j}{\beta_t}} \circ \Upsilon_\sigma^{\frac{\beta_t-1}{\beta_t}} \left(\partial_j \frac{\delta D_{\beta_t}(\rho||\sigma)}{\delta\rho}\Big|_{\rho=\rho_t}\right)$. As for T_2, by using (4.27) again (for the case $\alpha = 1$ in (4.27)),

$$
\begin{aligned}
T_2 &= -\left\langle \log(\rho_{\sigma,t}^{\beta_t}) - \log(\sigma), \mathcal{L}^\dagger(\rho_{\sigma,t}^{\beta_t})\right\rangle_{HS} \\
&= \left\langle \nabla\big(\beta_t \log(\rho_{\sigma,t}) - \log(\sigma)\big), \left[\rho_{\sigma,t}^{\beta_t}\right]_{\vec{\omega}} \nabla\big(\beta_t \log(\rho_{\sigma,t}) - \log(\sigma)\big)\right\rangle_{HS}.
\end{aligned}
$$

Consider the term

$$
\begin{aligned}
\partial_j\big(\beta_t \log(\rho_{\sigma,t}) - \log(\sigma)\big) &= \beta_t\Big(\partial_j \log(\rho_{\sigma,t}) - \frac{\omega_j}{\beta_t} V_j\Big) \\
&= \beta_t \left(V_j \log\left(e^{-\frac{\omega_j}{2\beta_t}} \rho_{\sigma,t}\right) - \log\left(e^{\frac{\omega_j}{2\beta_t}} \rho_{\sigma,t}\right) V_j\right) \\
&\overset{(4.28)}{=} Z_t\Big(\left[\rho_{\sigma,t}^{\beta_t-1}\right]^{-1}_{\frac{\beta_t-1}{\beta_t}\omega_j} \circ \Upsilon_\sigma^{\frac{\beta_t-1}{\beta_t}}\Big) \partial_j\Big(\frac{\delta D_{\beta_t}(\rho||\sigma)}{\delta\rho}\bigg|_{\rho=\rho_t}\Big) = Z_t A_j.
\end{aligned}
$$

Here, to obtain the second line, we have used $\partial_j \log(\sigma) \equiv [V_j, \log(\sigma)] = \partial_s\left(\Delta_\sigma^{-s} V_j\right)|_{s=0} =$

$\partial_s\left(e^{\omega_j s} V_j\right)|_{s=0} = \omega_j V_j$ (see [36, Lem. 5.9]). It follows immediately that

$$T_2 = Z_t^2 \sum_j \left\langle A_j, \left[\rho_{\sigma,t}^{\beta_t}\right]_{\omega_j}(A_j)\right\rangle_{HS}.$$

By plugging the expressions of T_1 and T_2 back into (4.45), and using the fact that $\frac{\dot{\beta}_t}{2K} = \frac{\eta 2K(\alpha_0-1)e^{\eta 2Kt}}{2K} = \eta(\beta_t - 1)$, we have

$$\beta_t^2 Z_t \dot{F}_t \leq$$

$$(\beta_t - 1)Z_t^2 \sum_j \left(\eta\left\langle A_j, \left[\rho_{\sigma,t}^{\beta_t}\right]_{\omega_j}(A_j)\right\rangle_{HS} - \left\langle A_j, \left[\rho_{\sigma,t}^{\beta_t-1}\right]_{\frac{(\beta_t-1)\omega_j}{\beta_t}} \circ \left[\rho_{\sigma,t}\right]_{\frac{\omega_j}{\beta_t}}(A_j)\right\rangle_{HS}\right).$$

Step (II): We show that $\dot{F}_t \leq 0$.

By observing the last equation, since $\beta_t - 1 > 0$ and $Z_t > 0$, it is sufficient to show that each term indexed by j on the right hand side of the last equation is less than or equal to zero, namely,

$$\eta\left\langle A_j, \left[\rho_{\sigma,t}^{\beta_t}\right]_{\omega_j}(A_j)\right\rangle_{HS} \leq \left\langle A_j, \left[\rho_{\sigma,t}^{\beta_t-1}\right]_{\frac{(\beta_t-1)\omega_j}{\beta_t}} \circ \left[\rho_{\sigma,t}\right]_{\frac{\omega_j}{\beta_t}}(A_j)\right\rangle_{HS},$$

for any time $t \in [0, T]$ and index j. In the commutative case, both left and right hand sides represent multiplication of the operator $\rho_{\sigma,t}^{\beta_t}$, and the above inequality holds trivially. For noncommutative (quantum) systems, the above inequality is non-trivial. In this step, we will simplify the expression in the last equation and find sufficient conditions for η so that the last inequality holds.

By definition (4.15) and let $B_j = A_j \rho_{\sigma,t}^{\frac{\beta_t}{2}}$, the above inequality can be simplified to

$$\begin{aligned}
&\eta\Big\langle B_j, \int_0^1 e^{\omega_j s}\rho_{\sigma,t}^{\beta_t s} B_j \rho_{\sigma,t}^{-\beta_t s}\,\mathrm{d}s\Big\rangle_{HS} \\
&\leq \Big\langle B_j, \int_0^1\int_0^1 e^{\frac{(\beta_t-1)\omega_j v}{\beta_t}} e^{\frac{\omega_j u}{\beta_t}} \rho_{\sigma,t}^{(\beta_t-1)v+u} B_j \rho_{\sigma,t}^{-(\beta_t-1)v-u}\,\mathrm{d}u\,\mathrm{d}v\Big\rangle_{HS} \\
&= \Big\langle B_j, \int_0^1 e^{\omega_j s}\rho_{\sigma,t}^{\beta_t s} B_j \rho_{\sigma,t}^{-\beta_t s}\left(\int_{U_s}\left|\frac{\partial(u,v)}{\partial(s,\widetilde{s})}\right|\mathrm{d}\widetilde{s}\right)\,\mathrm{d}s\Big\rangle_{HS} \\
&= \Big\langle B_j, \int_0^1 e^{\omega_j s}\rho_{\sigma,t}^{\beta_t s} B_j \rho_{\sigma,t}^{-\beta_t s} f(s)\,\mathrm{d}s\Big\rangle_{HS}.
\end{aligned} \tag{4.46}$$

From the second line to the third line, we have used the change of variable $s = \frac{(\beta_t-1)v+u}{\beta_t}$ and $\widetilde{s} = \frac{(\beta_t-1)v-u}{\beta_t}$. The set U_s is defined as

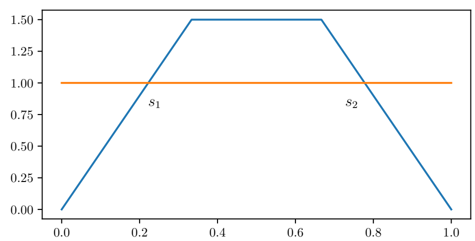
$$U_s := \Big\{\tilde{s}\in\mathbb{R} : 0 \leq \frac{\beta_t}{2(\beta_t-1)}(\widetilde{s}+s) \leq 1 \text{ and } 0 \leq \frac{\beta_t}{2}(s-\widetilde{s}) \leq 1\Big\}.$$

The weight function $f(s) := \int_{U_s}\left|\frac{\partial(u,v)}{\partial(s,\widetilde{s})}\right|\mathrm{d}\widetilde{s} = \int_{U_s}\frac{\beta_t^2}{2(\beta_t-1)}\,\mathrm{d}\widetilde{s}$ is a probability distribution on $[0,1]$, which can be expressed explicitly as

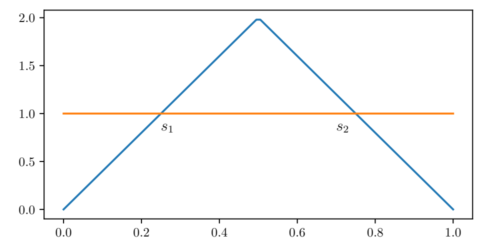
$$f(s) = \frac{\beta_t^2}{2(\beta_t-1)}\Big(\min(s, \frac{2(\beta_t-1)}{\beta_t} - s) - \max(-s, s - \frac{2}{\beta_t})\Big), \qquad s\in[0,1].$$

The graphs of $f(s)$ for $\beta_t = 1.5$, 2, and 3 are given in Fig. 4.3 for illustration. The largest value of $f(s)$ on the interval $[0,1]$ is β_t if $\beta_t \leq 2$, and is $\frac{\beta_t}{\beta_t-1}$ if $\beta_t \geq 2$. In either case, since $\beta_t > 1$, the largest value of $f(s)$ must be strictly larger than 1. Therefore, we can define s_1

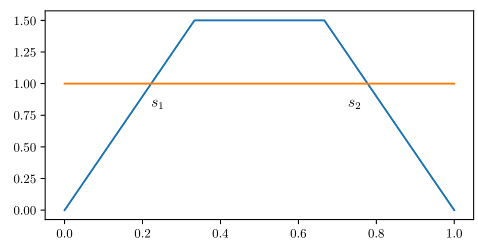
and s_2 as the two zeros of the function $f(s) - 1$, as visualized in Fig. 4.3. More explicitly, $s_1 = (\beta_t - 1)/\beta_t^2$ and by symmetry, $s_2 = 1 - s_1$.



(a) $\beta_t = 1.5$



(b) $\beta_t = 2$



(c) $\beta_t = 3$

Figure 4.3: Weight function $f(s)$ in blue color for $\beta_t = 1.5$, 2, 3.

To show (4.46), it is then sufficient to show that

$$\eta \int_0^1 e^{\omega_j s} \rho_{\sigma,t}^{\beta_t s}(\cdot)\rho_{\sigma,t}^{-\beta_t s}\, \mathrm{d}s \leq \int_0^1 e^{\omega_j s} \rho_{\sigma,t}^{\beta_t s}(\cdot)\rho_{\sigma,t}^{-\beta_t s} f(s)\, \mathrm{d}s,$$

in the sense of operators. In the above inequality, the notation $\rho_{\sigma,t}^{\beta_t s}(\cdot)\rho_{\sigma,t}^{-\beta_t s}$ is an operator mapping $A \in \mathfrak{M}$ to $\rho_{\sigma,t}^{\beta_t s} A \rho_{\sigma,t}^{-\beta_t s} \in \mathfrak{M}$. Notice that the function $f(s)$ is symmetric with respect to the axis $s = 1/2$, and recall the definition of $G_{X,\omega}(s)$ given by (4.44). We can rewrite the last inequality as

$$\eta \int_0^{1/2} G_{\rho_{\sigma,t}^{\beta_t},\omega_j}(s)\,\mathrm{d}s \leq \int_0^{1/2} G_{\rho_{\sigma,t}^{\beta_t},\omega_j}(s) f(s)\,\mathrm{d}s. \tag{4.47}$$

By Lem. 4.34 part (i), we have $\int_0^{1/2} G_{\rho_{\sigma,t}^{\beta_t},\omega_j}(s)\,\mathrm{d}s \geq \int_0^{1/2} G_{\rho_{\sigma,t}^{\beta_t},\omega_j}(s) f(s)\,\mathrm{d}s$. Apparently, we need $\eta \leq 1$ for the inequality (4.47) to hold. Thus, from now on, we shall restrict to $\eta \leq 1$. By subtracting the common parts in the integral (4.47), in order to show (4.47), it suffices to prove that

$$\eta \int_0^{s_1} G_{\rho_{\sigma,t}^{\beta_t},\omega_j}(s)(1-f(s))\,\mathrm{d}s \leq \int_{s_1}^{\frac{1}{2}} G_{\rho_{\sigma,t}^{\beta_t},\omega_j}(s)(f(s)-1)\,\mathrm{d}s. \tag{4.48}$$

Note that $\int_0^{s_1}(1-f(s))\,\mathrm{d}s = \int_{s_1}^{\frac{1}{2}}(f(s)-1)\,\mathrm{d}s$, since $f(s)$ is symmetric about the line $s = \frac{1}{2}$ and also $f(s)$ is a probability distribution on $[0,1]$. Then by Lem. 4.34 part (i),

$$\begin{aligned}\int_{s_1}^{\frac{1}{2}} G_{\rho_{\sigma,t}^{\beta_t},\omega_j}(s)(f(s)-1)\,\mathrm{d}s &\geq \int_{s_1}^{\frac{1}{2}} G_{\rho_{\sigma,t}^{\beta_t},\omega_j}\Big(\frac{1}{2}\Big)(f(s)-1)\,\mathrm{d}s \\ &= \int_0^{s_1} G_{\rho_{\sigma,t}^{\beta_t},\omega_j}\Big(\frac{1}{2}\Big)(1-f(s))\,\mathrm{d}s.\end{aligned}$$

By Lem. 4.34 part (ii), if we choose η such that

$$\eta \leq 2\sqrt{e^{\omega_j}\frac{\lambda_{\min}(\rho_{\sigma,t}^{\beta_t})}{\lambda_{\max}(\rho_{\sigma,t}^{\beta_t})}}\Big/\Big(1+e^{\omega_j}\frac{\lambda_{\max}(\rho_{\sigma,t}^{\beta_t})}{\lambda_{\min}(\rho_{\sigma,t}^{\beta,t})}\Big), \tag{4.49}$$

then $\eta G_{\rho_{\sigma,t}^{\beta_t},\omega_j}(s) \leq G_{\rho_{\sigma,t}^{\beta_t},\omega_j}\left(\frac{1}{2}\right)$ for all $0 \leq s \leq \frac{1}{2}$, which implies (4.48).

Hence, the remaining task is to show the existence of η such that (4.49) holds, for all $j \in \{1, 2, \cdots, \mathsf{J}\}$, all $t \in [0, T]$, and any initial condition ρ_0 satisfying the assumption that $D(\rho_0||\sigma) \leq \epsilon$ for some $0 < \epsilon < \frac{\lambda_{\min}(\sigma)^2}{2}$. The following step (III) is fully devoted into finding such an η.

Step (III): We show there exists $\eta \in (0, 1]$ satisfying (4.49).

By observing the term on the right hand side of (4.49), it suffices to show that $\frac{\lambda_{\max}(\rho_{\sigma,t}^{\beta_t})}{\lambda_{\min}(\rho_{\sigma,t}^{\beta,t})}$ is bounded from above. We first establish an upper bound for $\lambda_{\max}(\rho_{\sigma,t}^{\beta_t})$. In fact, from the quantum LSI (4.4) we have

$$D(\rho_t||\sigma) \leq D(\rho_0||\sigma)\, e^{-2Kt} \leq \epsilon e^{-2Kt} =: \epsilon_t.$$

Moreover, by the quantum Pinsker's inequality (4.42), we obtain that $\mathrm{Tr}|\rho_t - \sigma| \leq \sqrt{2\epsilon_t}$, which implies that

$$-\sqrt{2\epsilon_t}\, \mathbb{I}_n \leq \rho_t - \sigma \leq \sqrt{2\epsilon_t}\, \mathbb{I}_n.$$

As a result,

$$\begin{aligned}
\sigma^{\frac{1-\beta_t}{2\beta_t}} \rho_t \sigma^{\frac{1-\beta_t}{2\beta_t}} &= \sigma^{\frac{1-\beta_t}{2\beta_t}} (\rho_t - \sigma) \sigma^{\frac{1-\beta_t}{2\beta_t}} + \sigma^{\frac{1}{\beta_t}} \\
&\leq \sigma^{\frac{1}{\beta_t}} + \sqrt{2\epsilon_t}\sigma^{\frac{1-\beta_t}{\beta_t}} \leq \left(\lambda_{\max}(\sigma)^{\frac{1}{\beta_t}} + \sqrt{2\epsilon_t}\lambda_{\min}(\sigma)^{\frac{1-\beta_t}{\beta_t}}\right) \mathbb{I}_n.
\end{aligned} \tag{4.50}$$

It follows that

$$
\begin{aligned}
\lambda_{\max}(\rho_{\sigma,t}^{\beta_t}) &= \lambda_{\max}\Big(\big(\sigma^{\frac{1-\beta_t}{2\beta_t}}\rho_t\sigma^{\frac{1-\beta_t}{2\beta_t}}\big)^{\beta_t}\Big) \leq \Big(\lambda_{\max}(\sigma)^{\frac{1}{\beta_t}} + \sqrt{2\epsilon_t}\lambda_{\min}(\sigma)^{\frac{1-\beta_t}{\beta_t}}\Big)^{\beta_t} \\
&= \lambda_{\max}(\sigma)\left(1 + \frac{\sqrt{2\epsilon_t}}{\lambda_{\min}(\sigma)}\Big(\frac{\lambda_{\min}(\sigma)}{\lambda_{\max}(\sigma)}\Big)^{\frac{1}{\beta_t}}\right)^{\beta_t} \\
&\leq \lambda_{\max}(\sigma)\exp\left(\beta_t\frac{\sqrt{2\epsilon_t}}{\lambda_{\min}(\sigma)}\Big(\frac{\lambda_{\min}(\sigma)}{\lambda_{\max}(\sigma)}\Big)^{\frac{1}{\beta_t}}\right).
\end{aligned}
$$

In the last inequality, we used an elementary inequality $(1+x)^y = \exp(y\log(1+x)) \leq \exp(xy)$ for $x, y \geq 0$. By the definition of β_t, the exponent on the right side of the above equation can be bounded above by

$$
\begin{aligned}
&\beta_t\frac{\sqrt{2\epsilon_t}}{\lambda_{\min}(\sigma)}\Big(\frac{\lambda_{\min}(\sigma)}{\lambda_{\max}(\sigma)}\Big)^{\frac{1}{\beta_t}} \\
&= (1 + (\alpha_0 - 1)e^{\eta 2Kt})\frac{\sqrt{2\epsilon}}{\lambda_{\min}(\sigma)}\exp\left(-Kt - \frac{1}{\beta_t}\log\Big(\frac{\lambda_{\max}(\sigma)}{\lambda_{\min}(\sigma)}\Big)\right) \\
&\leq \frac{\sqrt{2\epsilon}}{\lambda_{\min}(\sigma)} + (\alpha_0 - 1)\frac{\sqrt{2\epsilon}}{\lambda_{\min}(\sigma)}\exp\left(\eta 2Kt - Kt - \frac{1}{\beta_t}\log\Big(\frac{\lambda_{\max}(\sigma)}{\lambda_{\min}(\sigma)}\Big)\right).
\end{aligned}
$$

If we restrict $\eta \leq \frac{1}{2}$, then the quantity above is less than $\alpha_0\frac{\sqrt{2\epsilon}}{\lambda_{\min}(\sigma)}$, and thus

$$
\lambda_{\max}(\rho_{\sigma,t}^{\beta_t}) \leq \lambda_{\max}(\sigma)\exp\left(\alpha_0\frac{\sqrt{2\epsilon}}{\lambda_{\min}(\sigma)}\right). \tag{4.51}
$$

Next, we prove a lower bound for $\lambda_{\min}(\rho_{\sigma,t}^{\beta,t})$. Similar to (4.50), one can show that

$$\sigma^{\frac{1-\beta_t}{2\beta_t}}\rho_t\sigma^{\frac{1-\beta_t}{2\beta_t}} = \sigma^{\frac{1-\beta_t}{2\beta_t}}(\rho_t-\sigma)\sigma^{\frac{1-\beta_t}{2\beta_t}} + \sigma^{\frac{1}{\beta_t}} \geq \sigma^{\frac{1}{\beta_t}} - \sqrt{2\epsilon_t}\sigma^{\frac{1-\beta_t}{\beta_t}}$$

$$\geq \left(\lambda_{\min}(\sigma)^{\frac{1}{\beta_t}} - \sqrt{2\epsilon_t}\lambda_{\min}(\sigma)^{\frac{1-\beta_t}{\beta_t}}\right)\mathbb{I}_n$$

$$= \lambda_{\min}(\sigma)^{\frac{1}{\beta_t}}\left(1-\frac{\sqrt{2\epsilon_t}}{\lambda_{\min}(\sigma)}\right)\mathbb{I}_n > 0,$$

and hence,

$$\lambda_{\min}(\rho_{\sigma,t}^{\beta_t}) = \lambda_{\min}\left(\left(\sigma^{\frac{1-\beta_t}{2\beta_t}}\rho_t\sigma^{\frac{1-\beta_t}{2\beta_t}}\right)^{\beta_t}\right) \geq \lambda_{\min}(\sigma)\left(1-\frac{\sqrt{2\epsilon_t}}{\lambda_{\min}(\sigma)}\right)^{\beta_t}$$

$$= \lambda_{\min}(\sigma)\exp\left(\beta_t\log\left(1-\frac{\sqrt{2\epsilon}}{\lambda_{\min}(\sigma)}e^{-Kt}\right)\right).$$

We show that the exponent on the right side of the above equation can be bounded from below. Indeed, note that $1-\frac{\sqrt{2\epsilon}}{\lambda_{\min}(\sigma)}e^{-Kt} \geq 1-\frac{\sqrt{2\epsilon}}{\lambda_{\min}(\sigma)} > 0$ and $\log(1-x) \geq -\frac{x}{1-x}$ for $x \in [0,1)$. If we restrict $\eta \leq 1/2$, then

$$0 > \beta_t\log\left(1-\frac{\sqrt{2\epsilon}}{\lambda_{\min}(\sigma)}e^{-Kt}\right) = (1+(\alpha_0-1)e^{\eta 2Kt})\log\left(1-\frac{\sqrt{2\epsilon}}{\lambda_{\min}(\sigma)}e^{-Kt}\right)$$

$$\geq \log\left(1-\frac{\sqrt{2\epsilon}}{\lambda_{\min}(\sigma)}\right) - (\alpha_0-1)e^{\eta 2Kt}\frac{\frac{\sqrt{2\epsilon}}{\lambda_{\min}(\sigma)}e^{-Kt}}{1-\frac{\sqrt{2\epsilon}}{\lambda_{\min}(\sigma)}e^{-Kt}}$$

$$= \log\left(1-\frac{\sqrt{2\epsilon}}{\lambda_{\min}(\sigma)}\right) - (\alpha_0-1)e^{\eta 2Kt-Kt}\frac{\frac{\sqrt{2\epsilon}}{\lambda_{\min}(\sigma)}}{1-\frac{\sqrt{2\epsilon}}{\lambda_{\min}(\sigma)}}$$

$$\geq \log\left(1-\frac{\sqrt{2\epsilon}}{\lambda_{\min}(\sigma)}\right) - (\alpha_0-1)\frac{\frac{\sqrt{2\epsilon}}{\lambda_{\min}(\sigma)}}{1-\frac{\sqrt{2\epsilon}}{\lambda_{\min}(\sigma)}} \geq -\alpha_0\frac{\frac{\sqrt{2\epsilon}}{\lambda_{\min}(\sigma)}}{1-\frac{\sqrt{2\epsilon}}{\lambda_{\min}(\sigma)}}.$$

This implies that

$$\lambda_{\min}(\rho_{\sigma,t}^{\beta_t}) \geq \lambda_{\min}(\sigma)\exp\Big(-\alpha_0 \frac{\frac{\sqrt{2\epsilon}}{\lambda_{\min}(\sigma)}}{1-\frac{\sqrt{2\epsilon}}{\lambda_{\min}(\sigma)}}\Big). \tag{4.52}$$

Combining the estimations (4.51) and (4.52) yields

$$\begin{aligned}\frac{\lambda_{\max}(\rho_{\sigma,t}^{\beta_t})}{\lambda_{\min}(\rho_{\sigma,t}^{\beta_t})} &\leq \frac{\lambda_{\max}(\sigma)\exp\Big(\alpha_0\frac{\sqrt{2\epsilon}}{\lambda_{\min}(\sigma)}\Big)}{\lambda_{\min}(\sigma)\exp\Big(-\alpha_0\frac{\frac{\sqrt{2\epsilon}}{\lambda_{\min}(\sigma)}}{1-\frac{\sqrt{2\epsilon}}{\lambda_{\min}(\sigma)}}\Big)} \\ &= \frac{\lambda_{\max}(\sigma)}{\lambda_{\min}(\sigma)}\exp\Big(\alpha_0\sqrt{2\epsilon}\frac{2\lambda_{\min}(\sigma)-\sqrt{2\epsilon}}{\lambda_{\min}(\sigma)(\lambda_{\min}(\sigma)-\sqrt{2\epsilon})}\Big) =: \Lambda.\end{aligned}$$

This leads to

$$\frac{2\sqrt{e^{\omega_j}\frac{\lambda_{\min}(\rho_{\sigma,t}^{\beta_t})}{\lambda_{\max}(\rho_{\sigma,t}^{\beta_t})}}}{1+e^{\omega_j}\frac{\lambda_{\max}(\rho_{\sigma,t}^{\beta_t})}{\lambda_{\min}(\rho_{\sigma,t}^{\beta_t})}} \geq \frac{2\sqrt{e^{\omega_j}\frac{1}{\Lambda}}}{1+e^{\omega_j}\Lambda}.$$

This finishes the proof of (4.49) with

$$\eta := \min\left(\frac{1}{2}, \min_{j=1}^{J}\left\{\frac{2\sqrt{e^{\omega_j}\frac{1}{\Lambda}}}{1+e^{\omega_j}\Lambda}\right\}\right) > 0.$$

□

4.5.6 Discussion on the quantum comparison theorem

Finally, we would like to comment on the quantum comparison theorem (Prop. 4.32).

Discussion on its proof and the hypercontractivity: The main idea in proving Prop. 4.32 is to verify the *hypercontractivity* in the sense of noncommutative $\mathbb{L}_\alpha$ space (see, e.g., [89, Def. 12]) under the assumption of the quantum LSI. More specifically, suppose $\beta_t = 1 + (\alpha_0 - 1)e^{\eta 2Kt}$ for some constant $\eta > 0$ and ρ_t follows the Lindblad equation, then for the time interval $[0, T]$, where $T = \frac{1}{2K\eta} \log\left(\frac{\alpha_1 - 1}{\alpha_0 - 1}\right)$, we need to show that

$$\left[\mathrm{Tr}\left(\left(\Upsilon_\sigma^{\frac{1-\beta_t}{\beta_t}}(\rho_t)\right)^{\beta_t}\right)\right]^{\frac{1}{\beta_t}} \text{ is a non-increasing function of time } t \in [0, T]. \tag{4.53}$$

The function F_t in the last subsection is simply the logarithm of this function. To see why this requirement (4.53) is closely related to the hypercontractivity in the noncommutative $\mathbb{L}_\alpha$ spaces, let us introduce the relative density $X_t = \Upsilon_\sigma^{-1}(\rho_t)$. The time-evolution of the relative density X_t has the generator $\mathcal{L}$, i.e., $X_t = e^{t\mathcal{L}}(X_0)$, which can be directly verified with the help of (4.9). With these notations and definitions, (4.53) means $\|e^{t\mathcal{L}}(X_0)\|_{\beta_t,\sigma} \leq \|X_0\|_{\alpha_0,\sigma}$, which has the same form as [89, Def. 12]. In a recent paper [14], for primitive Lindblad equations with GNS-detailed balance, a quantum Stroock-Varopoulos inequality (SVI) has been proved in [14, Thm. 14], and its implication (see [14, Cor. 17]) shows that one can pick $\eta_{SVI} = 2\kappa_2/K$ (see (4.17) for the definition of κ_2); the subscript *SVI* is used to emphasize using the quantum Stroock-Varopoulos inequality. Recall that $K = 2\kappa_1$ (4.20). The quantum Stroock-Varopoulos inequality shows that $\kappa_1 \geq \kappa_2$ (see also Thm. 4.18), hence $\eta_{SVI} \leq 1$, which is consistent with our observation in the proof that $\eta \leq 1$. We make some comments on the difference between Prop. 4.32 and the above estimation using the quantum Stroock-Varopoulos inequality:

(i) Using the estimation from the quantum Stroock-Varopoulos inequality, the prefactor $\frac{1}{2K\eta}$ in (4.41), is simply $\frac{1}{4\kappa_2}$; in this chapter, we directly use $K/2 \equiv \kappa_1$ and the information about the spectrum of σ.

(ii) There is a more stringent assumption on the initial condition ρ_0 in Prop. 4.32. Because of such a restriction on initial conditions, even though our proof is related to verify the hypercontractivity, strictly speaking, we haven't really proved the hypercontractivity by solely assuming the quantum LSI (4.4) for $\alpha = 1$.

(iii) Despite the restriction on initial conditions ρ_0, our proof is probably more similar to the classical result for the Fokker-Planck equation case in [29, Lem. 3.4]. Moreover, our proof for the quantum comparison theorem does not require any prior knowledge about noncommutative $\mathbb{L}_\alpha$ spaces.

Comparison with the classical result: By comparing Prop. 4.32 and a similar result for the Fokker-Planck equation (c.f. [29, Lem. 3.4]), it appears that the waiting time T in Prop. 4.32 is much larger than that for the Fokker-Planck equation. More specifically, the parameter $\eta \leq 1$ for the Lindblad equation, and $\eta = 1$ for the Fokker-Planck equation.

4.6 Conclusion

In this chapter, we have extended the gradient flow structure of the quantum relative entropy in [36] to sandwiched Rényi divergences of any order $\alpha \in (0, \infty)$, for primitive Lindblad equations with GNS-detailed balance. The necessary condition for the validity of such a gradient flow structure has been briefly discussed. Furthermore, we have proved the exponential decay of the sandwiched Rényi divergence of any order $\alpha \in (0, \infty)$.

Chapter 5

Conclusion and outlook

In the last three chapters, we have studied Lindblad equations from the physical, numerical, and information theory perspectives. Below are some potential future directions for each perspective, respectively. These questions have been presented in [27, 28, 30].

Physical perspective

It is useful to investigate how to perform further approximations to simplify the hopping coefficients in the LCME (see (2.21) and (2.22)) to obtain a simpler equation.

Numerical perspective

We are curious about whether the SDLR method can be extended to infinite-dimensional Hilbert spaces beyond $\mathbb{C}^n$, because (1) many interesting QMEs (as well as their unraveling schemes) evolve on infinite-dimensional Hilbert spaces; e.g. the Lindblad equation for the Anderson-Holstein model in Chp. 2; (2) many SPDEs are essentially SDEs on infinite-dimensional Hilbert spaces, e.g., stochastic Burgers' equation (3.20). Though finite truncation is a must in the numerical simulation, it is still desirable to study how our scheme could be applied to such SPDEs directly in theory.

Another interesting question is to develop adaptive schemes for the SDLR method with rank adjustment on-the-fly.

Information theory perspective

(More general entropy) In Chp. 4, we have considered the entropy production of the sandwiched Rényi divergence, which is a generalization of the quantum relative entropy. One natural question is whether the gradient flow structure and exponential decay proved for the sandwiched Rényi divergence can be extended to the more general (α, z)-Rényi divergence [9, 34], for certain ranges of (α, z).

(Gap between sufficient and necessary conditions for gradient flow structures) We have mentioned in Sec. 4.4 that there is still a gap between sufficient and necessary conditions to regard Lindblad equations as gradient flow dynamics of sandwiched Rényi divergences (including the quantum relative entropy [37]). Characterizing the QMS with various detailed balance conditions, beyond GNS-detailed balance, might be essential to resolve this issue, i.e., studying different versions of Thm. 4.9 by considering various detailed balance assumptions.

(Quantum Wasserstein distance) In the classical optimal transport theory, Wasserstein distance, introduced to capture the cost of transportation between two probability measures, has been widely studied [139]. Due to its success for classical systems, it is a natural question to explore the notion of *quantum Wasserstein distance*. There are many attempts to define this concept via the Monge formalism [145, 146], the Kantorovich formalism [1, 39, 79, 80, 143], and the Benamou-Brenier formalism [35, 36, 42–44, 87, 142]. The one that is most relevant to us is Benamou-Brenier formalism [15], which offers a dynamical picture to view the Wasserstein distance between any two states ρ_0 and ρ_1 as the minimal Lagrangian along all possible paths connecting ρ_0 and ρ_1. This Lagrangian could be defined via metric tensors in Riemannian manifolds. In [36, Sec. 8], with the variational formalism of primitive Lindblad equations with GNS-detailed balance for the quantum relative entropy, quantum Wasserstein distance is defined via the Benamou-Brenier formalism. With the generalization of this variational formalism to the case of the sandwiched Rényi divergence in Thm. 4.2, one could straightforwardly generalize the quantum Wasserstein distance defined in [36] to a family of quantum (α, q)-Wasserstein distances.

For any order $\alpha \in (0, \infty)$ and power $q \geq 1$, a quantum (α, q)-Wasserstein distance is defined in the following way: for any $\rho_0, \rho_1 \in \mathfrak{D}$,

$$\mathcal{W}_{\alpha,q}(\rho_0, \rho_1) := \inf_{\gamma_\cdot} \left(\int_0^1 \left(g_{\alpha,\gamma_s,\mathcal{L}} \left(\dot{\gamma}_s, \dot{\gamma}_s \right) \right)^{\frac{q}{2}} \mathrm{d}s \right)^{\frac{1}{q}},$$

where the infimum is taken over all smooth paths $\gamma_\cdot : [0, 1] \to \mathfrak{D}^+$ that connect ρ_0 and ρ_1. More specifically, $\gamma_0 = \rho_0$, $\gamma_1 = \rho_1$ and $\gamma_s \in \mathfrak{D}^+$ for all $s \in (0, 1)$. The study of the properties of this quantum Wasserstein distance might be an interesting topic for future works.

(Lindblad equation with energy-conservation term) In general, the Lindblad equation has both energy-conservation term and dissipative term; see (1.1) and (1.2). The Hamiltonian H, in general, does not vanish. Therefore, it is also an interesting question to explore how we could generalize our results to such general Lindblad equations with non-trivial H.

An exact construction of gradient flow structures in a Riemannian manifold (as in Sec. 4.2.1) does not seem to be possible, even for classical kinetic Fokker-Planck equations. However, for a generalized Kramers equation (probably regarded as the classical analog of Lindblad equations with a non-trivial Hamiltonian term), some JKO schemes have already existed in [61]. As far as we know, currently, there is no any JKO scheme for Lindblad equations, which is probably also an interesting question to study, due to the current active research on quantum Wasserstein distance, as we mentioned above.

As for the convergence rate to the equilibrium, there is another approach known as the *hypocoercivity* [54, 55, 138] for kinetic Fokker-Planck equations. We are curious about whether this approach admits a quantum extension to help study the convergence of Lindblad equations with a non-trivial Hamiltonian term.

Bibliography

[1] J. Agredo and F. Fagnola, *On quantum versions of the classical Wasserstein distance*, Stochastics **89** (2017), no. 6-7, 910–922.

[2] R. Alicki, *On the detailed balance condition for non-Hamiltonian systems*, Rep. Math. Phys. **10** (1976), no. 2, 249–258.

[3] R. Alicki and R. Kosloff, *Introduction to quantum thermodynamics: History and prospects*, Thermodynamics in the Quantum Regime: Fundamental Aspects and New Directions, 2018, pp. 1–33.

[4] L. Ambrosio, N. Gigli, and G. Savare, *Gradient flows: In metric spaces and in the space of probability measures*, 2nd ed., Lectures in Mathematics. ETH Zürich, Birkhäuser Basel, 2008.

[5] P. W. Anderson, *Localized magnetic states in metals*, Phys. Rev. **124** (1961), no. 1, 41–53.

[6] A. C. Antoulas, D. C. Sorensen, and S. Gugercin, *A survey of model reduction methods for large-scale systems*, Contemp. Math. **280** (2001), 193–219.

[7] G. Argentieri, F. Benatti, R. Floreanini, and M. Pezzutto, *Violations of the second law of thermodynamics by a non-completely positive dynamics*, EPL **107** (2014), no. 5, 50007.

[8] A. Arnold, J. A. Carrillo, L. Desvillettes, J. Dolbeault, A. Jüngel, C. Lederman, P. A. Markowich, G. Toscani, and C. Villani, *Entropies and equilibria of many-particle systems: An essay on recent research*, Monatsh. Math. **142** (2004), no. 1, 35–43.

[9] K. M. R. Audenaert and N. Datta, *α-z-Rényi relative entropies*, J. Math. Phys. **56** (2015), no. 2, 022202.

[10] H. Babaee, M. Choi, T. P. Sapsis, and G. E. Karniadakis, *A robust bi-orthogonal/dynamically-orthogonal method using the covariance pseudo-inverse with application to stochastic flow problems*, J. Comput. Phys. **344** (2017), 303–319.

[11] D. Bakry, I. Gentil, and M. Ledoux, *Analysis and geometry of Markov diffusion operators*, Springer, Cham, 2014.

[12] D. Bakry and M. Émery, *Diffusions hypercontractives*, Séminaire de probabilités de Strasbourg **19** (1985), 177–206 (fr).

[13] S. Beigi, *Sandwiched Rényi divergence satisfies data processing inequality*, J. Math. Phys. **54** (2013), no. 12, 122202.

[14] S. Beigi, N. Datta, and C. Rouzé, *Quantum reverse hypercontractivity: its tensorization and application to strong converses*, 2018. arXiv:1804.10100.

[15] J.-D. Benamou and Y. Brenier, *A computational fluid mechanics solution to the Monge-Kantorovich mass transfer problem*, Numer. Math. **84** (2000), no. 3, 375–393.

[16] F. Benatti and H. Narnhofer, *Entropy behaviour under completely positive maps*, Lett. Math. Phys. **15** (1988), no. 4, 325–334.

[17] M. Berta, O. Fawzi, and M. Tomamichel, *On variational expressions for quantum relative entropies*, Lett. Math. Phys. **107** (2017), no. 12, 2239–2265.

[18] K. Blum, *Density matrix theory and applications*, Springer-Verlag Berlin Heidelberg, 2012.

[19] F. Brandão, M. Horodecki, N. Ng, J. Oppenheim, and S. Wehner, *The second laws of quantum thermodynamics*, Proc. Natl. Acad. Sci. U.S.A. **112** (2015), no. 11, 3275–3279.

[20] H.-P. Breuer, *Quantum jumps and entropy production*, Phys. Rev. A **68** (2003), no. 3, 032105.

[21] H.-P. Breuer, E.-M. Laine, J. Piilo, and B. Vacchini, *Colloquium: Non-markovian dynamics in open quantum systems*, Rev. Mod. Phys. **88** (2016), no. 2, 021002.

[22] H.-P. Breuer and F. Petruccione, *The theory of open quantum systems*, Oxford University Press, 2007.

[23] H.-P. Breuer, B. Kappler, and F. Petruccione, *Stochastic wave-function method for non-Markovian quantum master equations*, Phys. Rev. A **59** (1999), no. 2, 1633–1643.

[24] A. Bruch, M. Thomas, S. Viola Kusminskiy, F. von Oppen, and A. Nitzan, *Quantum thermodynamics of the driven resonant level model*, Phys. Rev. B **93** (2016), no. 11, 115318.

[25] T. A. Brun, *Continuous measurements, quantum trajectories, and decoherent histories*, Phys. Rev. A **61** (2000), no. 4, 042107.

[26] D. Burgarth, G. Chiribella, V. Giovannetti, P. Perinotti, and K. Yuasa, *Ergodic and mixing quantum channels in finite dimensions*, New J. Phys. **15** (2013), 073045.

[27] Y. Cao and J. Lu, *Lindblad equation and its semiclassical limit of the Anderson-Holstein model*, J. Math. Phys. **58** (2017), no. 12, 122105.

[28] ———, *Stochastic dynamical low-rank approximation method*, J. Comput. Phys. **372** (2018), 564–586.

[29] Y. Cao, J. Lu, and Y. Lu, *Exponential decay of Rényi divergence under Fokker–Planck equations*, J. Stat. Phys. **176** (2019), no. 5, 1172–1184.

[30] ———, *Gradient flow structure and exponential decay of the sandwiched Rényi divergence for primitive Lindblad equations with GNS-detailed balance*, J. Math. Phys. **60** (2019), no. 5, 052202.

[31] Y. Cao and J. Lu, *Tensorization of the strong data processing inequality for quantum chi-square divergences*, Quantum **3** (2019), 199.

[32] Y. Cao, J. Lu, and L. Wang, *On explicit L^2-convergence rate estimate for underdamped Langevin dynamics*, 2019. arXiv: 1908.04746.

[33] ———, *Complexity of randomized algorithms for underdamped Langevin dynamics*, 2020. arXiv: 2003.09906.

[34] E. A. Carlen, R. L. Frank, and E. H. Lieb, *Inequalities for quantum divergences and the Audenaert–Datta conjecture*, J. Phys. A: Math. Theor. **51** (2018), no. 48, 483001.

[35] E. A. Carlen and J. Maas, *An analog of the 2-Wasserstein metric in non-commutative probability under which the Fermionic Fokker–Planck equation is gradient flow for the entropy*, Commun. Math. Phys. **331** (2014), no. 3, 887–926.

[36] ———, *Gradient flow and entropy inequalities for quantum Markov semigroups with detailed balance*, J. Funct. Anal. **273** (2017), no. 5, 1810 –1869.

[37] ———, *Non-commutative calculus, optimal transport and functional inequalities in dissipative quantum systems*, J. Stat. Phys. **178** (2020), no. 2, 319–378.

[38] H. Carmichael, *An open systems approach to quantum optics*, Lecture Notes in Physics Monographs, vol. 18, Springer Berlin Heidelberg, 1993.

[39] S. Chakrabarti, Y. Huang, T. Li, S. Feizi, and X. Wu, *Quantum Wasserstein generative adversarial networks*, Advances in Neural Information Processing Systems 32, 2019, pp. 6781–6792.

[40] A. Chatterjee, *An introduction to the proper orthogonal decomposition*, Curr. Sci. **78** (2000), no. 7, 808–817.

[41] H.-T. Chen, G. Cohen, A. J. Millis, and D. R. Reichman, *Anderson-Holstein model in two flavors of the noncrossing approximation*, Phys. Rev. B **93** (2016), no. 17, 174309.

[42] Y. Chen, W. Gangbo, T. T. Georgiou, and A. Tannenbaum, *On the matrix Monge–Kantorovich problem*, Eur. J. Appl. Math. (2019), 1–27.

[43] Y. Chen, T. T. Georgiou, and A. Tannenbaum, *Matrix optimal mass transport: A quantum mechanical approach*, IEEE Trans. Autom. Control **63** (2018), no. 8, 2612–2619.

[44] ______, *Wasserstein geometry of quantum states and optimal transport of matrix-valued measures*, Emerging Applications of Control and Systems Theory, 2018, pp. 139–150.

[45] M. Cheng, T. Y. Hou, and Z. Zhang, *A dynamically bi-orthogonal method for time-dependent stochastic partial differential equations I: Derivation and algorithms*, J. Comput. Phys. **242** (2013), 843–868.

[46] ______, *A dynamically bi-orthogonal method for time-dependent stochastic partial differential equations II: Adaptivity and generalizations*, J. Comput. Phys. **242** (2013), 753 –776.

[47] T. M. Cover and J. A. Thomas, *Elements of information theory*, Wiley-Interscience, 2006.

[48] T. L. Curtright, D. B. Fairlie, and C. K. Zachos, *A concise treatise on quantum mechanics in phase space*, World Scientific Publishing Co. Pte. Ltd, 2014.

[49] A. S. Dalalyan, *Theoretical guarantees for approximate sampling from smooth and log-concave densities*, J. R. Stat. Soc.: B **79** (2017), no. 3, 651–676.

[50] J. Dalibard, Y. Castin, and K. Mølmer, *Wave-function approach to dissipative processes in quantum optics*, Phys. Rev. Lett. **68** (1992), no. 5, 580–583.

[51] N. Datta and C. Rouzé, *Relating relative entropy, optimal transport and Fisher information: A quantum HWI inequality*, Ann. Henri Poincaré (2020).

[52] E. B. Davies, *Markovian master equations*, Comm. Math. Phys. **39** (1974), no. 2, 91–110.

[53] P. Diaconis and L. Saloff-Coste, *Logarithmic Sobolev inequalities for finite Markov chains*, Ann. Appl. Probab. **6** (1996), no. 3, 695–750.

[54] J. Dolbeault, C. Mouhot, and C. Schmeiser, *Hypocoercivity for kinetic equations with linear relaxation terms*, C. R. Math. **347** (2009), no. 9, 511–516.

[55] ______, *Hypocoercivity for linear kinetic equations conserving mass*, Trans. Amer. Math. Soc. **367** (2015), no. 6, 3807–3828.

[56] B. Dong, G. H. Ding, and X. L. Lei, *Full counting statistics of a single-molecule quantum dot*, Phys. Rev. B **88** (2013), no. 7, 075414.

[57] W. Dou, A. Nitzan, and J. E. Subotnik, *Frictional effects near a metal surface*, J. Chem. Phys. **143** (2015), no. 5, 054103.

[58] ______, *Surface hopping with a manifold of electronic states. II. Application to the many-body Anderson-Holstein model*, J. Chem. Phys. **142** (2015), no. 8, 084110.

[59] ______, *Surface hopping with a manifold of electronic states. III. Transients, broadening, and the Marcus picture*, J. Chem. Phys. **142** (2015), no. 23, 234106.

[60] W. Dou, C. Schinabeck, M. Thoss, and J. E. Subotnik, *A broadened classical master equation approach for treating electron-nuclear coupling in non-equilibrium transport*, J. Chem. Phys. **148** (2018), no. 10, 102317.

[61] M. H. Duong, M. A. Peletier, and J. Zimmer, *Conservative-dissipative approximation schemes for a generalized Kramers equation*, Math. Methods Appl. Sci. **37** (2014), no. 16, 2517–2540.

[62] F. Elste, G. Weick, C. Timm, and F. von Oppen, *Current-induced conformational switching in single-molecule junctions*, Appl. Phys. A **93** (2008), no. 2, 345–354.

[63] M. Erbar, *The heat equation on manifolds as a gradient flow in the Wasserstein space*, Ann. l'Inst. Henri Poincaré, Probab. Stat. **46** (2010), no. 1, 1–23.

[64] M. Erbar and J. Maas, *Gradient flow structures for discrete porous medium equations*, Discrete Contin. Dyn. Syst. Ser. A **34** (2014), no. 4, 1355–1374.

[65] M. Esposito and M. Galperin, *Transport in molecular states language: Generalized quantum master equation approach*, Phys. Rev. B **79** (2009), no. 20, 205303.

[66] ______, *Self-consistent quantum master equation approach to molecular transport*, J. Phys. Chem. C **114** (2010), no. 48, 20362–20369.

[67] F. Fagnola and R. Rebolledo, *Entropy production for quantum Markov semigroups*, Commun. Math. Phys. **335** (2015), no. 2, 547–570.

[68] F. Fagnola and V. Umanità, *Generators of detailed balance quantum Markov semigroups*, Infin. Dimens. Anal. Quantum. Probab. Relat. Top. **10** (2007), no. 3, 335–363.

[69] ______, *Generators of KMS symmetric Markov semigroups on $\mathcal{B}(\mathrm{h})$ symmetry and quantum detailed balance*, Commun. Math. Phys. **298** (2010), no. 2, 523–547.

[70] R. L. Frank and E. H. Lieb, *Monotonicity of a relative Rényi entropy*, J. Math. Phys. **54** (2013), no. 12, 122201.

[71] K. Fujii, *Quantum damped harmonic oscillator*, Advances in quantum mechanics, 2013.

[72] M. Galperin, M. A. Ratner, A. Nitzan, and A. Troisi, *Nuclear coupling and polarization in molecular transport junctions: Beyond tunneling to function*, Science **319** (2008), no. 5866, 1056–1060.

[73] M. Galperin, A. Nitzan, and M. A. Ratner, *Resonant inelastic tunneling in molecular junctions*, Phys. Rev. B **73** (2006), no. 4, 045314.

[74] ______, *Inelastic effects in molecular junctions in the Coulomb and Kondo regimes: Nonequilibrium equation-of-motion approach*, Phys. Rev. B **76** (2007), no. 3, 035301.

[75] M. Galperin, M. A. Ratner, and A. Nitzan, *Inelastic electron tunneling spectroscopy in molecular junctions: Peaks and dips*, J. Chem. Phys. **121** (2004), no. 23, 11965–11979.

[76] ______, *Molecular transport junctions: vibrational effects*, J. Phys.: Condens. Matter **19** (2007), no. 10, 103201.

[77] A. W. Ghosh, P. S. Damle, S. Datta, and A. Nitzan, *Molecular electronics: Theory and device prospects*, MRS Bull. **29** (2004), no. 6, 391–395.

[78] N. Gisin and I. C. Percival, *The quantum-state diffusion model applied to open systems*, J. Phys. A: Math. Gen. **25** (1992), no. 21, 5677.

[79] F. Golse, C. Mouhot, and T. Paul, *On the mean field and classical limits of quantum mechanics*, Commun. Math. Phys. **343** (2016), no. 1, 165–205.

[80] F. Golse and T. Paul, *Wave packets and the quadratic Monge-Kantorovich distance in quantum mechanics*, C. R. Math. **356** (2018), no. 2, 177–197.

[81] V. Gorini, A. Kossakowski, and E. C. G. Sudarshan, *Completely positive dynamical semigroups of N-level systems*, J. Math. Phys. **17** (1976), no. 5, 821–825.

[82] L. Gross, *Hypercontractivity and logarithmic Sobolev inequalities for the Clifford-Dirichlet form*, Duke Math. J. **42** (1975), no. 3, 383–396.

[83] ______, *Logarithmic Sobolev inequalities*, Am. J. Math. **97** (1975), no. 4, 1061–1083.

[84] A. Guionnet and B Zegarlinksi, *Lectures on logarithmic Sobolev inequalities*, Séminaire de probabilités XXXVI, 2003, pp. 1–134.

[85] D. J. Higham, *Runge-Kutta type methods for orthogonal integration*, Appl. Numer. Math. **22** (1996), no. 1, 217–223.

[86] T. Holstein, *Studies of polaron motion: Part I. The molecular-crystal model*, Ann. Phys. **8** (1959), no. 3, 325–342.

[87] D. F. Hornshaw, *L^2-Wasserstein distances of tracial W^*-algebras and their disintegration problem*, 2018. arXiv:1806.01073.

[88] R. Jordan, D. Kinderlehrer, and F. Otto, *The variational formulation of the Fokker–Planck equation*, SIAM J. Math. Anal. **29** (1998), no. 1, 1–17.

[89] M. J. Kastoryano and K. Temme, *Quantum logarithmic Sobolev inequalities and rapid mixing*, J. Math. Phys. **54** (2013), no. 5, 052202.

[90] J. Koch and F. von Oppen, *Franck-Condon blockade and giant fano factors in transport through single molecules*, Phys. Rev. Lett. **94** (2005), no. 20, 206804.

[91] O. Koch and C. Lubich, *Dynamical low-rank approximation*, SIAM J. Matrix Anal. Appl. **29** (2007), no. 2, 434–454.

[92] I. Kondov, U. Kleinekathöfer, and M. Schreiber, *Stochastic unraveling of Redfield master equations and its application to electron transfer problems*, J. Chem. Phys. **119** (2003), no. 13, 6635–6646.

[93] D. S Kosov, T. Prosen, and B. Žunkovič, *Lindblad master equation approach to superconductivity in open quantum systems*, J. Phys. A: Math. Theor. **44** (2011), no. 46, 462001.

[94] S. Kullback and R. A. Leibler, *On information and sufficiency*, Ann. Math. Stat. **22** (1951), no. 1, 79–86.

[95] C. Le Bris and P. Rouchon, *Low-rank numerical approximations for high-dimensional Lindblad equations*, Phys. Rev. A **87** (2013), no. 2, 022125. Extended version available at arXiv:1207.4580.

[96] C. Le Bris, P. Rouchon, and J. Roussel, *Adaptive low-rank approximation and denoised Monte Carlo approach for high-dimensional Lindblad equations*, Phys. Rev. A **92** (2015), no. 6, 062126.

[97] G. Lindblad, *Completely positive maps and entropy inequalities*, Commun. Math. Phys. **40** (1975), no. 2, 147–151.

[98] ———, *On the generators of quantum dynamical semigroups*, Commun. Math. Phys. **48** (1976), no. 2, 119–130.

[99] J. Lott, *On tangent cones in Wasserstein space*, Proc. Am. Math. Soc. **145** (2017), no. 7, 3127–3136.

[100] J. Maas, *Gradient flows of the entropy for finite Markov chains*, J. Funct. Anal. **261** (2011), no. 8, 2250–2292.

[101] J. M. McCracken, *Quantum channel negativity as a measure of system-bath coupling and correlation*, Phys. Rev. A **88** (2013), no. 3, 032103.

[102] A. Mitra, I. Aleiner, and A. J. Millis, *Phonon effects in molecular transistors: Quantal and classical treatment*, Phys. Rev. B **69** (2004), no. 24, 245302.

[103] ———, *Semiclassical analysis of the nonequilibrium local polaron*, Phys. Rev. Lett. **94** (2005), no. 7, 076404.

[104] M. Moodley and F. Petruccione, *Stochastic wave-function unraveling of the generalized Lindblad master equation*, Phys. Rev. A **79** (2009), no. 4, 042103.

[105] M. Mosonyi and T. Ogawa, *Quantum hypothesis testing and the operational interpretation of the quantum Rényi relative entropies*, Commun. Math. Phys. **334** (2015), no. 3, 1617–1648.

[106] L. Mühlbacher and E. Rabani, *Real-time path integral approach to nonequilibrium many-body quantum systems*, Phys. Rev. Lett. **100** (2008), no. 17, 176403.

[107] A. Müller-Hermes and D. S. Franca, *Sandwiched Rényi convergence for quantum evolutions*, Quantum **2** (2018), 55.

[108] A. Müller-Hermes and D. Reeb, *Monotonicity of the quantum relative entropy under positive maps*, Ann. Henri Poincaré **18** (2017), no. 5, 1777–1788.

[109] A. Müller-Hermes, D. Stilck França, and M. M. Wolf, *Entropy production of doubly stochastic quantum channels*, J. Math. Phys. **57** (2016), no. 2, 022203.

[110] ———, *Relative entropy convergence for depolarizing channels*, J. Math. Phys. **57** (2016), no. 2, 022202.

[111] M. Müller-Lennert, F. Dupuis, O. Szehr, S. Fehr, and M. Tomamichel, *On quantum Rényi entropies: A new generalization and some properties*, J. Math. Phys. **54** (2013), no. 12, 122203.

[112] E. Musharbash, F. Nobile, and T. Zhou, *Error analysis of the dynamically orthogonal approximation of time dependent random PDEs*, SIAM J. Sci. Comput. **37** (2015), no. 2, A776–A810.

[113] R. Olkiewicz and B. Zegarlinski, *Hypercontractivity in noncommutative L_p spaces*, J. Funct. Anal. **161** (1999), no. 1, 246–285.

[114] F. Otto and C. Villani, *Generalization of an inequality by Talagrand and links with the logarithmic Sobolev inequality*, J. Funct. Anal. **173** (2000), no. 2, 361 –400.

[115] F. Otto, *The geometry of dissipative evolution equations: The porous medium equation*, Commun. Partial Differ. Equ. **26** (2001), no. 1-2, 101–174.

[116] W. Ouyang, W. Dou, and J. E. Subotnik, *Surface hopping with a manifold of electronic states. I. Incorporating surface-leaking to capture lifetimes*, J. Chem. Phys. **142** (2015), no. 8, 084109.

[117] H. C. Ozen and G. Bal, *A dynamical polynomial chaos approach for long-time evolution of SPDEs*, J. Comput. Phys. **343** (2017), 300–323.

[118] G. A. Pavliotis, *Stochastic processes and applications: Diffusion processes, the Fokker-Planck and Langevin equations*, Springer-Verlag New York, 2014.

[119] P. Pechukas, *Reduced dynamics need not be completely positive*, Phys. Rev. Lett. **73** (1994), no. 8, 1060–1062.

[120] I. Percival, *Quantum state diffusion*, Cambridge University Press, 1998.

[121] D. Petz, *Quasi-entropies for finite quantum systems*, Rep. Math. Phys. **23** (1986), no. 1, 57–65.

[122] M. Raginsky and I. Sason, *Concentration of measure inequalities in information theory, communications, and coding*, Found. Trends Commun. Inf. Theory **10** (2013), no. 1-2, 1–246.

[123] C. Rouzé and N. Datta, *Concentration of quantum states from quantum functional and transportation cost inequalities*, J. Math. Phys. **60** (2019), no. 1, 012202.

[124] D. A. Ryndyk, *Theory of quantum transport at nanoscale electronic resource: An introduction*, Springer International Publishing, 2016.

[125] A. Rényi, *On measures of entropy and information*, Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Contributions to the Theory of Statistics, 1961, pp. 547–561.

[126] T. P. Sapsis, *Dynamically orthogonal field equations for stochastic fluid flows and particle dynamics*, Thesis, 2011.

[127] T. P. Sapsis and P. F. J. Lermusiaux, *Dynamically orthogonal field equations for continuous stochastic dynamical systems*, Phys. D: Nonlinear Phenom. **238** (2009), no. 23, 2347 –2360.

[128] ______ , *Dynamical criteria for the evolution of the stochastic dimensionality in flows with uncertainty*, Phys. D: Nonlinear Phenom. **241** (2012), no. 1, 60 –76.

[129] A. Shaji and E. C. G. Sudarshan, *Who's afraid of not completely positive maps?*, Phys. Lett. A **341** (2005), no. 1, 48 –54.

[130] R. Shankar, *Principles of quantum mechanics*, Springer, 1994.

[131] C. E. Shannon, *A mathematical theory of communication*, Bell Syst. Tech. J. **27** (1948), no. 3, 379–423.

[132] B. W. Shore and P. L. Knight, *The Jaynes-Cummings Model*, J. Mod. Opt. **40** (1993), no. 7, 1195–1238.

[133] H. Spohn, *Entropy production for quantum dynamical semigroups*, J. Math. Phys. **19** (1978), no. 5, 1227–1230.

[134] K. Temme, M. J. Kastoryano, M. B. Ruskai, M. M. Wolf, and F. Verstraete, *The χ^2-divergence and mixing times of quantum Markov processes*, J. Math. Phys. **51** (2010), no. 12, 122201.

[135] J. C. Tully, *Molecular dynamics with electronic transitions*, J. Chem. Phys. **93** (1990), no. 2, 1061–1071.

[136] S. Vempala and A. Wibisono, *Rapid convergence of the Unadjusted Langevin Algorithm: Isoperimetry suffices*, Advances in Neural Information Processing Systems 32, 2019, pp. 8094–8106.

[137] F. Verstraete, M. M. Wolf, and J. I. Cirac, *Quantum computation and quantum-state engineering driven by dissipation*, Nat. Phys. **5** (2009), no. 9, 633–636.

[138] C. Villani, *Hypocoercivity*, Memoirs of the American Mathematical Society **202** (2009), no. 950.

[139] ______ , *Optimal transport: Old and new*, Springer-Verlag Berlin Heidelberg, 2009.

[140] E. Wigner, *On the quantum correction for thermodynamic equilibrium*, Phys. Rev. **40** (1932), no. 5, 749–759.

[141] M. M. Wilde, A. Winter, and D. Yang, *Strong converse for the classical capacity of entanglement-breaking and Hadamard channels via a sandwiched Rényi relative entropy*, Commun. Math. Phys. **331** (2014), no. 2, 593–622.

[142] M. Wirth, *A noncommutative transport metric and symmetric quantum Markov semigroups as gradient flows of the entropy*, 2018. arXiv:1808.05419.

[143] N. Yu, L. Zhou, S. Ying, and M. Ying, *Quantum earth mover's distance, no-go quantum Kantorovich-Rubinstein theorem, and quantum marginal problem*, 2018. arXiv:1803.02673.

[144] M. Zworski, *Semiclassical analysis*, American Mathematical Society, 2012. Graduate Studies in Mathematics, Volume 138.

[145] K. Zyczkowski and W. Slomczynski, *The Monge distance between quantum states*, J. Phys. A: Math. Gen. **31** (1998), no. 45, 9095.

[146] ______, *The Monge metric on the sphere and geometry of quantum states*, J. Phys. A: Math. Gen. **34** (2001), no. 34, 6689–6722.

Biography

Yu Cao received his Bachelor's degree in Computing Mathematics at City University of Hong Kong in 2015. He will receive the Ph.D. degree in Mathematics at Duke University in May 2020. From this fall, he will be a Courant Instructor at the Courant Institute at New York University.