Exploring quantum many-body systems from the viewpoints of quantum computing and spectroscopy

Jinzhao Sun
Jesus College
University of Oxford

A thesis submitted for the degree of
Doctor of Philosophy

Trinity 2022

This thesis is dedicated to my family

Abstract

Understanding quantum many-body behaviours such as exotic phases and spectroscopic properties in quantum materials and molecular systems is a long-standing problem of both fundamental and practical interest in quantum physics. This understanding provides insights into the true underlying physics of quantum many-body systems, aids in the prediction of the microscopic and macroscopic properties of those systems, and also advances the rational design and synthesis of novel materials. However, our ability to understand quantum many-body behaviours has hitherto been limited, due to the excessive demands imposed on classical computing by the inherent complexity of describing and analysing those behaviours. While the advent of quantum computing has opened up new possibilities for examining these questions, the current generation of quantum technology does not yet present a feasible, standalone way to solve the above problem. However, a fusion of classical and quantum approaches could arguably provide a viable way of exploring interesting quantum phenomena. The central objective of this thesis is to achieve such a synthesis in practice, and to establish a corresponding framework for the study of quantum many-body systems. One area of particular interest is the intersection between quantum computing and spectroscopy, specifically in terms of the latter's potential to greatly assist in the investigation of quantum many-body phenomena.

Quantum many-body problems in general can be divided into two classes, static and dynamic problems, which correspond to the estimation of eigenstate properties (such as eigenenergies and order parameters in different phases), and dynamical properties (such as response to an external field). In Part II, I present a number of approaches to solve these static and dynamic problems. I initially establish a quantum computing framework based on hybrid quantum-classical tensor networks, which incorporate the inherent advantages of classical tensor networks and quantum computing to represent the quantum system. I then demonstrate how eigenstate properties can be estimated by a randomised linear-combination-of-unitary method, termed algorithmic cooling, with at most one ancillary qubit; this achieves a logarithmic circuit complexity $\mathcal{O}(\log(1/\varepsilon))$ with respect to precision ε in eigenstate property estimation, and reaches the Heisenberg limit in eigenenergy estimation. Turning to dynamic problems, I present an adaptive product formula approach to construct a low-depth

quantum circuit for quantum dynamics simulation. I further show how to enable large-scale dynamics simulation based on hybrid tensor networks, followed by a perturbative approach to simulating quantum many-body dynamics.

In Part III, I first demonstrate how spectroscopic features of quantum systems can be probed. Equipped with the framework and methods established and developed in this thesis, I study quantum many-body phenomena, and excitation spectra in interacting bosons, fermions, and quantum spins through numerics and experiments. In Part IV, the quantum resources required for the application of quantum computing to realistic problems in the near-future are assessed, together with the challenges that accompany such application. This encompasses a discussion of the estimated resources needed for estimating eigenstate properties of spins, fermions and molecules, in respect of both noisy quantum processors and fault-tolerant quantum computers. I then address some of the inherent challenges of using near-term noisy quantum devices, such as encountering unavoidable quantum process errors and statistical errors, by applying error mitigation, and efficient grouping measurement schemes proposed in this thesis. Finally, I conclude with a few remarks on the development of quantum computing in solving quantum many-body systems, and I pose outlooks for further research in this field.

Acknowledgements

I wish to express my deep gratitude to my supervisors Andrew Boothroyd and Vlatko Vedral. I have been extremely fortunate to be supervised by two great scientists, who guide me to pursue the research in physics. I will forever be indebted to my supervisors for starting me off on a journey of physics, and for giving me the privilege to explore in interesting and new fields. My thesis and DPhil research would not have been possible without their incredible help and support. Andrew; thank you for your guidance, feedback, and suggestions about science and careers throughout my study at Oxford. Thank you for navigating me to research on quantum materials and spectroscopy. I have benefited and enjoyed a lot from working with you. Vlatko; thank you for always being willing, excited, and supportive to discuss new research ideas and explore interesting physics. Thank you for continuously giving me the greatest support and encouragement since my DPhil study. It has been my great honour and fortune to get entangled with you.

I am very fortunate to have been studying at University of Oxford, where I have met so many excellent fellow academics and friends. I am very lucky that I have met Xiao Yuan at the starting point of my DPhil study. Xiao; thank you for supervising me on a lot of projects, and for giving me infinite supports over the last years. Working with you is always enjoyable, and you really light up my way forward in research. I am incredibly fortunate that I have this fantastic opportunity to work with you. Xiao is an excellent friend and is always considerate, positive, enthusiastic, trustworthy, and sincere. I will forever cherish our time-invariant friendship and the wonderful time we have ever had at Oxford, Beijing, Shenzhen and many places in the world.

I am immensely grateful to my college advisor Sarah Rugheimer. Sarah; thank you for giving me so many valuable advices throughout my study at Oxford, especially at my decision time, as well as for your kind care over the past four years. My college life is lightened up by you. I also thank Yulin Chen, my college advisor at my final year at Jesus College, for his kind and helpful advice on my career development. I am also grateful to be a member of Jesus College. I would like to thank our college principal, Nigel Shadbolt, for always taking care of my academic and social life, and offering many supports for my college life.

I am grateful to meet my collaborators and friends at Oxford. I wish to thank Tianhan Liu, who is an expert in condensed matter physics and always envisions to tackle big issues in this area. We have had many discussions about physics, quantum, superconductivity, topology as well as so many related topics. Tianhan is my close friend and is always supportive of me. We frequently had dinner and chatted about interesting things at Oxford and Beijing, which has left me with wonderful memories.

I would like to thank Suguru Endo. I have had a lot of close, deep, and fascinating collaboration with Suguru. Suguru always have countless insightful ideas, and he is really a nice advisor as well as a genius in quantum computing. The discussion with you was always inspiring, and I really enjoyed working with you. I also thank my big brother at Oxford, Heng Shen, for your sincereness, kindness, continuous support and encouragement since I started my PhD.

I would like to thank my group members and colleagues at Oxford. I am grateful to Tristan Farrow for supporting me a lot since I came to Oxford, leading me on the track, and offering generous help at my hard time. I wish to thank Jian-Rui Soh for the mentoring, training as well as the excellent collaboration ever since. I have learned so many experimental skills and theories of magnetism from you. I would like to thank Tian Zhang for always being helpful, kind, and sincere to me. I thank my DPhil co-starter, Aditya Iyer. We have had long-range strong interactions over the last four years. It is lucky and great to join the group at the same time with you. I also thank Lucia Estevez for doing such a great work in our collaboration and inspiring my a lot in research. I also thank my group members, Chiara Marletto, Christian Schilling, David Felce, Hanlin Nie, Samuel Kuypers, Benjamin Yadin, Anicet Tibau Vidal, Siobhan Tobin, Dharmalingam Prabhakaran, and Felix Tennie. It is great to work with all of you. I would also like to thank the director of graduate studies, Arzhang Ardvan, for giving me valuable suggestions for the research direction.

I am grateful to have had many fantastic collaborations during my DPhil study. I wish to thank Pei Zeng. He is an excellent researcher and has solid background in quantum information and quantum computing. I appreciate the close collaboration with you. I wish to thank my undergraduate advisor, Yuan Li; thank you for inspiring me a lot for my undergraduate research on superconductivity and always being helpful at my early stage of research career. I would like to thank my close collaborators: Ying Li, Simon Benjamin, Man-Hong Yung, Chenbing Wang, Qi Huang, Dingshun lv, Partrick Hayden, Wenjun Yu, Xiao-Ming Zhang, Zeyao Han, Takahiro Tsunoda, Bujiao Wu, Xiao-si Xu, Weitang Li, Yifei Huang, Yukun Zhang, Ting Zhang, Zigeng Huang, Qi Zhao, You Zhou, Junning Li, Xiaowei Li, Kun Liu, Xiangrun Bao, Changsu Cao, Junyu Liu, Ming Gong, Zijian Zhang, and He Lu. I would also thank Alessandro

Bombardi, Duc Le, Tatiana Guidi, Helen Walker, Navid Qureshi, Sonia Francoual, Jennifer Sears, Pascal Manuel, and Fabio Orlandi for giving the supports at the beamline at Diamond, ISIS, ILL, and DESY. Working with all of you is my great fortune.

I would also like to thank my colleagues for the insightful discussions and friendships: Mingwei Ma, Lichen Wang, Shuxiang Cao, Jiahao Chen, Jiawei Dai, Haoxiong Yan, Zhenyu Cai, Tianfeng Feng, Kecheng Liu, Chengcai Li, Wenjie Chen, Chao Yin, Guoqing Wang, Nan Sheng, Tongyang Li, Minshuai Wang, Wenjie Chen, Tyson Jones, and Jianwei Wang. I would like to thank Banghai Wang for hosting me in Guang Dong. I am also grateful to my colleagues at Peking University: Xiaoyang Wang, Yusen Wu, Kecheng Liu, Chaojian Li, Qiming Ding, Junxiang Huang, Yiming Huang, Ying Li, Huiping Lin, and Peilin Zheng.

I would like to thank all my friends at Jesus College. Thank my college roommates Wenyuan Zhang for the wonderful two years at 121 Woodstock Road and 11 Ship Street. Thank Li Shen for his support and valuable advices. Thank Jiahe Cui, Yongchen Wang, Zhaotian Chi, Zihan Yu, and Sharon Chang for our good memory at Jesus College, at 121 Woodstock Road, and at 11 Ship Street. With your company, I went through a happy and joyful life at Jesus College.

I would like to thank my friends at Oxford: Xiaonan Lu, Yining Zhang, Hang Zhou, Tianyao Zhou, Jiate Luo, Zheng Zhang, Yifei Chen, Jingxiu Wu, Chaoying Wang, Zihui Su, Han Peng, Ding Pei, Yiwei Li, Zifan Wang, Jiayu Zhou, Lianlong Wu, Litong Jiang, Yinan Wang, Dingxi Zhou, Bing Wang, Feng Wang, Guanhua Ying, Aming Li, and Xiaofan Wang. I cherish the lovely cooking experience with Xiaonan at St Johns College and our travel time at Bristol and Cardiff. I would thank my friends for the long-lasting friendship: Zhenyu Zhang, Yu Huo, Lulu Song, Bo Li, Tianming Wu, Wendi Yang, Jiaming Zheng, Sitong Tao, Li Lin, Jiaqi Sun, Ruoxiao Chen, Eryue Du, Jian Cui, Bing Han, Huanfa Chen, Lin Chai, Donglin Li, Erchen Bo, Anqi Moxi, Runqing Yang, Xiaoshi Zhang, Simeng Lu, and Xinwei Li. Thank you for being a constant presence in my life. I truly cherish our friendship, and I keep this precious and unforgettable memories deeply in my heart.

I would like to express my sincere gratitude to Sholto Alexander Pridgeon for proofreading my DPhil thesis. Your feedback and suggestions have been incredibly insightful, and more importantly, have enabled me to express my ideas exactly as what I intend. Your revision undoubtedly has contributed significantly to the improvement of my thesis. I thoroughly enjoyed the interaction with you during the stage of thesis corrections. I would like to thank my collaborators and colleagues, Aditya Iyer, Qi

Huang, Adam Taylor, Man Cheng, Xiao-Ming Zhang, Tianhan Liu, Kecheng Liu, and Heng Shen for their careful proofreading. It is their efforts that make this thesis fewer errors.

Finally, I thank my parents for their unconditional and infinite love to me. I sincerely appreciate your love, care, and support throughout my entire life. Thank you for always having my back and being there for me. I wish that we accompany each other during our life journey forever. I wish good heath and happiness to my family.

Contents

Part I

Introduction

Chapter 1

Introduction

1.1 Background and motivation

Characterising quantum many-body behaviour such as exotic phases and collective excitations in quantum materials and molecular systems is a long-standing problem of theoretical and practical interest in quantum physics. Accurate simulation of quantum many-body systems provides us with valuable physical insights into their exotic behaviour. It helps us predict the microscopic and macroscopic properties of quantum many-body systems and unveil the underlying physics. In addition, it also has the potential to advance the rational design and synthesis of novel materials, which could in turn accelerate industrial development in many areas. The ability to accurately understand and further predict properties of quantum many-body systems, including materials and molecules, is consequently a central topic in quantum physics of both scientific and industrial interest.

A general quantum many-body system with nontrivial interactions is believed to be hard to compute classically. In this context, "computing" of a quantum many-body system is a general term which encompasses a broad range of problems, including the computation of the system's eigenstates or dynamic properties. There are substantial challenges involved in the description and computation of quantum systems. For instance, the Hilbert space of a quantum system grows exponentially with the system size. This exponential explosion is in general unavoidable, due to the quantum nature of the system. Unlike classical systems whose evolution can be described by a single trajectory in phase space, quantum many body systems can be difficult to characterise, and hence to solve, due to the exponential complexity of the Hilbert space size. Although approximations are available for specific problems, such as the mean-field approximation, which renders them classically solvable, these approximations have some limitations, and may not always be suitable for application to a general quantum many-body problem. Indeed, finding the ground state energy of a k-local ($k \geq 2$) Hamiltonian belongs to the computational complexity class nondeterministic polynomial (NP) time complete[1] [2–5]; solving a quantum many-body

[1] An n-qubit Hamiltonian $H = \sum_l H_l$ is k-local if H_l acts nontrivially on at most k qubits. The Ising spin glass, which is a special case of the 2-local Hamiltonian problem, is NP-hard [1].

problem remains a challenge in general.

A potential solution to this problem is the creation of an entirely new type of computer. This idea was clearly articulated in a lecture by Richard Feynman in 1982 [6], in which he set out a generic definition of a machine whose operations are, in principle, quantum. Feynman postulated that this system, which he termed a quantum computer, would be built out of quantum mechanical elements. There are various experimental platforms and instruments that can be regarded as quantum computers [7], including ultracold atoms in optical lattices, trapped ions, nuclear spins, neutron spectroscopy, superconducting circuits, photonic waveguides, and spins in semiconductors [8–11]. Depending on its ability to deal with universal problems as opposed to task-specific problems, this new class of computers is generally divided into two categories: digital quantum simulators and analogue quantum simulators. Digital quantum simulators can be regarded as a natural generalisation of digital classical computers [7, 12, 13]. Digital quantum simulators allow universal quantum operations on an ensemble of well-defined qubits, and are usually referred to as universal quantum computers as well. A universal quantum computer offers a universal simulation capability, and can be programmed to execute various tasks. The other category of quantum computer is the analogue quantum simulator, which directly emulates the target system by mapping it to specialised quantum hardware. They sacrifice the universality of solving general problems, but instead simulate specific systems in an analogue fashion. While its operations are restricted, analogue quantum simulation enables us to probe certain many-body phenomena which classical computers are incapable of examining as a result of limitations in computing power.

Nevertheless, machines capable of carrying out quantum operations were in existence, and use, prior to the conceptual development of quantum simulation. One example of this can arguably be found in spectroscopy. Spectroscopy initially began as the study of the interaction between matter and an external excitation, and was subsequently established as an experimental means to probe matter through external perturbations, such as electric or magnetic fields. In this thesis, the definition of spectroscopy also encompasses the characterisation of a generic quantum system. It is worth noting the differences and similarities between spectroscopy and quantum simulation. On one hand, quantum simulation relies on the usage of quantum devices to emulate the behaviour of another quantum system of interest (e.g. Hubbard models). Therefore, quantum simulation is a virtual rather than a real probe, and consequently differs from spectroscopy. On the other hand, spectroscopy and analogue quantum simulation share similarities in both their operation and goals. Intuitively, they both follow the rules of quantum mechanics, and detect features of quantum systems by measurement. For instance, in inelastic neutron scattering, the observable is the dynamical structure factor, with the measurement outcome detected by counting the scattered neutrons with different momenta [14, 15]. Analogue quantum simulators are built out of an ensemble of physical elements (qubits). These qubits evolve under certain driving

Hamiltonians, which are realised by applying external perturbations, and simulation results are obtained by measuring the qubits, similar to that in spectroscopy experiments. In addition, we can pursue an engineered approach to spectroscopy [16–21] which can be used to identify and characterise eigenstate properties of quantum systems through measurements on the time-evolved states (such as quench spectroscopy [16, 19]). This experimental technique of engineered spectroscopy shares similarities with quantum simulation, from the viewpoint of their goals.

The possible applications of quantum computing have continued to broaden in scope. With the development of quantum information science and computer science, these include linear algebra, machine learning, principal component analysis, optimisation, finance, drug synthesis, and many other areas [12, 22–25]. This thesis will focus on the applications of quantum computing in quantum many-body problems, including materials and molecular systems, although the methods developed in this thesis have potential applicability in other contexts.

Assuming the reality of quantum computers, a central question is whether we can use them to solve a quantum many-body problem of interest rigorously and efficiently. The answer to this seemingly straightforward question is not obvious. Finding the ground state of a quantum many-body system, to take one example, remains problematic: deciding whether the ground state energy of a 2-local Hamiltonian is greater than a_1 or smaller than a_2 for $a_1 > a_2$ is quantum Merlin-Arthur (QMA) complete [2], which is the quantum analogue of NP complete in a probabilistic setting[2]. This indicates that solving a quantum many-body problem is still a challenge, even for a quantum computer. Although this outcome for a general problem setting appears disappointing, quantum computing retains considerable promise in its application to special quantum many-body systems [12]. First of all, quantum computing could potentially address the above ground state problem under certain assumptions. In this context, the common assumptions include a nonvanishing energy gap and a nonvanishing overlap of the initial state and the target state. With these assumptions, we can use quantum algorithms, such as quantum phase estimation [27, 28], to obtain the ground state energy. There has been considerable progresses in the development of quantum computing methods, including quantum phase estimation [29, 30], quantum signal processing [31–34], and projection by linear combination of unitaries (LCU) [29, 35–39]. These algorithms are believed to hold significant promise to deliver solutions to ground state problems with a theoretical gurantee, given a sufficiently powerful universal quantum computer. I will discuss the solution to the eigenstate problem, including eigenstate preparation and eigenenergy estimation, in detail in Section 2.3 and Part II. The use of quantum computing is also promising in respect to addressing some quantum dynamics problems. For instance, simulating quantum dynamics for a local Hamiltonian belongs to the computational complexity class bounded-error quantum polynomial time (BQP), which can be efficiently simulated on a quantum computer in polynomial time [13, 27].

[2]The name QMA was given to it in Ref. [26]. Problems in this complexity class are not believed to be efficiently solvable with a quantum computer [4].

An interesting question is whether we are able to design quantum schemes that harness the power of quantum computing in solving quantum many-body problems. Typical quantum many-body problems of physical interest include strongly correlated electron systems, molecular systems, and quantum field theories. The central objective of this thesis is to facilitate the study of these classically challenging quantum many-body tasks by employing the power of quantum computing along with advanced classical computing approaches. This includes the use of universal quantum computing methods, hybrid quantum-classical computing methods, quantum simulation, and spectroscopy. It is an extremely rich research topic which draws on a wide variety of interdisciplinary approaches, ranging from analytic theory and numerical simulation to experiments with controlled quantum engineering.

In this thesis, I will develop several quantum computing and spectroscopic methods to estimate the properties of correlated quantum systems, such as their energy spectra and dynamics, which will subsequently enable us to access the exotic behaviour of many-body systems. Before moving on to further discussion of these topics, I would like to give some comments on aspects of quantum many-body physics and quantum computing. While quantum computing can be employed to facilitate our understanding of a quantum system, the success of quantum computing essentially depends on whether we could have a comprehensive, deep understanding of the quantum nature of the target system. From the perspective of physics, the development of quantum computing is not only a subject of computing, but also that of how we understand a quantum many-body system, especially its quantum nature. The development of quantum computing methods will in turn stimulate and advance the understanding of the underlying relation between different descriptions of the correlation effects in many-body systems at a fundamental level; it could potentially examine how quantum effects matter in relation to its interesting behaviour. From a practical point of view, it will enable a systematic study of quantum many-body behaviours as well as the prediction of properties in real materials. In this thesis, I establish a general, useful framework for describing and finding solutions to quantum many-body problems. I also apply the methods developed in this thesis to explore quantum many-body systems of interest, such as lattice models, magnetic materials and molecular systems.

1.2 Quantum computing of quantum many-body problems

The central objective of this section is to provide a description of quantum many-body problems and a general recipe for dealing with them. Quantum many-body problems can be divided into two general classes: static problems and dynamic problems. The first one corresponds to estimation of the eigenstate energy or properties (e.g., magnetisation) of a quantum system. The second one corresponds to property estimation for dynamical processes[3], such as response to

[3]Note that this thesis is primarily concerned with simulation of quantum dynamics, in which the quantum state is evolved under time evolution and the change of an observable is observed. In this context, this problem can also

an external field, quantum quenches, scattering, dissipation, and others. At this juncture it is worth noting that we can only obtain an estimation of the outcome from a quantum computer due to the uncertainty principle, as opposed to a definite outcome from a classical computer. Moreover, the only way that the information from a quantum computer can be accessed is through carrying out quantum measurements. This is the fundamental reason why problems are defined in probabilistic, rather than deterministic, terms, through the use of property estimation.

This section builds up a framework for quantum computing of quantum many-body problems, providing the foundation for the discussion in the thesis. I first formulate the two primary quantum problems considered in this thesis in Section 1.2.1. In Section 1.2.2, I provide a paradigm of quantum computing, which consists of four key components towards the solution of a quantum many-body problem. The introduction to these components will be expanded in Chapter 2. The methods developed herein will be further discussed in Part II in detail.

1.2.1 Problem description

Before moving on to a specific quantum many-body problem of interest in the real world, I formulate the problem in a general way. The problem is to find the state $|\Psi\rangle$ of a quantum system which is governed by the Hamiltonian H, described by its wave function Ψ, and to compute the expectation value of a physical quantity $\langle\Psi|\hat{O}|\Psi\rangle$, where $\hat{O}$ is a Hermitian operator representing an observable of interest.

The static problems and dynamic problems will be elaborated upon now. The static problems that are concerned in this thesis include eigenstate preparation and eigenenergy estimation. This already includes a large variety of problems in condensed matter physics. For instance, we can infer phase transitions by calculating a well-chosen order parameter[4].

Consider an N-qubit system with the Hamiltonian H. The eigenstate $|u_i\rangle$ and the corresponding eigenenergy E_i of the Hamiltonian satisfy,

$$H\,|u_i\rangle = E_i\,|u_i\rangle\,, \quad i = 0, 1, ..., 2^N - 1. \tag{1.1}$$

The task is to estimate the eigenenergy E_i, and the eigenstate property, characterised by an observable expectation on the target eigenstate $\langle u_i|\hat{O}|u_i\rangle$, up to certain precision.

The second type of task is dynamical property estimation of a quantum state under evolution, such as particle scattering and quench dynamics. The quantum state follows the laws of quantum mechanics, and its dynamics in a closed system could be described by the Schrödinger equation

$$\partial_t\,|\psi(t)\rangle = -iH\,|\psi(t)\rangle \tag{1.2}$$

be referred to as a dynamics problem, with an focus on dynamical property estimation.

[4] An order parameter is a measure which can be used to distinguish different phases, such as the magnetisation in a magnetic system. It is usually the expectation value of an observable on the ground state at zero temperature or an ensemble average at a finite temperature.

with the time derivative defined as $\partial_t := \partial/\partial t$. In this thesis, for simplicity, we set $\hbar = 1$ if there is no ambiguity. If we further assume the Hamiltonian is time-independent, the quantum state at time t in the Schrödinger picture is given by

$$|\psi(t)\rangle = e^{-iHt}\,|\psi(t=0)\rangle \tag{1.3}$$

where $|\psi(t=0)\rangle$ is referred to as the initial quantum state at $t = 0$.

When $U = e^{-iHt}$ is an identity operation, i.e., $H = 0$, and we choose the initial state as an eigenstate, the first task is reduced to a special case of Eq. (1.3). Therefore, Eq. (1.3) could be regarded as a unified description of the two tasks. Part II of this thesis discusses the solutions to these two tasks in detail.

This thesis mainly focuses on quantum dynamics in a closed system, in which quantum dynamics is described by unitary operations. These dynamical processes include quench dynamics, response to external fields, scattering dynamics, etc. If we focus on a subsystem interacting with another quantum system (bath), the subsystem dynamics is not unitary anymore. Instead of describing the joint system by a unitary operator, the subsystem dynamics may be described by the Lindblad master equation

$$\partial_t \rho(t) = -i[H(t), \rho(t)] + \lambda \mathcal{L}\big[\rho(t)\big]. \tag{1.4}$$

Here, the density matrix representation of this open system, ρ, is used, and $\mathcal{L}[\rho] = \frac{1}{2}\sum_k (2L_k \rho L_k^\dagger - L_k^\dagger L_k \rho - \rho L_k^\dagger L_k)$ is the superoperator that describes inherent coupling with the environment with dissipation strength λ and coupling operator L_k.

1.2.2 A general recipe

There are three primary components in computing the quantum state and the corresponding expectation value given by Eq. (1.3): the initial-state preparation, unitary evolution $U = e^{-iHt}$, and the final measurement. To obtain the solution given by Eq. (1.3), the classical computational resources, such as storage and running time, grow exponentially with the increase in the system size if the quantum state is treated in a classical state-vector form, and if there is no further assumption on the quantum state.

Alternatively, a quantum computer could simulate this dynamical process natively. Assume that the Hamiltonian is sparse and could be decomposed into a sum of local interaction terms as $H = \sum_{l=1}^{L} H_l$. The sparsity indicates that the number of terms in the Hamiltonian L grows polynomially with respect to the system size N as $L = \mathcal{O}(\mathrm{Poly}(N))$. This condition usually holds in realistic applications, such as the Ising model $L = \mathcal{O}(N)$, the Hubbard model $L = \mathcal{O}(N)$, and quantum materials and molecular systems $L = \mathcal{O}(N^4)$ in general (see Section 2.1). Here, we have the assumption of sparsity to make quantum simulation efficient, since the complexity of quantum simulation generally depends on the number of terms in the target Hamiltonian L (see

Section 2.4 for more details). On a universal digital quantum computer, introduced in Section 1.1, we could decompose the unitary evolution into elementary quantum operations up to an additive error ε and sequentially apply these elementary quantum operations to realise the joint evolution within an error threshold. Here, it should be noted that the initial state preparation could itself prove a challenge, if, for instance, we aim to prepare the ground state of the system.

To solve a specific problem, we should first encode the target problem into a problem that is compatible with a quantum computer. That is to say, the original problem, such as a fermionic or bosonic system, or a spin system, needs to be formulated into a qubit system with a proper number of qubits. This problem encoding could be challenging in itself. For instance, when it comes to real materials, the number of electrons in the material is of the order of Avogadro's number ($\mathcal{O}(10^{23})$). It would thus be impossible to process such a huge number of particles on a quantum computer. To enable quantum simulation on a quantum computer, the first step is to identify the most relevant degrees of freedom that characterise the interesting behaviour of the electrons. The description of a quantum system involves, for instance, a proper choice of bases and the active space where the degrees of freedom at this level contribute dominantly to the relevant physics. The construction of an effective Hamiltonian is discussed in Section 2.1.1. It is worth emphasising that the identification of the relevant degrees of freedom of the target system is not well defined. Usually, we can employ first-principle methods, such as density functional theory (DFT) [40], to precompute the functional of the material in order to have a rough sense of its properties, such as energy scale and energy dispersions. This first-principles method, albeit imprecise, provides an effective description of the system that captures its main features.

To concentrate the discussion on quantum simulation, let us first assume that the Hamiltonian has been constructed. In condensed matter physics, the behaviour in the free theory, i.e., where there is no interaction, is usually well known, and serves as a good reference. We attempt to study the behaviour in the interacting theory when the interaction is turned on. Successful methods include perturbation theory, adiabatic evolution, etc, which have been widely used in a variety of cases. These also inspires quantum algorithms for simulating the physics in the interacting theory. This idea was first proposed in the context of quantum field theories by Jordan, Lee, and Preskill [41].

The following is an example of quantum simulation of interacting systems at a high level:

1. Problem encoding: identification and construction of the effective Hamiltonian as well as the encoding of the quantum system in a proper basis.

2. Initial state preparation in the interacting theory.

3. Hamiltonian simulation: time evolution of the initial state under the Hamiltonian.

4. Estimating the physical quantity of interest by quantum measurement.

Below, I elaborate on the implementation of each step. At Step 1, firstly the original problem should be encoded into a qubit form that is compatible with quantum hardware, which comprises the representation of the Hamiltonian and the quantum state. The important stage is to identify the relevant degrees of freedom that captures the interesting physics, and have an effective description of the system Hamiltonian. Problem encoding depends on the choice of basis, such as the spin-orbital basis in molecular systems, which should be adapted according to the purpose of the simulation and the hardware to be used. In the context of bosonic lattice quantum field theories, there are four common choices [42]: field basis or harmonic oscillator basis in coordinate space or momentum space, respectively. For a fermionic problem, there are also many choices of bases, such as plane wave basis, orbital basis, Wannier basis, and band basis, either in first quantisation or second quantisation [12, 43, 44].

In Section 2.1.1, I demonstrate how to have an effective description of the Hamiltonian represented with a proper spin-orbital basis. Then, in Section 2.1.4, I show a general representation of quantum many-body systems. Based on that, the original problem can be encoded into a compact qubit form.

At Step 2, the initial state is prepared as the eigenstate of the interacting Hamiltonian. In general, preparing an eigenstate could be a challenge even for a quantum computer. Therefore, quantum algorithms need to be appropriately designed to achieve this goal. One solution is to first prepare the state in the free theory at a mean-field level. In a condensed matter or a quantum chemistry problem, we may first prepare a Hartree-Fock (mean-field) state as a reference state, which is usually easy to prepare. In the Gaussian orbital basis in second quantisation under proper fermionic-to-qubit mapping, for instance, the Hartree-Fock state is simply a tensor product state in the Pauli basis. For the preparation of a wave packet in the free theory in first quantisation, we may prepare the Gaussian vacuum state by the Kitaev-Webb algorithm on a quantum computer [45], as discussed in [41] or the author's paper [42]. A multi-particle wave packet states with given momenta could be prepared from the creation operator acting on the Gaussian vacuum.

The next step of eigenstate preparation is to evolve the initial state (a wave packet) in the free theory to a state in the interacting theory, by slowly turning on the interaction. The time scale in adiabatic evolution, for instance, is closely related to the energy gap of the quantum system [41]. There are various schemes for eigenstate preparation, which have been actively developed both in theory and through experiments over the past few decades. These include adiabatic evolution, phase estimation, quantum signal processing, linear combinational of unitaries, and variational state preparation. In this thesis, I will introduce several representative schemes for the state preparation in Section 2.3 and elaborate in Part II.

At Step 3, the initial state is evolved under the Hamiltonian drive. This is also referred to as Hamiltonian simulation. I will introduce several typical Hamiltonian simulation schemes in

Section 2.4 and elaborate on them in Part II.

Finally, at Step 4, a physical quantity of interest is accessed by measuring the final quantum state after evolution. Measurement should be implemented in polynomial time to preserve quantum advantages. Due to statistical fluctuations set by the laws of quantum mechanics, we require a proper number of measurements N_s to obtain the measurement accuracy to within an error threshold ε. To achieve estimation precision ε, the number of measurements scales as $N_s = \mathcal{O}(1/\varepsilon^2)$ [46], known as the standard quantum limit. A typical high-precision measurement strategy is quantum phase estimation [4,46,47]. The estimation of a physical quantity can be obtained by performing measurements on the ancillary qubits. The number of the ancillary qubits is dependent on the desired simulation accuracy. In some instance, the scaling of measurement complexity may be improved to $N_s = \mathcal{O}(1/\varepsilon)$ [48]. To avoid the cost of the increase in the circuit depth and qubit number, we can directly measure the observable on the final quantum state. For instance, we can decompose an arbitrary observable into Pauli bases, and consequently it can be directly measured on a quantum computer without additional quantum circuits, which is discussed in Chapter 10.

1.3 Quantum computing, quantum simulation, and spectroscopy

In Section 1.1, I briefly introduced the concept of quantum computing in a broad sense. Quantum computing draws upon the principles of quantum simulation, allowing quantum operations on carefully designed systems following the laws of quantum mechanics. On the basis of this description, spectroscopy shares similarities with quantum simulation. From a theoretical point of view, spectroscopy refers to observation of the energy spectrum of a quantum system, and reveals its spectral information in either the time or the frequency domain. From a first principle standpoint, this bears a resemblance to the two types of problems outlined in Section 1.2, static problems and dynamic problems. I discuss the close relation between the static problem case and the dynamic problem case from the viewpoint of spectroscopy in Section 1.3.3. From an experimental point of view, spectroscopy is a well-established approach to probe and study the properties of quantum matter. To measure the properties of a quantum system, such as electronic and magnetic structures and excitations, a generic spectroscopy approach is to perturb the system by an external field, and observe the response to the perturbation. This is also similar to quantum simulation in both the methodology and the actual allowable operations. Moreover, spectroscopic methods provide valuable insights in studying both dynamic and static properties, which are introduced in Section 1.3.3 and Section 7.3 and elaborated upon in Chapter 7 and Appendix E.

This section extends the discussion in Section 1.1. I review quantum computing, quantum simulation, and spectroscopy, and detail some aspects of their close relationships in the context of quantum many-body problems introduced in Section 1.2.1.

1.3.1 Spectroscopy

In Section 1.1, I introduced concepts of spectroscopy and quantum simulation and the relation between them. The objective of quantum simulation is to emulate the behaviour of another quantum system with a controllable quantum device, and therefore it is a virtual, rather than real, probe. This is in contrast to spectroscopy, which usually refers to an experimental technique to probe matter through external perturbations, such as electric or magnetic fields. Although spectroscopy and quantum simulation are distinctive in their approach, they nevertheless have a similar objective, namely characterising complex systems of interest. They are also currently the most powerful means we have to achieve this objective, and it could therefore be interesting to discuss the relation between the two.

Spectroscopy initially began as the theoretical and experimental study of the interaction between matter and an external excitation, and it is used in this thesis to encompass the characterisation of a generic quantum system, such as eigenstate properties. Experimental spectroscopy techniques include angle-resolved photoemission spectroscopy (ARPES), and neutron or X-ray scattering spectroscopy. For example, in neutron scattering, we explore the properties of materials by injecting neutrons into the materials and analysing the distribution of scattered neutrons [15]. An incident neutron has an interaction with the nuclei or electrons (both spin and orbital degrees of freedom) in the materials and is scattered following the laws of quantum mechanics. The neutron here acts as both the external perturbation and the probe. Measurement data is collected by counting the scattered neutrons with different energy and momenta, with the results reflecting the collective behaviour of the target material. In quantum simulation experiments, quantum operations are applied on an experimental platform, which can be superconducting circuits [8, 49], trapped ions [9, 50–52], Rydberg atoms [53, 54], NMR [55–57], optical lattice [58], etc. In analogue quantum simulation, the initial state is simply evolved under the system Hamiltonian drive for a certain period of time, after which an observable is measured and the statistical results are obtained. Observables include, for example, the magnetisation and the particle number. We observe that spectroscopy and quantum simulation share certain similarities in their operation: the scattering experiments described above are carried out using an instrument which follows the rules of quantum mechanics and detects the quantum nature of observables through measurement, as is also the case with a quantum simulator. However, spectroscopy has limited degrees of freedom relative to that of analogue quantum simulators. Conventional spectroscopy experiments are only able to perform restricted operations and the system cannot be engineered. By contrast, an analogue quantum simulator usually permits an engineering of quantum systems, such as single-qubit addressing (though not universal quantum gates, as opposed to a digital quantum computer).

A recent notable development is the creation of spectroscopic techniques for engineered systems [16–20, 59], which have more controllable degrees of freedom. For example, Ref. [16, 19, 20]

introduces quench spectroscopy, which can be used to look for spectral response with periodic drive or quench dynamics. Spectroscopy also provides insights into quantum computing schemes designed to obtain eigenstate information. The procedure can be summarised as follows. The initial state is prepared, and then evolved under the simulator's Hamiltonian. Following that, the expectation values of a well-chosen observable $\hat{O}$ at different times t_i are obtained by measurements. By post-processing the expectation values $\{\langle\hat{O}(t_i)\rangle\}$ under time evolution, we can obtain the eigenstate properties of the Hamiltonian (see [18–20] and Chapter 7).

1.3.2 Spectroscopy and Hamiltonian learning

As seen from the above discussion, the spectroscopic features of a system can be accessed by carrying out spectroscopy experiments. Spectroscopy also serves as a useful tool for quantum Hamiltonian learning problems. This section discusses the relation between spectroscopy and Hamiltonian learning.

Recall the main procedure of the spectroscopy experiment. To study the properties of the quantum material, we firstly precompute core properties of the material. Here, precomputation refers to employing the available, relatively low-cost methods to calculate material properties before we perform real experiments or high performance computing, which usually require a significant allocation of resources. For instance, we can employ first-principle methods, such as density functional theory (DFT), to precompute the functional of the material and have an effective description of the system that captures its main features. While the Hamiltonian may not be very accurate, with this prior knowledge, we can gauge the ground state and the spectroscopic features that will be further revealed by spectroscopy experiments.

In spectroscopy experiments, the incident neutron interacts with the phonons or electrons in the material, and the information about the material is then recorded by the scattered neutrons. By comparing the experimental results obtained by measurements and the predicted results using the effective Hamiltonian generated by DFT, we can modify the effective Hamiltonian, determine its parameters, and hence have a better description of the system. This could be regarded as a Hamiltonian learning process.

The Hamiltonian generated by DFT can be expressed in a general form as

$$\hat{H} = \sum_i h_i \hat{P}_i, \tag{1.5}$$

where $\hat{P}_i$ is a general operator, which, for instance, could be a spin operator, and h_i characterises its strength. Formally, the measurement outcome can be expressed as $S_{\hat{O}} = \mathrm{Tr}(\hat{O}\rho)$ where the observable, for instance, could be the dynamical structure factor (see Section 1.3.3 and Appendix E). With the measurement outcome, the idea of learning the effective Hamiltonian

is to minimise the distance between the experimental outcome and the expectation value of observables obtained by theoretical caculation over different parameters $\{h_i\}$,

$$\min_{\{h_i\}} \mathrm{Tr}(\hat{O}\rho) - \mathrm{Tr}(\hat{O}\rho_{\mathrm{expm}}). \tag{1.6}$$

Here, the state is in thermal equilibrium, $\rho = e^{-\beta\hat{H}}/\mathrm{Tr}(e^{-\beta\hat{H}})$. Note that challenges remain in connecting the data measured on the experimental apparatus with Hamiltonian learning. Ideally, one should be able to directly interface the experimental apparatus with other quantum systems. One proposal for Hamiltonian learning is using the interactive quantum likelihood evaluation method, which requires interactive quantum operations on two coupled quantum systems, see [60–62] for reference. This is clearly not the case for spectroscopy.

It is worth noting that given a universal quantum computer, a straightforward choice of learning a Hamiltonian is via quantum process tomography [27], which gives the complete knowledge of the Hamiltonian. However, since the dimension of a many-body Hamiltonian grows exponentially with the system size, carrying out a complete tomography of the Hamiltonian is costly and formidable in real experiments [63]. Recently, a more efficient Hamiltonian learning protocol was proposed by the author [64], which requires only single-qubit operations for learning a Hamiltonian, which can already be realised by most engineered quantum simulators.

Compared to digital quantum simulation, a spectroscopy experiment only permits very restricted operations and measurements, in which only collective behaviour can be observed. In addition, only classical data can be accessed, while further additional operations are not permitted. Despite the inherent restrictions, spectroscopic approaches, as an alternative, have some degrees of freedom, which includes temperature, pressure, external fields, etc. Learning effective physics from spectroscopic experiments would be an interesting future direction. I direct interested readers to [65–67] for the Hamiltonian learning process in this specific context.

1.3.3 Dynamics, response, and spectral function

In this section I employ a selection of basic quantum mechanics formulae to provide a concrete demonstration of the relation between dynamics, response and spectral function. This is to provide the reader with a reference to existing theory, which serves as a basis for the development of engineered spectroscopy and quantum computing methods outlined in this thesis.

To illustrate the relation between dynamic properties and static properties, I will first give an illustrative example using the most basic quantum mechanics and spectroscopic analysis. Suppose that the initial state is decomposed into the eigenbases $\{|u_i\rangle\}$ of the Hamiltonian as $|\psi_0\rangle = \sum_i c_i |u_i\rangle$. The state under real-time evolution in the Schrödinger picture is given by $|\psi(t)\rangle = e^{-iHt}|\psi_0\rangle = \sum_i c_i e^{-iE_i t}|u_i\rangle$, and we have the observable expectation on the time-evolved state as

$$\langle\psi(t)|\hat{O}|\psi(t)\rangle = \sum_{ij} c_i^* c_j e^{i(E_i - E_j)t} \langle u_i|\hat{O}|u_j\rangle . \tag{1.7}$$

This simple toy example indicates that dynamic properties of the state could be accessed given the eigenenergies of the Hamiltonian.

In a spectroscopy experiment, the system is in its thermal equilibrium $[\rho, \hat{H}] = 0$; the state is diagonal in the eigenbasis of $\hat{H}$, which can be written as $\rho = \sum_n \rho^{nn} |u_n\rangle \langle u_n|$. In thermal equilibrium, ρ^{nn} takes the explicit form as $\rho^{nn} = e^{-\beta E_n}/Z$ where $\beta = T^{-1}$ is the inverse of temperature, Z is the partition function $Z = \mathrm{Tr}(e^{-\beta \hat{H}})$, and the Boltzmann constant is set as $k_B = 1$ in natural units. The observable in neutron spectroscopy is the dynamical structure factor, which is related to the unequal time correlator and reflects the dynamic property of the system. Consider a generic unequal time correlator, which is defined as

$$C(t, t') = \langle \hat{O}_1(t) \hat{O}_2(t') \rangle, \tag{1.8}$$

in the Heisenberg picture with $\hat{O}(t) = e^{i\hat{H}t} \hat{O} e^{-i\hat{H}t}$ and $\hat{O} := \hat{O}(t = 0)$. Here, the $\langle \cdots \rangle$ is taken as the canonical ensemble average. To simplify the notation, the eigenstate is denoted as $|n\rangle := |u_n\rangle$. Expanded on the eigenbasis of $\hat{H}$, we have

$$\begin{aligned} C(t, t') &= \mathrm{Tr}[\hat{O}_1(t) \hat{O}_2(t') \hat{\rho}] \\ &= \sum_{n, n', m} \rho^{n'n} e^{i(E_n t - E_{n'} t')} e^{-iE_m (t - t')} \times \langle n | \hat{O}_1 | m \rangle \langle m | \hat{O}_2 | n' \rangle \delta_{n, n'} \\ &= \sum_{n, m} \rho^{nn} e^{i(E_n - E_m)(t - t')} \times \langle n | \hat{O}_1 | m \rangle \langle m | \hat{O}_2 | n \rangle. \end{aligned} \tag{1.9}$$

Taking the Fourier transform of an unequal time correlator in the time domain, we have

$$\begin{aligned} S(\omega) :=& \frac{1}{2\pi} \int dt e^{i\omega t} C(t, t') \\ =& \sum_{n, m} \rho^{nn} \langle n | \hat{O}_1 | m \rangle \langle m | \hat{O}_2 | n \rangle \times \delta(E_n - E_m + \omega). \end{aligned} \tag{1.10}$$

$S(\omega)$ is also commonly referred to as the dynamical structure factor (or dynamic structure factor) in the literature. For later convenience, the purely dynamic part of $S(\omega)$ is defined as $\tilde{S}(\omega) = S(\omega) - \langle \hat{O}_1 \hat{O}_2 \rangle \, \delta(\omega)$. If the spectral weight $\langle n | \hat{O}_1 | m \rangle \langle m | \hat{O}_2 | n \rangle$ coupled by the two eigenstates $|m\rangle$ and $|n\rangle$ is nonvannishing, the energy difference of these two coupled eigenstates $|m\rangle$ and $|n\rangle$ could be revealed by finding the local maximum of the spectral function, which peaks sharply at resonance, $\omega = E_m - E_n$. This discussion holds in the linear response regime where the perturbation is weak. Otherwise, the eigenstates of the new perturbed Hamiltonian will be different from the unperturbed eigenstates, as well as the eigenenergies.

The above discussion starts from a time-dependent correlation function as a generic quantity reflecting the dynamics of the quantum system. Below, I will briefly review how this quantity is related to the response of the system. It is natural to denote the original unperturbed Hamiltonian as $\hat{H}_0$ and the perturbation as $\hat{V}$, and the Hamiltonian reads

$$\hat{H} = \hat{H}_0 + \hat{V}. \tag{1.11}$$

In the Heisenberg picture, the operators are time-dependent, and to make the discussion more general, the perturbation is written as

$$\hat{V}(t) = \sum_i \hat{O}_i(t) f_i(t), \tag{1.12}$$

where the observable may be local as $\hat{O}(\mathbf{x}, t)$. The system is initialised as the ground state of the original Hamiltonian. A perturbation $\hat{V}$ is then applied at time $t = 0$, and an observable $\hat{O}$ at time $t > 0$ is measured. The perturbation drives the system out of equilibrium, and the system shows a dynamical response against the applied perturbation. In the interaction picture, the time-evolved state is

$$\rho_I(t) = U_I(t)\rho_{0,I}(t)U_I^\dagger(t) \tag{1.13}$$

with $U(t) = U(t; t_0 \to -\infty)$ and

$$U_I(t; t_0) := \mathcal{T} \exp(-i \int_{t_0}^{t} \hat{V}(t_1) dt_1). \tag{1.14}$$

Under the perturbation, the time-dependent expectation of an observable can be written as

$$\begin{aligned} \langle \hat{O}_i(t) \rangle &= \mathrm{Tr}\, \rho_I(t) \hat{O}_i(t) \\ &= \mathrm{Tr}\, \rho_{0,I}(t) U_I^\dagger(t) \hat{O}_i(t) U_I(t) \\ &= \langle \hat{O}_i(t) \rangle|_{f=0} - i \int_{-\infty}^{t} dt_1 \langle [\hat{O}_i(t), \hat{V}(t_1)] \rangle + \mathcal{O}(\|\hat{V}\|^2), \end{aligned} \tag{1.15}$$

which is restricted to the first-order expansion, and again the expectation is taken as the ensemble average on the thermal state.

It is common to define a response function

$$\chi_{ij}(t, t_1) = -i\theta(t - t_1) \langle [\hat{O}_i(t), \hat{O}_j(t_1)] \rangle, \tag{1.16}$$

and then we have

$$\begin{aligned} \langle \hat{O}_i(t) \rangle - \langle \hat{O}_i(t) \rangle|_{f=0} &= -i \int_{-\infty}^{t} dt_1 \sum_j \langle [\hat{O}_i(t), \hat{O}_j(t_1)] \rangle f_j(t_1) \\ &= \int_{-\infty}^{\infty} dt_1 \sum_j \chi_{ij}(t, t_1) f_j(t_1). \end{aligned} \tag{1.17}$$

This is the celebrated Kubo formula, which connects the dissipative quantity on the left-hand side with the equilibrium average of the correlation function. Taking the Fourier transform of the response function,

$$\chi(\omega) = \lim_{\eta \to 0^+} \int_0^\infty dt e^{i\omega t + \eta t} \chi(t, 0),$$

we further have the relation

$$\mathrm{Im}\, \chi(\omega) = -\pi(1 - e^{-\beta\omega}) \tilde{S}(\omega), \tag{1.18}$$

at thermal equilibrium, which is known as the fluctuation-dissipation theorem [15].

The response function can also be calculated from the eigenvalues and eigenstates of H. Denote the thermal probability $p_n = \rho^{nn} = e^{-\beta E_n}/Z$; it can be shown that

$$\begin{aligned}\chi_{\hat{O}_1\hat{O}_2}(\omega) =& -i\sum_{n,m}(p_n - p_m)\langle n|\hat{O}_2|m\rangle\langle m|\hat{O}_1|n\rangle \int_0^\infty dt \exp(i(\omega - E_m + E_n + i\eta)t) \\ =& \sum_{n,m}(p_n - p_m)\frac{\langle n|\hat{O}_2|m\rangle\langle m|\hat{O}_1|n\rangle}{\omega - E_m + E_n + i\eta},\end{aligned} \tag{1.19}$$

which is known as the Lehmann representation. In the zero-temperature limit, the response function in the frequency domain can be expressed as

$$\chi_{\hat{O}_1\hat{O}_2}(\omega) = \sum_n \frac{\langle u_0|\hat{O}_1|u_n\rangle\langle u_n|\hat{O}_2|u_0\rangle}{\omega - (E_n - E_0) + i\eta} + \sum_n \frac{\langle u_0|\hat{O}_2|u_n\rangle\langle u_n|\hat{O}_1|u_0\rangle}{\omega + (E_n - E_0) + i\eta}. \tag{1.20}$$

The imaginary part of the response function is given by

$$\operatorname{Im}\chi(\omega) = -\pi\sum_{n,m}(p_n - p_m)\langle n|\hat{O}_2|m\rangle\langle m|\hat{O}_1|n\rangle\delta(E_m - E_n - \omega), \tag{1.21}$$

and one can recognise its relation with the dynamical structure factor. This equivalence of the representation reveals that the information could be accessed either in the time domain or in the frequency (energy) domain.

In the above discussion, we can find that the spectral function can be obtained by taking the Fourier transform of the time-dependent correlation function $C(t, t')$. This motivates how eigenenergy differences can be determined using quantum computing. The time-dependent correlation function $C(t, t')$ can be simulated and measured on a quantum computer: simulate real-time evolution using the Hamiltonian simulation methods introduced in this thesis, measure $C(t, t')$ with a Hadamard-test quantum circuit, and finally obtain the spectral properties by appropriate post-processing. Related discussion on engineered spectroscopy methods can be found in Chapter 7.

From Eq. (1.7) we see that the spectral information is contained in the dynamics of the state. The inherent relation of static properties and dynamic properties also motivates both classical and quantum computing methods. As introduced in Section 1.2 and will be elaborated Section 2.4, quantum computers are able to simulate real-time evolution $U = e^{-i\hat{H}t}$ of a physical sparse Hamiltonian efficiently. An interesting question is whether we can infer static properties from dynamic properties. In Section 7.3, two approaches that accomplish this task are demonstrated: the spectroscopy analysis methods and the quantum subspace expansion method.

In addition, the retarded Green's function $G^r(t, t')$ is a special case of the response function with $\hat{O} = \hat{\psi}$ and $\hat{O}' = \hat{\psi}^\dagger$, and $\hat{\psi}$ and $\hat{\psi}^\dagger$ is the field operator. In equilibrium, the Greens functions can be related to the spectral function $A(\omega) = -\frac{1}{\pi}\operatorname{Im} G^r(\omega)$, which can be directly measured

by using spectroscopy techniques, such as neutron scattering, angle-resolved photoemission spectroscopy (ARPES). We can also employ spectroscopic methods as an effective means to infer static properties of materials, discussed in detail in Chapter 7. Simulation results are shown in Chapter 8.

1.4 Quantum computing and quantum advantage in the near future

At the beginning of the thesis in Section 1.1, quantum computing, quantum simulation, and quantum many-body problems were introduced, followed by a general recipe for quantum computing of quantum many-body physics problems in Section 1.2. It seems that for a wide range class of quantum problems, we are in theory able to immediately use the aforementioned strategy to deal with the problems at hand, that is, to design quantum computing methods that are compatible with quantum devices, implement these on the devices, and ultimately obtain a solution to problems of interest. However, in reality, the situation is more complex.

From the perspective of quantum computing, it is not always easy to prepare the initial state, evolve it, and measure it with polynomial resources. The amount of physical resources (such as number of qubits, number of operations, number of steps, etc.) needed for the quantum simulation is heavily contingent upon the type of the target problem and the particularities of the simulator.

In terms of implementation, quantum simulation should be executed on a quantum device. In the above discussion, it is assumed that the quantum simulation devices are universal and error-free, in the sense that they allow perfect, universal quantum operations. The limitation of the quantum devices have not been addressed, and the computational or simulation resources needed for execution of the problem have only briefly been touched on. However, current hardware is limited in terms of quality (e.g., qubit quality, gate fidelity, and measurement fidelity) and size (qubit number). In order to run a quantum simulation, we have to confront, and overcome, the challenges imposed by quantum hardware technology. The primary challenges and the corresponding restrictions are summarised as follows:

1. Controllable qubit number $\leftrightarrow$ computational size
2. Circuit depth $\leftrightarrow$ allowed problem complexity
3. Noise rate $\leftrightarrow$ quality and reliability of the solution

With the development of quantum technology, it is expected that in the long term, we may overcome all these challenges, allowing for a universal, fault-tolerant quantum computer. With this fault-tolerant quantum computer, a problem could in theory be programmed on a quantum computer and solved, if it fell into the BQP class. The computation cost mainly arises from the

complexity of the problem and the algorithm, which determine run-time; an efficient and fast implementation of quantum algorithms is unquestionably desirable.

The implementation of fault-tolerant quantum computing requires 10^3 or more physical qubits per logical qubit to suppress the physical error [68]. This is far beyond current technological capabilities. At present in 2022, state-of-the-art programmable quantum hardware is able to control tens to hundreds of noisy qubits, with the error rate of two-qubit operation above 10^{-3} [69, 70]. It is more pragmatic in the near future to focus on the noisy intermediate-scale quantum (NISQ) regime, a notion originally proposed by John Preskill [71]. While it is still challenging to perform error-corrected universal quantum operation, we can utilise the present generation of quantum computers to perform specific tasks that are already challenging for classical computers. Currently, we are able to run some quantum algorithms on a shallow noisy circuit without implementing full error correction [12, 25, 72–74], which largely reduces the quantum hardware requirements in terms of quality. This includes not only algorithms running on gate-based digital quantum devices, but also analogue quantum simulators, as discussed in previous sections.

The outcome of the landmark quantum supremacy/advantage experiments were initially announced by Google's Team in 2019 using superconducting qubits for random circuit sampling with the qubit number 53 [11]. Subsequently, in 2021 a group led by Pan at USTC demonstrated random circuit sampling with up to 56 qubits on quantum processor, Zuchongzhi, which is composed of 66 functional qubits in a tunable coupling architecture [75]. In 2022, a team led by Lukin at Harvard University used arrays of neutral atoms trapped in optical tweezers with 289 qubits to investigate quantum algorithms for solving the maximum independent set problem [76], although this arguably cannot be seen as a fully functional, programmable quantum computer.

It is worth noting that the question of whether these experiments have demonstrated a concrete quantum advantage remains controversial. This is because as classical computing systems and methods have evolved, these problems may not be as unfeasible or impractical for them as previously assumed. For instance, Google claimed in [11] that classically sampling a 53-qubit quantum circuit with a circuit depth of 20 a million times would take approximately 10,000 years, while the Sycamore processor took about 200 seconds to execute the same sampling task. Subsequently, however, several works proposed efficient classical simulation methods based on tensor networks, which are able to simulate Google's quantum supremacy circuits in days ([77]) and even seconds ([78]), undermining the assertion of quantum supremacy.

Additionally, the problems tackled in the experiments and solutions advanced, such as random circuit sampling [11, 75] and Gaussian boson sampling [79], may lack relevance in terms of real-world application. An important question is whether the advantages of quantum computing over classical computing can be clearly demonstrated through solving classically challenging problems of physical interest on quantum devices. This thesis considers applications and implementations

of quantum computing with near-term noisy quantum computers in the noisy intermediate-scale quantum (NISQ) era, and universal fault-tolerant quantum computers in the fault-tolerant quantum computing (FTQC) era.

In the near future, we may need to consider quantum computation with shallow noisy quantum circuits, which are only capable of performing a constant number of restricted operations without full error correction. Hybrid quantum-classical computing, such as variational quantum algorithms, serves as a promising candidate for being compatible with near-term noisy quantum computers [25, 72, 80–83]. At this point, it is worth quoting a definition of hybrid quantum-classical computing from a recent perspective paper. Hybrid quantum-classical computing is a computing structure that "requires nontrivial amounts of both quantum and classical computational resources to run, and which cannot be sensibly described, even abstractly, without reference to the classical computation" [84]. Intuitively, because a large portion of the computational burden is processed on a classical computer, deep coherent quantum circuits may not be required.

One of the most influential and representative developments in the field of NISQ-suitable quantum algorithms is variational quantum algorithms (VQAs). These algorithms make use of short-depth parameterised quantum circuits, which are particularly suited to NISQ hardware, embedded in an otherwise classical variational loop [72, 85]. They directly realise the target quantum state with a controllable quantum device and reduce the requirement of quantum resource requirements by leveraging the power of classical optimisation. VQAs have demonstrated the potential computing power through the co-development of quantum and classical resources together; hybrid quantum-classical algorithms have been an integral part of quantum algorithms research ever since.

The success of hybrid quantum-classical computation is contingent on the following (1) sufficiently powerful quantum circuits to express the optimal solution of target problems [86], (2) efficient sampling and classical optimisation [87, 88], and (3) robustness to noise [89]. An important question is whether a clear and robust quantum advantage persists for such hybrid quantum-classical algorithms given these challenges. Notwithstanding these challenges, noisy intermediate-scale quantum processors still hold promise in solving certain problems that are already posing a challenge to the computational power of classical computers. A natural question flowing from this is whether quantum computation could address open problems appearing in the study of quantum many-body problems, where many of them have strong correlations. Part IV of this thesis addresses this important point.

Another similar question relate to quantum simulation is how far this could be feasibly developed in the near-future, both for NISQ and early fault-tolerant quantum computers. The discussion on this topic is concluded in Chapter 9, which provides a quantum resource estimation for eigenenergy estimation and eigenstate property estimation in several typical problems.

This addresses the question as to what circumstances and conditions are required to solve a physical problem and hence to achieve a quantum advantage, with the aim of shedding further light on the identification of the boundary of classical and quantum limits.

1.5 Thesis outline

The four subsections in Chapter 1 have introduced a general recipe for solving a quantum many-body problem, discussing in broad terms the close relation between quantum computing, quantum simulation and spectroscopy, as well the realistic application of different approaches in the near future. Chapter 2 covers the essential concepts introduced in Chapter 1 and further developed in Part II, Part III, and Part IV. Chapter 2 serves as a preliminary for the following chapters by providing a self-contained introduction to the techniques and methodologies introduced in this thesis.

Part II concerns quantum computing of quantum many-body problems, focusing on the issue of how to estimate static and dynamic properties of quantum systems, as a response to the points raised in Section 1.2. In Chapter 3, a general quantum computing framework is established for the static and dynamic problems herein. I introduce a new quantum-classical architecture, the hybrid tensor network, which serves as a basis for a general representation of quantum system representation which is applicable to a broad range of problems. In Chapter 4, I focus on static property estimation, which consists of two tasks (see Section 1.2.1), eigenenergy estimation and eigenstate property estimation. Motivated by imaginary-time evolution, which provides a natural way of driving the initial state to the ground state, I introduce quantum algorithmic cooling which realises a generalised imaginary-time evolution with the best asymptotic scaling to date in eigenstate energy and property estimation. I explicitly demonstrate how to search for the eigenenergy, and estimate the eigenstate property. The algorithm is proven to reach the Heisenberg limit in eigenenergy estimation and achieve an exponential speed-up for the circuit complexity compared to phase estimation.

In Chapter 5 and Chapter 6, I focus on dynamic problems formalised in Section 1.2.1 and detailed in Section 2.4. In Chapter 5, I propose adaptive product formulae to estimate dynamic properties of quantum states under real-time evolution. I further show how to enable large-scale quantum dynamics simulation by using the methods developed in Chapter 3. In Chapter 6, I consider a generic interacting quantum system described by $H = H^{\mathrm{loc}} + V^{\mathrm{int}}$. I propose a perturbative approach for such a problem, which shows advantages over conventional perturbative expansion and quantum simulation, and has wide applications in probing interacting dynamics.

In Part III, I explore interesting emergent quantum phenomena with the methods developed in this thesis. I discuss the close relation between quantum simulation and quantum computing, from both the theoretical, experimental, and operational point of view, as a response to the discussion in Section 1.3. In Chapter 7, I show how spectroscopy motivates quantum computing

methods to predict spectroscopic features of generic quantum many-body systems. I introduce a theoretical framework of engineered spectroscopy and demonstrate how the spectroscopic features of quantum systems with general and tuneable interactions can be probed. In Chapter 8, several quantum many-body phenomena in typical physical systems are explored, including interacting bosons, fermions, and quantum spins, in which the interacting dynamics is simulated using the perturbative approach.

Part IV discusses the challenges of employing quantum computing towards realistic applications in the near future, as a response to the discussion in Section 1.4. With the development of theoretical and experimental methods, the most interesting question is how far away we are from solving a practical problem. To answer this question, the quantum resources required for eigenstate property estimation at quantum gate level are estimated using a new methodological approach developed by the author in Chapter 9. This provides a clear illustration of the resource requirements for physical systems of interest, such as the Ising model, the Hubbard model, and molecular systems, with a focus on their application in both the NISQ and FTQC era. In Chapter 10, I discuss the challenges of noisy quantum computers, as errors are unavoidable when using such hardware. These include device errors (such as process errors[5] and readout errors) and statistical errors. Accordingly, I introduce a stochastic error mitigation scheme to mitigate process errors in NISQ devices, including both analogue quantum simulators and digital quantum simulators, and an efficient quantum state measurement scheme, called overlapped grouping measurement, to mitigate statistical errors and enable fast quantum processing.

Chapter 11 summarises the works carried out in this thesis, and concludes with some perspectives on quantum computing, quantum many-body physics, and quantum materials, as well as the current outlook for this rapidly growing field.

The contents presented from Chapter 2 to Chapter 10 in this thesis are based on the original research work of the author and collaborators, unless otherwise specified. The author conceived the key idea of the works that are presented in this thesis, and was responsible for the derivation of the analytic results, numerical and experiment verification, and manuscript writing. The author is the first author and/or the corresponding author of these publications. Chapter 3 is related to a published work [90] and a manuscript under preparation [91]. Chapter 4 is related to a manuscript currently under preparation [92] and a preprint [93]. Chapter 5 is related to a manuscript under preparation [91] and the theoretical analysis in a published work [94]. Chapter 6 is related to a published work [95]. Chapter 7 is related to a manuscript under preparation [96]. Chapter 8 presents the results for the applications of quantum computing and spectroscopy methods developed in preceding chapters, and is related to [95, 96]. Chapter 9 is related to Chapter 4 and [92]. Chapter 10 consists of two parts, which are related to [97] and the theoretical part in [98], respectively.

[5]Quantum processes include quantum gates in a digital quantum simulator and time evolution in an analogue quantum simulator.

Chapter 2

Framework and methods

In this chapter, I introduce the framework and methods for quantum computing of quantum many-body systems dealt with in this thesis. Recall that a general recipe of quantum computing of quantum many-body problems was introduced in Section 1.2. This is comprised of four parts: problem encoding, state preparation, dynamics simulation, and quantum measurement. I will introduce the essential ingredients of these four parts in Section 2.1, Section 2.3, and Section 2.4.

The role of this chapter is to give an overview of the motivation, framework, and methods that are introduced and developed in this thesis. This chapter serves as a preliminary for the following chapters. The contents are self-contained in order for the reader to find all the essential information from this chapter. It is especially intended for interested readers who wish to learn more about quantum computing, quantum many-body physics and quantum information, and to capture the key ideas of the techniques and methods developed in this thesis.

2.1 Problem encoding

2.1.1 Effective Hamiltonian

A time-independent non-relativistic Hamiltonian could be expressed as

$$H = -\sum_{i}^{N_e} \frac{\hbar^2}{2m_e}\nabla_i^2 - \sum_{I}^{N_n} \frac{\hbar^2}{2M_I}\nabla_I^2 - \sum_{i}^{N_e}\sum_{I}^{N_n} \frac{Z_I e^2}{4\pi\varepsilon_0^2|\mathbf{r}_i - \mathbf{R}_I|} + \frac{1}{2}\sum_{i\neq j}^{N_e} \frac{e^2}{4\pi\varepsilon_0^2|\mathbf{r}_i - \mathbf{r}_j|} + \frac{1}{2}\sum_{I\neq J}^{N_j} \frac{Z_I Z_J e^2}{4\pi\varepsilon_0^2|\mathbf{R}_I - \mathbf{R}_J|}. \tag{2.1}$$

Here, M_I, $\mathbf{R}_I$, and Z_I denote the mass, position, and atomic number of the Ith nucleus, and $\mathbf{r}_i$ is the position of the ith electron. N_e and N_n are the number of electrons and nuclei. In this thesis, natural units are used, $\hbar = m_e = e^2/4\pi\varepsilon_0 = 1$. The first two sums in H are the kinetic terms of the electrons and nuclei, respectively. The last three sums represent the Coulomb repulsion between the electrons and nuclei, the electrons themselves, and the nuclei themselves, respectively.

Since a realistic physical system involves a prohibitively large number of particles, it is impossible to solve the Hamiltonian given by a full description of the material in Eq. (2.1). However, a

solution to the "theory of everything" is unnecessary. As discussed in previous sections, interesting phenomena could be captured with much fewer degrees of freedom at a finite scale. In most cases, we could restrict the energy scale of the problems, and extract an effective Hamiltonian that models the phenomena of interest.

This thesis is primarily concerned with the behaviour of electrons. From the observation that the nucleus is much heavier than an electron, we can separate the dynamics of the nuclei and the dynamics of the electrons using the Born-Oppenheimer approximation. Under the approximation of fixed $\mathbf{R}$, the Hamiltonian describing the behaviour of the electrons and their interactions with nuclei is simplified as

$$H_{\text{el}} = -\frac{1}{2}\sum_i \nabla_i^2 + \sum_i U_{\text{lattice},i}(\mathbf{r}_i) + \frac{1}{2}\sum_{i\neq j}\frac{1}{|\mathbf{r}_i - \mathbf{r}_j|}, \tag{2.2}$$

where the Coulomb potential between the electrons and the nucleus ions is denoted as $U_{\text{lattice},i}(\mathbf{r}_i)$ which is a one-body scalar term, and only the valence electrons are considered here.

The electron Hamiltonian in the position coordinate may be written in the operator form using the Dirac notation as

$$\hat{H}_{\text{el}} = \int \mathrm{d}\mathbf{r}\ \mathrm{d}\mathbf{r}'|\mathbf{r}\rangle\langle\mathbf{r}|\hat{H}_{\text{el}}|\mathbf{r}'\rangle\langle\mathbf{r}'|, \tag{2.3}$$

where $H_{\text{el}} = \langle\mathbf{r}|\hat{H}_{\text{el}}|\mathbf{r}'\rangle$. The kinetic term and the Coulomb potential of the electron and the nuclear is diagonal in the coordinate basis, and we have

$$\nabla_i^2 = \langle\mathbf{r}|\hat{\nabla}_i^2|\mathbf{r}'\rangle\delta(\mathbf{r}-\mathbf{r}'), \quad U_{\text{lattice},i}(\mathbf{r}_i) = \langle\mathbf{r}|\hat{U}_{\text{lattice},i}|\mathbf{r}'\rangle\delta(\mathbf{r}-\mathbf{r}'). \tag{2.4}$$

The Hamiltonian could be represented with the field operator in second quantisation as

$$\hat{H} = \int d\mathbf{r}\hat{\psi}^\dagger(\mathbf{r})(-\frac{\nabla^2}{2} + U(\mathbf{r}))\hat{\psi}(\mathbf{r}) + \frac{1}{2}\int d\mathbf{r}d\mathbf{r}\hat{\psi}^\dagger(\mathbf{r})\hat{\psi}^\dagger(\mathbf{r}')V(|\mathbf{r}-\mathbf{r}'|)\hat{\psi}(\mathbf{r}')\hat{\psi}(\mathbf{r}) \tag{2.5}$$

where $\hat{\psi}(\mathbf{r})$ is the fermionic field operator defined in the position coordinate basis $\{|\mathbf{r}\rangle\}$, and $\hat{\psi}^\dagger(\mathbf{r})$ is its conjugate. In the following discussion, I will use the notation in Ref. [44]

Since the spin degrees of freedom are generally a topic of interest, they have been added to the formulae in the following discussion. The field operator could be expanded in a basis of single-particle wave functions as

$$\begin{aligned}\hat{\psi}_\sigma(\mathbf{r}) &\equiv \sum_\lambda \hat{c}_{\lambda,\sigma}\phi_\lambda(\mathbf{r}) \\ \hat{\psi}_\sigma^\dagger(\mathbf{r}) &\equiv \sum_\lambda \hat{c}_\lambda^\dagger\phi_{\lambda,\sigma}^*(\mathbf{r})\end{aligned} \tag{2.6}$$

where $\sigma \in \{\uparrow,\downarrow\}$ represents the spin degrees of freedom, λ represents the collection of good quantum numbers of the particles, excluding the spin, $\hat{c}$ ($\hat{c}^\dagger$) is the fermionic annihilation (creation) operation acting as

$$\hat{c}_{\lambda,\sigma}^\dagger\,|\text{vac}\rangle = |\lambda,\sigma\rangle\,,$$

with the anti-commutation relation $\{\hat{c}_{\lambda_1,\sigma_1}, \hat{c}^\dagger_{\lambda_2,\sigma_2}\} = \delta_{\lambda_1,\lambda_2}\delta_{\sigma_1,\sigma_2}$, and the wave function is $\phi_\lambda(\mathbf{r}) = \langle \mathbf{r}|\lambda\rangle$.

The Hamiltonian in the new basis becomes

$$\hat{H} = \sum_{\lambda_1,\lambda_2} t_{\lambda_1\lambda_2}\hat{c}^\dagger_{\lambda_1,\sigma}\hat{c}_{\lambda_2,\sigma} + \sum_{\lambda_1,\lambda_2,\lambda_3,\lambda_4} V_{\lambda_1\lambda_2\lambda_3\lambda_4}\hat{c}^\dagger_{\lambda_1,\sigma'}\hat{c}^\dagger_{\lambda_2,\sigma}\hat{c}_{\lambda_3,\sigma}\hat{c}_{\lambda_4,\sigma'}. \tag{2.7}$$

Here, the hopping term is defined as

$$t_{\lambda_1\lambda_2} = \int d\mathbf{r}\phi^*_{\lambda_1}(\mathbf{r})(-\frac{\nabla^2}{2} + U(\mathbf{r}))\phi_{\lambda_2}(\mathbf{r}) \tag{2.8}$$

and the Coulomb interaction term is defined as

$$V_{\lambda_1\lambda_2\lambda_3\lambda_4} = \frac{1}{2}\int d\mathbf{r}'d\mathbf{r}\phi^*_{\lambda_1}(\mathbf{r})\phi^*_{\lambda_2}(\mathbf{r}')V(|\mathbf{r}-\mathbf{r}'|)\phi_{\lambda_3}(\mathbf{r}')\phi_{\lambda_4}(\mathbf{r}). \tag{2.9}$$

The matrix is Hermitian, i.e., $t_{\lambda_1\lambda_2} = t^*_{\lambda_1\lambda_2}$, and $V_{\lambda_1\lambda_2\lambda_3\lambda_4} = V^*_{\lambda_4\lambda_3\lambda_2\lambda_1}$. The Coulomb interaction matrix has an additional symmetry $V_{\lambda_1\lambda_2\lambda_3\lambda_4} = V_{\lambda_2\lambda_1\lambda_4\lambda_3}$.

From the Hamiltonian in Eq. (2.5), one can deduce various model electron Hamiltonians appropriate for different physical circumstances. The explicit form of the Hamiltonian depends on the single-particle bases introduced in Eq. (2.6). A representative model Hamiltonian is illustrated below, following the discussion in [44].

Electrons in a periodic potential conserves translational symmetry. A natural and common basis is the plane wave basis which is useful to describe assemblies of atoms that are close to free electrons. The operator takes the form of

$$\hat{\psi}_\sigma(\mathbf{r}) = \frac{1}{\sqrt{V_c}}\sum_{\mathbf{k},\mathbf{G}} e^{i(\mathbf{k}+\mathbf{G})\cdot\mathbf{r}}\hat{c}_{\mathbf{k}+\mathbf{G},\sigma}, \tag{2.10}$$

where $\mathbf{k}$ is the crystal momentum and $\mathbf{G}$ is a reciprocal lattice vector. The Hamiltonian in the plane wave basis reads

$$\hat{H} = \sum_{\mathbf{k},\mathbf{G},\mathbf{G}',\sigma} h_{\mathbf{k},\mathbf{G}-\mathbf{G}'}\hat{c}^\dagger_{\mathbf{k}+\mathbf{G},\sigma}\hat{c}_{\mathbf{k}+\mathbf{G}',\sigma} + \sum_{\substack{\mathbf{k},\mathbf{k}',p\\ \mathbf{G},\mathbf{G}',\sigma,\sigma'}} V_p\hat{c}^\dagger_{\mathbf{k}+\mathbf{G}+p,\sigma}\hat{c}^\dagger_{\mathbf{k}'+\mathbf{G}'-p,\sigma'}\hat{c}_{\mathbf{k}'+\mathbf{G}',\sigma'}\hat{c}_{\mathbf{k}+\mathbf{G},\sigma} \tag{2.11}$$

where $h_{\mathbf{k},\mathbf{G}-\mathbf{G}'} = \left[(\mathbf{k}+\mathbf{G})^2\delta_{\mathbf{G},\mathbf{G}'} + U_{\mathbf{G}-\mathbf{G}'}\right]$ and $V_p = 1/(2V_c)\int d\mathbf{r}e^{-ip\cdot\mathbf{r}}V(|\mathbf{r}|)$ [44]. I refer to Ref. [44] for other examples in solids and Ref. [43] for molecule Hamiltonians.

2.1.2 Low rank factorisition

The above Hamiltonian in Eq. (2.7) is written in physicist convention, and one could rearrange the form in chemist convention as

$$\hat{H} = \sum_{\lambda_1,\lambda_2} T_{\lambda_1\lambda_2}\hat{c}^\dagger_{\lambda_1,\sigma}\hat{c}_{\lambda_2,\sigma} + \sum_{\lambda_1,\lambda_2,\lambda_3,\lambda_4} G_{\lambda_1\lambda_2\lambda_3\lambda_4}\hat{c}^\dagger_{\lambda_1,\sigma'}\hat{c}_{\lambda_2,\sigma}\hat{c}^\dagger_{\lambda_3,\sigma}\hat{c}_{\lambda_4,\sigma'} \tag{2.12}$$

with the ordering $\hat{c}^\dagger\hat{c}\hat{c}^\dagger\hat{c}$ rather than $\hat{c}^\dagger\hat{c}^\dagger\hat{c}\hat{c}$. Using real basis functions, such as molecular orbitals, the interaction matrix has the symmetry

$$T_{\lambda_1\lambda_2} = T_{\lambda_2\lambda_1}, \tag{2.13}$$

and

$$G_{\lambda_1\lambda_2\lambda_3\lambda_4} = G_{\lambda_4\lambda_3\lambda_2\lambda_1} = G_{\lambda_1\lambda_2\lambda_4\lambda_3} = G_{\lambda_2\lambda_1\lambda_3\lambda_4} = G_{\lambda_3\lambda_4\lambda_1\lambda_2} = G_{\lambda_3\lambda_4\lambda_2\lambda_1} = G_{\lambda_4\lambda_3\lambda_1\lambda_2} = G_{\lambda_4\lambda_3\lambda_1\lambda_2}. \tag{2.14}$$

The Coulomb interaction matrix is a rank-4 tensor with dimension $N/2$ along each axis and could be reshaped to a $N^2/4 \times N^2/4$ matrix W.

The reason why the interaction term is reordered is to exploit the low rank nature of the Coulomb interaction matrix through matrix factorisation. As discussed in Ref. [99], the matrix W by Eq. (2.7) is a full rank matrix with $L = N^2/4$ and no reduction is possible. In the case of molecular electronic Hamiltonians, the matrix W by Eq. (2.12) is not full rank, and can be factorised by singular value decomposition with the number of terms $L = \mathcal{O}(N)$, which is much less than the worst case scenario [99]. The total number of distinct terms with different coefficients in the Hamiltonian is of the order $\mathcal{O}(N^2L)$, which shows an improvement over the original representation, which contains $\mathcal{O}(N^4)$ distinct terms. Recall that in the Hamiltonian simulation, the cost would be related to the number of terms in the Hamiltonian. Therefore, this rearrangement would save the resource cost in Hamiltonian simulations. This strategy has been intensively used in quantum chemistry simulation (see [99–101] for examples).

2.1.3 Fermionic to qubit mapping

The original Hamiltonian is expressed in a fermionic form. To represent fermions on a quantum computer, one should specify a fermion-to-qubit mapping. There are many choices of mapping, including Jordan-Wigner mapping, Bravyi-Kitaev mapping, and parity encoding (see [12] for a review). The choice of mapping will result in different circuit depths and qubit numbers. One of the commonly used mapping is the Jordan-Wigner mapping, which maps the fermionic operator $\hat{c}_j$ on the jth site (or mode) to the qubit Pauli operator as

$$\hat{c}_j \mapsto \frac{1}{2}\left(\hat{\sigma}_j^x + i\hat{\sigma}_j^y\right)\bigotimes_{i=1}^{j-1}\hat{\sigma}_i^z, \quad \hat{c}_j^\dagger \mapsto \frac{1}{2}\left(\hat{\sigma}_j^x - i\hat{\sigma}_j^y\right)\bigotimes_{i=1}^{j-1}\hat{\sigma}_i^z. \tag{2.15}$$

with Pauli operators $\hat{\sigma}_j^\alpha$, $\alpha = (x, y, z)$ acting on the jth qubit. It is straightforward to have the occupation number operator

$$\hat{n}_j = \hat{c}_j^\dagger\hat{c}_j \mapsto \frac{1}{2}\left(\hat{\mathcal{I}} - \hat{\sigma}_j^z\right),$$

and the two-body term

$$\hat{n}_j\hat{n}_k \mapsto \frac{1}{4}\left(\hat{\mathcal{I}} + \hat{\sigma}_j^z\hat{\sigma}_k^z - \hat{\sigma}_j^z - \hat{\sigma}_k^z\right).$$

With the mapping, an N-qubit Hamiltonian can be expressed in the Pauli basis as

$$\hat{H} = \sum_i h_i P_i, \tag{2.16}$$

where $P_i \in \{\hat{\mathcal{I}}, \hat{\sigma}_j^x, \hat{\sigma}_j^y, \hat{\sigma}_j^z\}^{\otimes N}$ is a tensor product of single-qubit Pauli operators and h_i is the corresponding strength.

2.1.4 A general representation of quantum systems

A major challenge in studying quantum many-body physics stems from the hardness of efficient representation of many-body wave functions. To motivate a general description of a quantum system, I start from a general description of a quantum state in the tensor representation. A general rank-n tensor is a multi-dimensional array with n indices denoted as $T_{i_1,i_2,\ldots,i_n}$. In quantum mechanics, it represents the wave function of an n-partite quantum state in the computational basis,

$$|\psi\rangle = \sum_{j_1,j_2,\ldots,j_n} \psi_{j_1,j_2,\ldots,j_n} |j_1\rangle |j_2\rangle \ldots |j_n\rangle . \tag{2.17}$$

One can see that directly storing a general quantum state in a classical memory is highly inefficient, with the cost of space resources increasing exponentially with the number of parties. This thus motivates us to find more efficient ways to represent quantum states. The deep observation by physicists is that quantum states in nature may only lie in a small subset of the whole Hilbert space, where the area law scaling may exist; for example, the ground state of certain gapped local Hamiltonians [102–104]. This enables the possibility of efficient classical representation of these quantum states. The overall idea is to decompose the rank-n tensor into a network of low-rank tensors. Take the matrix product state (MPS) ansatz as an example. As shown in Figure 2.1(a), the rank-n tensor is now decomposed into n low-rank tensors as

$$|\psi\rangle = \sum_{j_1,j_2,\ldots,j_n} \mathrm{Tr}[\alpha_1^{j_1} \ldots \alpha_n^{j_n}] |j_1\rangle |j_2\rangle \ldots |j_n\rangle . \tag{2.18}$$

Here each $\alpha_k^{j_k}$ is a rank-3 tensor (except for $\alpha_1^{j_1}$ and $\alpha_n^{j_n}$ whose rank is 2), and the index j_k is the physical index, with dimension 2 for the qubit case.

The MPS representation compresses the space of an n-partite state from $O(2^n)$ to $O(n\kappa^2)$, which is reduced from exponential to linear scaling with the particle number n. This enormous reduction is based on the pre-knowledge of the weakly entangled state under the geometrically local interactions in a one-dimensional gapped system [103]. Many different classical tensor networks have been proposed for different problems (see [105] for a review).

Nevertheless, it is likely that certain quantum systems, such as the Fermi-Hubbard model and molecular systems, which have nonlocal interactions, and 2D lattice problems [5], may not be efficiently described via classical methods. This motivates the idea of quantum simulation, i.e.,

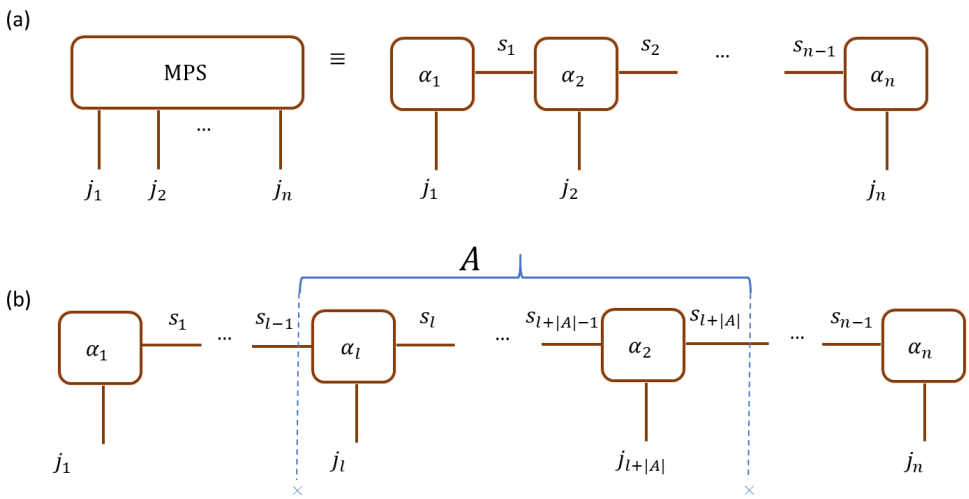


Figure 2.1: Illustration of a typical classical tensor network — the matrix product state.

using a controllable quantum system to simulate a target quantum problem. A quantum state generated from applying a unitary circuit to a certain initial state forms a quantum tensor, which could be high-rank, and can be naturally stored and manipulated with a quantum computer. In the literature, classical tensor network theory and quantum simulation are generally used as separate, distinct techniques in classical and quantum computing. In this thesis, we introduce quantum tensors to be general n-partite quantum states prepared by a quantum computer and classical tensors to be low-rank tensors stored in a classical computer. We refer to a combination of quantum and classical tensors as a hybrid tensor network.

Suppose we generate an n-partite quantum state by applying a unitary U_ψ to an initial state $|\bar{0}\rangle$ as $|\psi\rangle = U_\psi |\bar{0}\rangle$. As shown in Eq. (2.17), the quantum state can be regarded as a rank-n tensor in the computational basis. We can also introduce a classical index to the quantum state by applying different unitary operators as

$$|\psi^i\rangle = U_{\psi^i} |\bar{0}\rangle = \sum_{j_1,j_2,\ldots,j_n} \psi^i_{j_1,j_2,\ldots,j_n} |j_1\rangle |j_2\rangle \ldots |j_n\rangle , \tag{2.19}$$

where the classical index i relates to U_i which is a unitary operator. As a result, it as a whole forms a rank-$(n+1)$ tensor $\psi^i_{j_1,j_2,\ldots,j_n}$. We regard all these cases as quantum tensors, and the network connected with quantum tensors and classical tensors as a hybrid tensor network.

We can find that this hybrid quantum-classical tensor network provides the basis for general representation of many-body wave functions that is applicable to a broad range of problems. In Chapter 3, I will discuss this hybrid quantum-classical tensor network in detail.

2.2 Framework of hybrid quantum-classical computing

Section 2.1 introduces a framework of quantum state representation. In this section, I discuss hybrid quantum-classical computing introduced in Section 1.4. In particular, I introduce varia-

tional approaches to both static and dynamic problems, which provides useful tools throughout this thesis.

2.2.1 Variational algorithms

The variational method is a powerful classical tool for simulating quantum many-body systems [106–110]. The core idea is based on the intuition that some physical states with low energy belong to a small manifold of the whole Hilbert space [103, 104], similar to the discussion in Section 2.1. Recently, variational methods have been recently generalised to the quantum regime [72–74,83,111–126], which employ the power of quantum computing in representing quantum states. The trial state in variational quantum algorithms is prepared with a parameterised shallow quantum circuit [127–130], which is robust against a certain amount of device noise and is compatible with near-term noisy intermediate-scale quantum (NISQ) hardware [83,116]. Variational quantum algorithms can be utilised for finding energy spectra, also known as variational quantum eigensolvers (VQE) [69, 72, 73, 83, 113, 121], and simulating real-time Schrödinger evolution [80, 116] of closed systems. It is also worth noting that variational algorithms are not limited to energy minimisation and the simulation of unitary processes. Variational algorithms have been developed by the author and collaborators to simulate dissipative quantum dynamics, non-Hermitian dynamics, and imaginary-time evolution, which cannot be straightforwardly mapped to unitary gates [80, 120, 121, 131].

In Section 2.2.2, I show the basics of variational quantum algorithms, and exemplify their application in ground state preparation. In Section 2.2.3, I introduce a general framework of variational dynamics simulation. With this framework, I discuss the solution to two types of dynamic problems, namely, real-time evolution and imaginary-time evolution, in Section 2.2.3.2 and Section 2.2.3.3, respectively.

2.2.2 Variational state preparation

The variational quantum algorithm starts by preparing a quantum state with a quantum circuit, which can be written as

$$|\psi\rangle = \prod_{j=1}^{L} U_j(\theta_j) V_j |\psi_0\rangle \tag{2.20}$$

where U_j and $V_j \in \mathrm{SU}(D)$ represent the parameterised and unparameterised unitary quantum operations with $D = 2^n$, respectively.

The parameterised unitary quantum operation could be expressed by $U_j(\theta_j) = \exp(-i\theta_j G_j)$ with the Hermitian generator $G_j \in \mathrm{Herm}[\mathbb{C}^{D\times D}]$. The Hermitian generator can be decomposed into the Pauli operators, and $U_j(\theta_j)$ becomes $U_j(\theta_j) = \exp(-i\theta_j \sum_k \alpha_{j,k} P_{j,k}/2)$, with $P_{j,k} \in \{I, X, Y, Z\}^{\otimes n}$ being a tensor product of single-qubit Pauli operators. The parameterised quantum circuit given by Eq. (2.20) is also referred to as an ansatz, which is used

to approximate the target state. For simplicity, we consider the parameterised operation as $U_j(\theta_j) = \exp(-i\theta_j P_j/2)$ with P_j being a multi-qubit Pauli operator, which is widely used in quantum circuit construction. Although the circuit given by Eq. (2.20) is fixed, we can construct a quantum circuit in an adaptive way to increase its expressivity. This will be discussed in Chapter 5.

Despite its quantum nature, the original ground state preparation problem is converted into an optimisation problem over the variational parameters $\vec{\theta} = \{\theta_j\}$. The target state will be approximated by some optimised variational parameters $\vec{\theta}^*$, which may be found using the variational principle [131]. For example, a typical problem in quantum simulation is to find the ground state. If this is the objective, then we could minimise the energy with respect to the variational parameters $\mathcal{L}(\vec{\theta}) := \langle\psi(\vec{\theta})|H|\psi(\vec{\theta})\rangle$[1]. The ground state energy is given by

$$E_0 = \min_{\theta} \mathcal{L}(\vec{\theta}). \tag{2.21}$$

Rather than search for the ground state from the whole Hilbert space, with exponentially increasing computational complexity, we only search from a subset of the whole Hilbert space, characterised by the parameter θ, to find an approximate solution of the true ground state. This is also known as the Rayleigh-Ritz method.

At the kth iteration, a general strategy for searching for ground states is by updating the parameters

$$\vec{\theta}(k+1) = \vec{\theta}(k) - \eta(k)A^{-1}(\vec{\theta}(k))\nabla_{\vec{\theta}}\mathcal{L}(\vec{\theta}(k)), \tag{2.22}$$

where $\vec{\theta}(k)$ represent parameters that characterises the optimisation dynamics at the kth iteration, and the learning rate is denoted by $\eta(t)$. Here, we use $A(\vec{\theta}(k))$ to represent an invertible metric matrix with parameters $\theta(k)$. When the metric matrix is an identity matrix $A = I$, it reduces to conventional gradient descent.

Many classical algorithms have been proposed as a solution to the optimisation problem. These include the gradient descent, simultaneous perturbation stochastic approximation (SPSA), quasi-Newton methods, and natural gradient algorithms [132–135]. Subsequently, quantum versions have been adapted in this context. Among these algorithms, quantum natural gradient method [135], or equivalently variational imaginary-time evolution [136], is promising, since from a physics viewpoint, imaginary-time evolution defines a natural way of driving the initial state to the target state. I show in Section 2.3 that given a nonvanishing energy gap between the ground state and the first excited state, the state reached under imaginary-time evolution is exponentially close to the ground state. We shall see how variational imaginary-time evolution can be implemented in Section 2.2.3.3.

[1]Here, the Hamiltonian is mapped from an operator form $\hat{H}$ to a qubit form H. Nevertheless, these two forms are used interchangeably in this thesis when there is no ambiguity.

As noted in the literature, optimisation in variational quantum algorithms in general is hard [88], and one cannot guarantee avoiding local minima during the optimisation process. More rigorous state preparation methods without relying on variational ansatz will be discussed in Section 2.3.

2.2.3 Variational quantum dynamics simulation

2.2.3.1 General framework

To begin with, I introduce a generalised framework for variational simulation of quantum dynamics. Consider a differential equation

$$B(t)\frac{d}{dt}\left|v(t)\right\rangle = \left|dv(t)\right\rangle. \tag{2.23}$$

Here,

$$\left|dv(t)\right\rangle := \sum_j D_j(t)\left|v_j'(t)\right\rangle,$$

where $D_j(t)$ and $B(t)$ are general time dependent sparse (non-Hermitian) operators, $|v(t)\rangle$ is the system state, and each of $|v_j'(t)\rangle$ can be either $|v(t)\rangle$ or any known state that can be efficiently implemented with quantum hardware. The states $|v(t)\rangle$ and $|v_j'(t)\rangle$ can be unnormalised states as

$$\left|v(t)\right\rangle = \alpha(t)\left|\psi(t)\right\rangle$$

and

$$\left|v_j'(t)\right\rangle = \alpha_j'(t)\left|\psi_j'(t)\right\rangle$$

with normalisation factors $\alpha(t)$ and $\alpha_j'(t)$, respectively. In practice, $D_j(t)$ ($B(t)$) is assumed to be decomposed as a linear combination of Pauli operators $D_j(t) = \sum_i \lambda_i^j(t)\sigma_i$ with complex coefficients λ_i and a polynomial (with respect to the system size) number of tensor products of Pauli matrices $\sigma_i = \bigotimes_{i_k}\sigma_{i_k}$ with i_k denoting the i_kth qubit to be efficiently realised with quantum hardware.

In variational quantum simulation, instead of directly simulating the dynamics, the state is approximated by a parameterised quantum state as

$$\left|v(\vec{\theta}(t))\right\rangle = \alpha(\vec{\theta}_0(t))\left|\varphi(\vec{\theta}_1(t))\right\rangle$$

with the variational parameters $\vec{\theta} := (\vec{\theta}_0, \vec{\theta}_1)$. The original evolution can be projected to the evolution of the parameters via McLachlan's principle [137], which is given by

$$\min\left\|B(t)\frac{d}{dt}\left|v(\vec{\theta}(t))\right\rangle - \sum_j D_j(t)\left|v_j'(t)\right\rangle\right\|, \tag{2.24}$$

where $\| |\psi\rangle \| = \sqrt{\langle\psi|\psi\rangle}$ and the minimisation is over the parameters. By minimising the distance between the true evolution and the evolution of the parameterised state, the minimisation problem is mapped to solving a linear equation of parameters as

$$\sum_j A_{k,j}\dot{\theta}_j = C_k, \tag{2.25}$$

where $\dot{\theta}_j = d\theta_j/dt$ and the coefficients are linear sums of state overlaps that can be efficiently measured with quantum circuits [138].

The most physically relevant examples are real-time and imaginary-time evolution, which correspond to $B(t) = 1$ and $|dv(t)\rangle = -iH|v(t)\rangle$ or $|dv(t)\rangle = -(H - \langle v(t)|H|v(t)\rangle)|v(t)\rangle$, respectively. The expression of the coefficients A and C are specified in Section 2.2.3.2 and Section 2.2.3, respectively.

We remark that this method provides a technique for solving more general dynamic problems, such as dissipative dynamics in an open system. Since dissipative dynamics is not the focus of this thesis, I direct the interested readers to [80] for further details.

2.2.3.2 Variational simulation of real-time evolution

Real-time evolution governed by a time-independent Hamiltonian H is described by the Schrödinger equation,

$$\frac{d|\psi(t)\rangle}{dt} = -iH|\psi(t)\rangle, \tag{2.26}$$

which is apparently a special case of Eq. (2.23). Instead of directly simulating real-time dynamics with Hamiltonian simulation algorithms [13, 33, 38, 139, 140], variational quantum dynamics methods assume that the quantum state $|\psi(t)\rangle$ is prepared by a parameterised quantum circuit, $|\varphi(\vec{\theta}(t))\rangle = R_N(\theta_N)\dots R_k(\theta_k)\dots R_1(\theta_1)|\bar{0}\rangle$ with each gate $R_k(\theta_k)$ controlled by the real parameter θ_k, the reference state $|\bar{0}\rangle$, and parameters $\vec{\theta} = (\theta_1, \theta_2, \dots, \theta_N)$. According to McLachlan's variational principle [137], real-time dynamics of $|\psi(t)\rangle$ can be mapped to the evolution of the parameters $\vec{\theta}(t)$ by minimising the distance between the ideal evolution and the evolution induced by the parameterised trial state,

$$\delta\|(d/dt + iH)|\varphi(\vec{\theta}(t))\rangle\| = 0, \tag{2.27}$$

where $\| |\varphi\rangle \| = \sqrt{\langle\varphi|\varphi\rangle}$. The solution is found to be

$$\sum_j A_{k,j}\dot{\theta}_j = C_k, \tag{2.28}$$

with the matrix elements of A and C given by

$$A_{k,j} = \mathrm{Re}\left(\frac{\partial\langle\varphi(\vec{\theta}(t))|}{\partial\theta_k}\frac{\partial|\varphi(\vec{\theta}(t))\rangle}{\partial\theta_j}\right), \quad C_k = \mathrm{Im}\left(\langle\varphi(\vec{\theta}(t))|H\frac{\partial|\varphi(\vec{\theta}(t))\rangle}{\partial\theta_k}\right). \tag{2.29}$$

2.2.3.3 Variational simulation of imaginary-time evolution

The normalised state at imaginary-time τ is $|\psi(\tau)\rangle = \frac{e^{-H\tau}|\psi(0)\rangle}{\sqrt{\langle\psi(0)|e^{-2H\tau}|\psi(0)\rangle}}$ and the Wick-rotated Schrödinger equation is

$$\frac{d|\psi(\tau)\rangle}{d\tau} = -\left(H - E_\tau\right)|\psi(\tau)\rangle \tag{2.30}$$

where $E_\tau = \langle\psi(\tau)|H|\psi(\tau)\rangle$ is the expected energy at imaginary-time τ. It is also easy to check that it admits a similar form to Eq. (2.23). The ground state can be determined from the long time limit of the Wick-rotated Schrödinger equation $|\psi\rangle_{\mathrm{GS}} = \lim_{\tau\to\infty}|\psi(\tau)\rangle$. Consider a normalised trial state $|\varphi(\vec{\theta}(\tau))\rangle$ with real parameters $\vec{\theta}$ representing all the parameters. Imaginary-time evolution of the Schrödinger equation on the trial state space is given by

$$\sum_i \frac{\partial|\varphi(\vec{\theta}(\tau))\rangle}{\partial\theta_i}\dot{\theta}_i = -\left(H - E_\tau\right)|\varphi(\vec{\theta}(\tau))\rangle. \tag{2.31}$$

Applying McLachlan's variational principle, which minimises the distance between the evolution of trial state $\frac{\partial|\varphi(\vec{\theta}(\tau))\rangle}{\partial\tau}$ and $-(H - E_\tau)|\varphi(\vec{\theta}(\tau))\rangle\rangle$, we have

$$\delta\|\left(d/d\tau + H - E_\tau\right)|\varphi(\vec{\theta}(\tau))\rangle\| = 0, \tag{2.32}$$

which determines the evolution of the parameters

$$\sum_j A_{k,j}\dot{\theta}_j = -C_k, \tag{2.33}$$

with the matrix elements of A and C given by

$$A_{k,j} = \mathrm{Re}\left(\frac{\partial\langle\varphi(\vec{\theta}(\tau))|}{\partial\theta_k}\frac{\partial|\varphi(\vec{\theta}(\tau))\rangle}{\partial\theta_j}\right),\ C_k = \mathrm{Re}\left(\frac{\partial\langle\varphi(\vec{\theta}(\tau))|}{\partial\theta_k}H|\varphi(\vec{\theta}(\tau))\rangle\right). \tag{2.34}$$

Therefore, we can effectively simulate imaginary-time evolution by tracking the evolution of the parameters.

The measurement in the variational schemes for both real-time evolution and imaginary-time evolution can be implemented using the methods introduced in Section 2.5.1. Compared to realising imaginary-time evolution directly, variational methods are more efficient in their use of quantum resources, but may not guarantee an accurate solution since they rely on variational optimisation. In contrast, methods based on imaginary-time evolution realised by using linear-combination-of-unitaries formulae can guarantee the simulation accuracy. A comparison between the two methods be found in Section 2.3. The variational imaginary-time evolution method serves as a subroutine in variational quantum simulations. I will show its application and numerical verification in the context of hybrid quantum-classical computing in Chapter 3.

2.3 Eigenstate preparation

Recall the main objective and the formulation of the static problem in Section 1.2. In this thesis, we mainly focus on two tasks, which are to estimate the eigenenergy E_i, and to estimate an observable expectation on the target eigenstate $\langle u_i|\hat{O}|u_i\rangle$ within an error ε. A crucial step to achieving both aims is the effective preparation of the eigenstate[2]. In this section, I will mainly discuss eigenstate preparation by applying a spectral projector of the Hamiltonian H to an initial state. Compared to the hybrid quantum-classical computing methods that were introduced in Section 2.2, this method is rigorous and universal, with a theoretical guarantee of the simulation accuracy.

Before I introduce the methods developed in this thesis, I will first review related work on eigenstate preparation. Section 2.3.1 serves as a reference for a better understanding of and point of comparison of our work.

2.3.1 Related works

I first provide some comments on the variational algorithms introduced in Section 2.2. In the variational methods, one attempts to search for the ground state of a Hamiltonian using an ansatz circuit, which can be characterised by the associated variational parameters. One then adjusts and optimises the parameters in the quantum circuit based on the measured value of energy. The quantum approximate optimisation algorithm (QAOA) [111] is another type of variational method that aims to solve the combinatorial optimisation problems, and is based on a variational evolution combined with adiabatic evolution. Variational algorithms are suitable for near-term quantum computers, since they usually require a shallow circuit with very few ancillary qubits. The weakness of the variational methods is that the effectiveness is contingent on the choice of the ansatz, the validity of which varies for different Hamiltonian problems.

Adiabatic state preparation [141] is an experimentally-friendly non-variational ground-state method based on a time-dependent Hamiltonian evolution. In adiabatic state preparation, the ground state of a simple Hamiltonian H_0 is first prepared, and then slowly evolved under a Hamiltonian that gradually changes from H_0 to the target Hamiltonian H. Based on the adiabatic theorem, the resulting state is then close to the ground state. Unlike variational algorithms, the adiabatic algorithms are universal and valid without ansatz assumptions. In terms of practical usage, the adiabatic algorithm has two drawbacks. First, the required evolution time t depends inverse polynomially on the minimum spectral gap along the entire path from H_0 to H. Second, the required evolution time t is $\mathcal{O}(1/\varepsilon)$, where ε is the infidelity of the target state.

[2]Here, we do not have to prepare the eigenstate state deterministically, since only the expectation value with respect to the state is required.

Phase estimation [27, 28] is one commonly-used non-variational method, which prepares the eigenstate of a given unitary U based on controlled-U evolution, and complementary basis measurement on ancillary qubits. In the canonical phase estimation algorithm [27], in order to prepare the eigenstate of the Hamiltonian H, we introduce ω qubits prepared in the state $|+\rangle$, perform controlled-$e^{-2\pi i H t}$ gate sequentially from one ancillary qubit to the original system with different times $\{t\}$, perform inverse Fourier transform on the ancillary qubits, and finally measure the ancillary qubits. The required number of qubits ω is logarithmic in the inverse of the eigenenergy precision. The circuit depth, determined by the total controlled evolution time t, is $\mathcal{O}(1/\varepsilon)$ with target precision ε. The phase estimation can be improved to an iterative version with only one ancillary qubit [28]. However, similar to adiabatic evolution methods, the required circuit depth of the phase estimation is fundamentally $\mathcal{O}(1/\varepsilon)$, which is unfavourable if a high-precision preparation is required, but the coherent time is relatively short.

To further improve the efficiency of the above two nonvariational methods, linear-combination-of-unitary (LCU) methods have been proposed [29,142]. In Ref. [29], $\cos^M(H)$ is used as a spectral projector to prepare the ground state of H. The Hamiltonian function $\cos^M(H)$ is expressed as a linear combination of unitary $\cos^M(H) = \sum_i \alpha_i U_i$ with unitary operator U_i and coefficient α_i. The implementation of LCU formulae relies on some circuit oracles, and I refer to Refs. [36, 38] for details. The query complexity of the LCU method is $\mathcal{O}(\frac{1}{\Delta}\log(\frac{1}{\varepsilon})$ and the number of ancillary qubits required is $\mathcal{O}(\log(\frac{1}{\Delta}\log\frac{1}{\varepsilon}))$, where Δ is a known lower bound of the Hamiltonian energy gap and ε is the target precision. The query complexity shows an exponential improvement when compared to the phase estimation algorithm. At a later stage, Lin and Tong [31] proposed a more efficient method for ground state preparation. The key idea is to realise a polynomial approximation of the sign function, which serves as a spectral projector of H, using quantum signal processing (QSP) based on a block-encoding of H [32] (see Section 2.4.3). This method will be discussed in detail in Section F.2.

The difficulty of realising either the LCU- or QSP-based methods in the near-term is that they both require many ancillary qubits and a deep circuit; the oracles for realising a block-encoding of H in LCU and QSP are difficult to compile. The resource cost for implementing a block-encoding of H will be discussed in Section 9.3.

2.3.2 Imaginary-time evolution

The key idea in this section is that access to the physical properties of eigenstates can be achieved through applying a spectral projector to the initial state, which projects out the contributions from the other eigenstates. A natural way is to consider imaginary-time evolution,

$$g_\tau(H) := e^{-H\tau}, \tag{2.35}$$

which drives the system to the ground state in the long-time evolution under imaginary-time.

To expand on this point more concretely, the normalised imaginary-time-evolved state $\frac{e^{-H\tau}|\Psi(\tau)\rangle}{\|e^{-H\tau}|\Psi(\tau)\rangle\|}$ can be expressed as $e^{-(H-E_\tau)\tau}|\Psi(\tau)\rangle$, where $\|\cdot\|$ is the state vector norm, and we can show that $E_\tau = \langle\Psi(\tau)|H|\Psi(\tau)\rangle$. Indeed, the state satisfies the Wick-rotated Schördinger equation

$$\partial_\tau |\Psi(\tau)\rangle = -(H - E_\tau)|\Psi(\tau)\rangle \tag{2.36}$$

Intuitively, the ground state can be determined from the long time limit of the Wick-rotated Schrödinger equation

$$|u_0\rangle = \lim_{\tau\to\infty} |\psi(\tau)\rangle .$$

In the following, we shall see that under imaginary-time evolution, the spectral weight on the excited states are exponentially suppressed, and the ground state energy is exponentially close to E_0. The time-evolved state can be decomposed into the eigenstate basis of the Hamiltonian as

$$|\psi\rangle = \sum_j c_j |u_j\rangle , \tag{2.37}$$

where we denote energy eigenstates $|u_j\rangle$ and eigenenergies E_j, and the normalisation condition holds $\sum_j |c_j|^2 = 1$. We assume the eigenstates are non-degenerate and the energy gap $\Delta_{01} := E_1 - E_0$ is not exponentially small. Indeed, exponentially small Δ_{01} indicates that these states are indistinguishable, and we can hence regard these states as degenerate states. The time evolved state is

$$|\psi(\tau)\rangle = \sum_j c_j e^{-(E_j-E_0)\tau} |u_j\rangle , \tag{2.38}$$

and the distance to the ground state energy is now given by

$$\begin{aligned} \frac{\langle\psi(\tau)|H|\psi(\tau)\rangle}{\langle\psi(\tau)|\psi(\tau)\rangle} - E_0 &= \frac{\sum_j |c_j|^2 (E_j - E_0) e^{-2(E_j-E_0)\tau}}{\sum_j |c_j|^2 e^{-2(E-E_0)\tau}} \\ &\leq \frac{\sum_j |c_j|^2 (\max E - E_0) e^{-2(E_1-E_0)\tau}}{|c_0|^2} = \mathcal{O}(e^{-2\Delta_{01}\tau}). \end{aligned} \tag{2.39}$$

I also remark that the time-evolved state is exponentially close to the ground state since

$$\frac{\langle u_0|\psi(\tau)\rangle \langle\psi(\tau)|u_0\rangle}{\langle\psi(\tau)|\psi(\tau)\rangle} = \frac{|c_0|^2}{\sum_j |c_j|^2 e^{-2(E_j-E_0)\tau}} \geq \frac{|c_0|^2}{|c_0|^2 + (1-|c_0|^2) e^{-2(E_1-E_0)\tau}}. \tag{2.40}$$

2.3.3 Generalised imaginary-time evolution

In additional to the original imaginary-time evolution $g_\tau(H) = e^{-H\tau}$, we could consider generalised imaginary-time evolution, such as a Gaussian type function, $g_\tau(H) = e^{-H^2\tau^2}$. Here, we define a general matrix function acting on the Hamiltonian, which is expressed as

$$g(H) := \sum_{i=0}^{N-1} g(E_i) |u_i\rangle \langle u_i| , \tag{2.41}$$

where $g(h) : \mathbb{R} \to \mathbb{C}$ is a generic continuous-variable function determining the transformation of the energy spectrum of the Hamiltonian. To realise this eigenstate preparation process, we require $g(h)$ satisfying strictly non-increasing absolute value, $|g(h')| < |g(h)|$, $\forall h' > h > 0$ or $h' < h < 0$, and vanishing asymptotic value, $\lim_{\tau\to\infty} |g(\tau h')/g(\tau h)| = 0$, $\forall h' > h > 0$ or $h' < h < 0$ or alternatively $\lim_{\tau\to\infty} |g(\tau h)/g(0)| = 0$, $\forall |h| > 0$.

While the projector[3] is nonunitary by construction, we can effectively realise it by using a combination of real-time dynamics e^{-iHt_i} with real-time length t_i, which is given by

$$g_\tau(H) = \sum_i p_i e^{-iHt_i}. \tag{2.42}$$

The Fourier transformation provides such a natural and universal decomposition. Therefore, we can realise this nonunitary operation by a series of unitary operations. These unitary operations can be implemented using Hamiltonian simulation methods on a digital quantum simulator or an analogue quantum simulator. For the most commonly used projection operator, we have

$$e^{-\tau H} = \frac{1}{\pi} \int dx \frac{1}{1+x^2} e^{i\tau x H}, \quad \tau > 0 \tag{2.43}$$

$$e^{-\tau^2 H^2} = \frac{1}{2\sqrt{\pi}} \int dx e^{-x^2/4} e^{i\tau x H}. \tag{2.44}$$

also known as a Hubbard-Stratonavich transformation which is widely used in condensed matter physics and quantum field theories.

The maximum time complexity $\max_i t_i$, which determines the circuit complexity required to achieve an additive error ε, is logarithmic in the inverse error $\mathcal{O}(\log(\varepsilon^{-1}))$.

2.4 Quantum dynamics simulation

Simulation of quantum dynamics $U = e^{-iHt}$ is one of the most natural and promising application of quantum computing. As discussed in Section 2.3, dynamics simulation is an essential component for eigenstate preparation. In addition, it is also an important subroutine for many other quantum algorithms, like optimisation problems and open-system dynamics [25, 143]. In this section, I briefly review methods for quantum dynamics simulation.

2.4.1 Overview

Quantum dynamics simulation of $U = e^{-iHt}$ is also called Hamiltonian simulation in the community of quantum computing. The problem of approximating the unitary operator can be abstracted as follows.

Problem 1. *Given a Hamiltonian H on n qubits, evolution time t, output a unitary approximation, U, of the ideal unitary $U_0 = e^{-iHt}$ within an error ε, such that $\|U - U_0\| \leq \varepsilon$.*

[3] Herein, a projector is referred to as a spectral projector of H.

Suppose the Hamiltonian can be decomposed as $H = \sum_{l=1}^{L} H_l$. It is assumed that the Hamiltonian is sparse, that is, the number of terms in the Hamiltonian grows polynomially with respect to the system size, $L = \mathcal{O}(\text{Poly}(N))$. The most natural and direct Hamiltonian simulation method is to apply the Lie-Trotter-Suzuki (Trotter) formulae, which approximate the real-time evolution operator $U(t)$ by the product of the evolution of the summands, e^{itH_l}, which are relatively easy to implement. Historically, Lloyd proposed first-order Trotter methods for universal Hamiltonian simulation [13]; Suzuki extended it to high-order product formula [139]; Berry *et al.* provided a detailed error analysis [35]; Childs *et al.* tightened the error analysis using the commutator information in the Hamiltonian [144]. To improve the dependence on the Hamiltonian sparsity L, Campbell proposed the qDRIFT algorithm, which randomly sample the summands to realise the Trotter formula [145]. However, the remainder of a kth-order Trotter formula (Trotter error) is usually large, which is polynomial in the order k, and the required gate number is consequently polynomially dependent on the accuracy requirement as $\mathcal{O}(\text{Poly}(1/\varepsilon)$.

In recent years, there have been developments in Hamiltonian simulation algorithms by implementing different linear-combination-of-unitary (LCU) formulae [36]. In 2015, Berry *et al.* proposed a simulation algorithm using the LCU formula introduced by Taylor series expansion of $U(t)$ [38]. Low and Chuang subsequently proposed the quantum signal processing (QSP) algorithm [32,33], which achieves optimal Hamiltonian simulation. In the QSP algorithm, H is embedded into a larger unitary operator, usually based on a LCU decomposition of H, $H = \sum_{l=1}^{L} \alpha_l U_l$; a polynomial of the Hamiltonian (such as Taylor series [38] or Jacobi-Anger expansion [32]) can be realised through coherent phase iterations and post-selection. These post-Trotter methods are able to capture dominant terms in the time evolution $U(t)$ with few resources, leading to a logarithmic gate-number dependence on the accuracy requirement as $\mathcal{O}(\text{Poly}(\log(1/\varepsilon))$.

Nearly optimal simulation by quantum signal processing will be reviewed in Section 2.4.3, which achieves the scaling

$$\mathcal{O}(\lambda t + \log(1/\varepsilon)/\log\log(1/\varepsilon)) \tag{2.45}$$

where $\lambda = \sum_l |\alpha_l|$ is the sum of absolute values of the Hamiltonian coefficients and ε is the target precision. It is worth noting that the scaling on t and ε is separate, which shows advantages over LCU. This is attributed to the fact that Jacobi-Anger expansion approximates the operator during a full time period t, instead of dividing the time into ν segments; the latter unavoidably results in a coupling of t and ε, and thus has the multiplicative error dependence in the circuit complexity as $\mathcal{O}(\lambda t \log(1/\varepsilon)/\log\log(1/\varepsilon))$. I refer to Section 2.4.3 for a detailed discussion of the complexity, which will be used in Chapter 9.

These advanced algorithms, however, require the implementation of LCU formulae or block encoding of Hamiltonians, which often costs many ancillary qubits and multi-controlled Toffoli gates. Resource analysis at gate level will be discussed in Chapter 9.

2.4.2 Product formula

One notable quantum dynamics simulation method is the product formula [144]. Typical product formulae are the Trotter-Suzuki formulae, which shall be referred to as Trotter formulae or Trotter methods in this thesis. In the Trotter methods, we first divide the real-time evolution into ν short-time segments,

$$e^{-iHt} = \left(e^{-iHx}\right)^{\nu} \tag{2.46}$$

where each time step is defined as $x := t/\nu$, and approximates the unitary by a successive product of short-time evolution. The first-order Trotter formula is

$$S_1(x) = \prod_{l=1}^{L} e^{-ixH_l} \tag{2.47}$$

and the second-order Trotter formula is

$$S_2(x) = \prod_{l=L}^{1} e^{-i(x/2)H_l} \prod_{l=1}^{L} e^{-i(x/2)H_l}. \tag{2.48}$$

A general $(2k)$th-order Trotter formula is

$$S_{2k}(x) = \left[S_{2k-2}\left(p_k x\right)\right]^2 S_{2k-2}\left(\left(1-4p_k\right)x\right)\left[S_{2k-2}\left(p_k x\right)\right]^2 \tag{2.49}$$

with $p_k := 1/\left(4 - 4^{1/(2k-1)}\right)$ for $k \geq 1$. The zeroth-order Trotter formula is denoted as $S_0(x) = I$.

Denote the (multiplicative) remainder of the Trotter formulae as

$$\begin{aligned} V_0(x) &= U(x), \\ V_1(x) &= U(x)S_1(x)^{\dagger}, \\ V_{2k}(x) &= K_{2k}(-x)U(x)K_{2k}(x)^{\dagger}, \end{aligned} \tag{2.50}$$

with $U(x) = e^{-ixH}$. Suzuki proves that,

$$S_k(x) = U(x) + \mathcal{O}(x^{k+1}) \tag{2.51}$$

for $k = 1$ or even positive k. As a result, the remainder $V_k(x)$ will only contain terms of x^q with $q \geq k+1$.

To improve the error dependence of the Trotterisation method, it is crucial to study the remainder of the Trotter formulae. I will show how the Trotter error can be compensated by LCU in Chapter 9.

2.4.3 Nearly optimal simulation by quantum signal processing

The quantum signal processing (QSP) algorithm by Low and Chuang [32,33,37] aims to expand the real-time evolution using some polynomial functions of eigenvalues. To realise this, we first encode the n-qubit Hamiltonian $H = \sum_l \alpha_l U_l$ to a $(n_L + n)$-qubit unitary, select(H)

$$\text{select}(H) := \sum_{l=1}^{L} |l\rangle\langle l| \otimes H_l, \tag{2.52}$$

where $n_L = \lceil \log(L) \rceil$ and U_l is a unitary operator. Denote

$$|G\rangle := \frac{1}{\sqrt{\lambda}} \sum_{l=1}^{L} \sqrt{\alpha_l} |l\rangle, \tag{2.53}$$

with $\lambda = \sum_{l=1}^{L} |\alpha_l|$, then we have

$$H = \lambda(\langle G| \otimes I) \operatorname{select}(H)(|G\rangle \otimes I), \tag{2.54}$$

which indicates that H is block-encoded into $\operatorname{select}(H)$.

The eigenvalues of H are related to the ones of $\operatorname{select}(H)$. Denote the spectral decomposition of H as follows,

$$H = \sum_j E_j |u_j\rangle\langle u_j| = \lambda \sum_j h_j |u_j\rangle\langle u_j| \tag{2.55}$$

where $h_j = E_j/\lambda$ is the jth normalised eigenvalue. We have $h_j \in [-1, 1]$ since $\|H\| \leq \lambda$ by the triangle inequality. For each eigenvalue $h_j \in (-1, 1)$, the following operator

$$-iQ := -i((2|G\rangle\langle G| - I) \otimes I) \operatorname{select}(H), \tag{2.56}$$

has two corresponding eigenvalues $h_j^{\pm}$, which are given by

$$h_j^{\pm} = \mp\sqrt{1 - h_j^2} - ih_j = \mp e^{\pm i \arcsin(h_j)} = \mp e^{\pm i\theta_j}, \tag{2.57}$$

with eigenvectors $\left|h_j^{\pm}\right\rangle = \left(|G_j\rangle \pm i \left|G_j^{\perp}\right\rangle\right)/\sqrt{2}$, where

$$\begin{aligned} |G_j\rangle &:= |G\rangle \otimes |u_j\rangle \\ \left|G_j^{\perp}\right\rangle &:= \frac{h_j |G_j\rangle - \operatorname{select}(H) |G_j\rangle}{\sqrt{1 - h_j^2}}. \end{aligned} \tag{2.58}$$

The QSP algorithm aims to construct the Hamiltonian evolution operator using the operator $-iQ$ with a sequence of operations called phase iterates. To realise this, we introduce an additional ancillary qubit and define the phase iteration operator, following the discussion in [37] rather closely,

$$V_\phi := \left(e^{-i\phi\sigma^z/2} \otimes I\right) (|+\rangle\langle+| \otimes I + |-\rangle\langle-| \otimes (-iQ)) \left(e^{i\phi\sigma^z/2} \otimes I\right). \tag{2.59}$$

Denote the spectral decomposition of $-iQ$ to be

$$-iQ = \sum_l e^{i\theta_l} |q_l\rangle \langle q_l|, \tag{2.60}$$

where θ_l contains the values of $\pm\theta_j = \pm \arcsin(h_j)$. Then, the phase iteration operator V_ϕ can be rewritten as

$$V_\phi = \sum_l e^{i\theta_l/2} R_\phi(\theta_l) \otimes |q_l\rangle \langle q_l| \tag{2.61}$$

where

$$R_\phi(\theta) := e^{-i\theta\sigma(\phi)/2}, \quad \sigma(\phi) := \cos(\phi)\sigma^x + \sin(\phi)\sigma^y \tag{2.62}$$

As a result, each eigenvalue $e^{i\theta_l}$ of $-iQ$ is embedded into V_ϕ as a rotation operator $R_\phi(\theta_l)$ with a tuneable angle ϕ.

The idea of QSP is to concatenate the rotation operators $\{R_{\phi_m}(\theta_l)\}_{m=1}^M$ with well-designed angles $\phi_1, \ldots, \phi_M$ so that the resulting operator can be realised,

$$R_{\phi_M}(\theta)\ldots R_{\phi_1}(\theta) = A\left(\cos\frac{\theta}{2}\right)I + iB\left(\cos\frac{\theta}{2}\right)\sigma^z + i\cos\frac{\theta}{2}C\left(\sin\frac{\theta}{2}\right)\sigma^x + i\cos\frac{\theta}{2}D\left(\sin\frac{\theta}{2}\right)\sigma^y \tag{2.63}$$

where the polynomials $A(x)$ and $C(x)$ can be engineered to be some target functions of the eigenvalues θ_l and hence the functions of h_j. The unwanted phase $e^{i\theta_l/2}$ in the phase iteration V_ϕ in Eq. (2.61) will be cancelled out by alternating between V_ϕ and $V_{\phi+\pi}^\dagger$,

$$V := V_{\phi_M+\pi}^\dagger V_{\phi_{M-1}} \ldots V_{\phi_2+\pi}^\dagger V_{\phi_1}. \tag{2.64}$$

The polynomials with only $A(x)$ and $C(x)$ can be extracted by preparing the ancillary qubit in the state $|+\rangle$, performing the operations in Eq. (2.64), and post-selecting the ancillary qubit to $|+\rangle$ (so that the polynomials $B(x)$ and $D(x)$ will be cancelled out). The resulting polynomial is

$$\mathrm{Poly}(\theta) = A\left(\cos\frac{\theta}{2}\right) + i\cos\left(\frac{\theta}{2}\right)C\left(\sin\frac{\theta}{2}\right) \tag{2.65}$$

In our case, the target is to realise the time evolution operator, which can be written as,

$$e^{-iHt} = \sum_{j=0}^{N-1} e^{-i(\lambda t)h_j}|u_j\rangle\langle u_j| = \sum_{j=0}^{N-1} e^{-i(\lambda t)\sin(\theta_j)}|u_j\rangle\langle u_j| \tag{2.66}$$

To realise e^{-iHt} using Eq. (2.64), the Jacobi-Anger expansion is used,

$$e^{i\sin(\theta)t} = \sum_{k=-\infty}^{\infty} J_k(t)e^{ik\theta} \tag{2.67}$$

To approximate $e^{-i(\lambda t)\sin\theta_j}$, we use the polynomial Poly (θ_j) constructed from M phase iteration operators. It is shown in [33] that with proper choices of the angles $\phi_1, \ldots, \phi_M$, we can realise the expansion in Eq. (2.67) at order $q := \frac{M}{2} + 1$, giving an approximation with error at most

$$2\sum_{k=q}^{\infty}|J_k(t)| \leq \frac{4t^q}{2^q q!} \tag{2.68}$$

The detailed classical algorithm to find the angles $\phi_1, \ldots, \phi_M$ to realise the target polynomials will not be introduced here.

Remarkably, quantum signal processing achieves nearly optimal scaling in terms of query complexity as

$$\mathcal{O}(\lambda t + \log(1/\varepsilon)/\log\log(1/\varepsilon)), \tag{2.69}$$

where $\lambda = \sum_l |\alpha_l|$ is the sum of coefficients of the Hamiltonian and ε is the target precision. Note that a block encoding of H could be difficult to implement, as it requires a huge overhead in circuit compilation. The resource cost for circuit compilation will be discussed in Section 9.3.

2.4.4 Approximation of time-evolved states

In many applications, we are simply interested in the evolution of a specific quantum state. Therefore, it is unnecessary to approximate the unitary $U = e^{-iHt}$. This problem is less demanding, and can be abstracted as follows.

Problem 2. *Given an initial state $|\psi(t=0)\rangle$, a Hamiltonian H on n qubits, evolution time t, output a state approximation, $|\Phi(t)\rangle$, of the ideal time-evolved state $|\psi(t)\rangle = e^{-iHt}|\psi(t)\rangle$ within an error ε, such that $\| |\Phi(t)\rangle - |\psi(t)\rangle \| \leq \varepsilon$.*

The error of the unitary approximation is the upper bound of the state error as

$$\|U - U_0\| = \max_{\phi} \|U|\phi\rangle - U_0|\phi\rangle\|. \tag{2.70}$$

We see that the resource requirement for state approximation is lower than that for the approximation of the unitary operation. It is also worth noting the connection to the variational methods introduced in Section 2.2.3. Variational methods aim to solve such a problem and approximate the evolution from a specific state, although it cannot guarantee the simulation accuracy. In Chapter 5, I will introduce a possible approach to Problem 2 with a theoretical guarantee of the simulation accuracy.

There are also other specific settings, such as a fixed electron or spin number [146], local observables [147], lattice models [144] and fixed observables [148], where quantum resources can be further reduced.

2.4.5 Circuit implementation of multi-qubit Pauli rotation e^{-iPt}

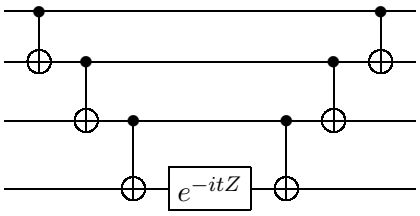
In this section, I give a pedagogical explanation of how an operator e^{-iPt} can be implemented with a quantum circuit. Assume $P = P_1 \otimes P_2 \otimes \cdots \otimes P_n$, where $P_j \in \{X, Y, Z\}$. Notice that $H_\mathrm{d} X H_\mathrm{d} = Z$ (H_d means the Hadamard gate) and $R_x(\frac{\pi}{2}) Y R_x(-\frac{\pi}{2}) = Z$. Define a map G from a Pauli operator to a gate, such that $G(Z) = I, G(X) = H_\mathrm{d}, G(Y) = R_x(-\frac{\pi}{2})$ and $G^{-1}(P_j) P_j G(P_j) = Z$. Then e^{-iPt} is decomposed as

$$e^{-iPt} = \exp(-it \bigotimes_{j=1}^{n} P_j) = \exp(-it \bigotimes_{j=1}^{n} G^{-1}(P_j) Z^{\otimes n} \bigotimes_{j=1}^{n} G(P_j)) = \bigotimes_{j=1}^{n} G^{-1}(P_j) e^{-iZ^{\otimes n}t} \bigotimes_{j=1}^{n} G(P_j), \tag{2.71}$$

where $G(P_j)$ can be implemented by simple single-qubit gates. To implement $e^{-iZ^{\otimes n}t}$, one can first apply a series of CNOT gates

$$CNOTs = CNOT_{1\to 2} CNOT_{2\to 3} \cdots CNOT_{n-1\to n}$$

so that, for any input $|\vec{z}\rangle$ from the computational basis, the last qubit will be transformed into $|z_1 \oplus z_2 \oplus \cdots \oplus z_n\rangle$. Then, we can simply apply e^{-itZ} on the last qubit and finally apply $CNOTs^{-1}$. An example for the case of $n = 4$ is shown below.



Through such a construction, one can easily check that for all inputs $|z\rangle$ from the computational basis, a phase e^{-it} will be applied to it when $z_1 \oplus z_2 \oplus \cdots \oplus z_n = 1$ and all the other inputs will remain the same. Thus, the effect of the circuit is equivalent to the operator $e^{-iZ^{\otimes n}t}$.

Finally, we can use the circuit for $e^{-iZ^{\otimes n}t}$ to implement e^{-iPt}, which can be realised by applying single-qubit operations without introducing additional two-qubit gates. In such a construction, $2n - 2$ CNOT gates will be used. In practical implementation, the depth of quantum circuits may be further reduced on a case-by-case basis.

2.5 Quantum operations

2.5.1 Generalised quantum operation

In this section, I introduce the concept of *generalised quantum operations* and its implementation, which was initially proposed in [95]. The concept will frequently be used in this thesis, such as in Section 2.2, and the entirety of Part II. The generalised quantum operation is defined as

$$\Phi(\rho) = \mathrm{Tr}_E[U(\rho \otimes |0\rangle \langle 0|_E)V^\dagger], \tag{2.72}$$

where U and V could be different unitary operators that apply jointly on ρ and $|0\rangle_E$, and $|0\rangle_E \langle 0|_E$ is short-handed as $|0\rangle \langle 0|_E$. The following shows several properties of the generalised quantum operation $\Phi(\rho)$.

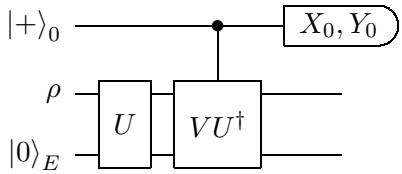
The generalised quantum operation $\Phi(\rho)$ has a bounded Schatten norm. Specifically, the Schatten norm of a matrix is $\|M\|_p = \mathrm{Tr}[|M|^p]^{1/p}$ for $p \in [1, \infty)$ and we have

$$\|\Phi(\rho)\|_p \leq \|\Phi(\rho)\|_1 \leq \|U(\rho \otimes |0\rangle \langle 0|_E)V^\dagger\|_1 = \|\rho \otimes |0\rangle \langle 0|_E\|_1 = \|\rho\|_1. \tag{2.73}$$

Here the two inequalities follow from the non-increase of the Schatten norm over p and the non-increase of the trace norm under a partial trace. Nevertheless, since $\Phi(\rho)$ could be complex, it might not be a quantum channel in general.

The real and imaginary part of $\Phi(\rho)$ could be expressed as a linear combination of completely positive trace non-increasing quantum channels. Specifically, they could be obtained with the

following circuit.



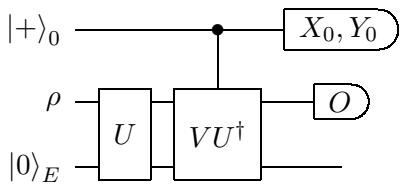
The output state before measurement is

$$\begin{aligned}\rho_{\text{out}} = &|0\rangle\langle 0|_0 U\rho\otimes|0\rangle\langle 0|_E U^\dagger + |0\rangle\langle 1|_0 U\rho\otimes|0\rangle\langle 0|_E V^\dagger \\ &+ |1\rangle\langle 0|_0 V\rho\otimes|0\rangle\langle 0|_E U^\dagger + |1\rangle\langle 1|_0 V\rho\otimes|0\rangle\langle 0|_E V^\dagger.\end{aligned} \tag{2.74}$$

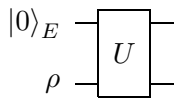
The real and imaginary part of $\Phi(\rho)$ can be obtained from the X and Y measurements

$$\text{Re}[\Phi(\rho)] = \text{Tr}_0[\rho_{\text{out}} X_0], \quad \text{Im}[\Phi(\rho)] = \text{Tr}_0[\rho_{\text{out}} Y_0]. \tag{2.75}$$

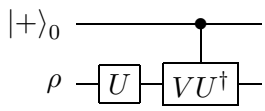
The measurement of the output state is realised in a similar way. For example, the real and imaginary parts of $\text{Tr}[\Phi(\rho)O]$ could be realised with the following circuit.



When $U = V$, it reduces to a quantum channel $\mathcal{N}$ with $\mathcal{N}(\rho) = \text{Tr}_E[U(\rho\otimes|0\rangle\langle 0|_E)U^\dagger]$ and the quantum circuit implementation



When there is no ancillary E, it becomes $\Phi(\rho) = U\rho V^\dagger$, with the circuit



which plays a key role in our explicit scheme in Section 6.3.

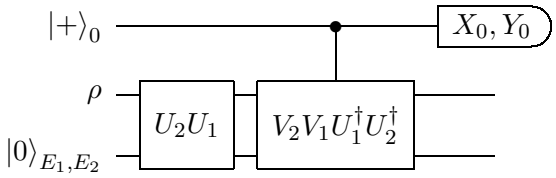
Given two generalised quantum operations

$$\begin{aligned}\Phi_1(\rho) &= \text{Tr}_{E_1}[U_1(\rho\otimes|0\rangle\langle 0|_{E_1})V_1^\dagger], \\ \Phi_2(\rho) &= \text{Tr}_{E_2}[U_2(\rho\otimes|0\rangle\langle 0|_{E_2})V_2^\dagger],\end{aligned} \tag{2.76}$$

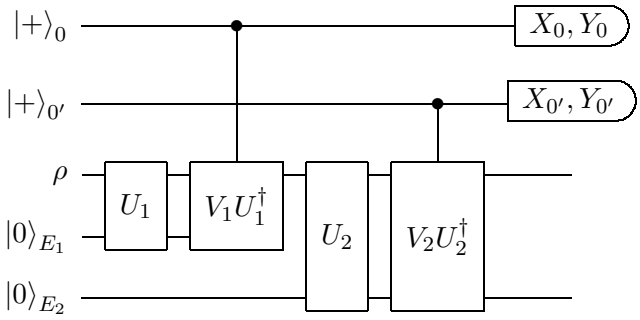
the concatenated operation

$$\Phi_2\circ\Phi_1(\rho) = \text{Tr}_{E_1E_2}[U_2U_1(\rho\otimes|0\rangle\langle 0|_{E_1,E_2})V_1^\dagger V_2^\dagger], \tag{2.77}$$

is also a generalised quantum operation. The real and imaginary part of $\Phi_2 \circ \Phi_1(\rho)$ could be obtained from measuring the ancillary qubit on the X_0 and Y_0 basis with the following circuit.



It can be equivalently realised as follows using two ancillary qubits.



In particular, the above circuit factorises into two independent circuits when (1) ρ is a tensor product of two states $\rho = \rho_1 \otimes \rho_2$ (2) U_1 and V_1 applies on ρ_1 and $|0\rangle_{E_1}$; U_2 and V_2 apply on ρ_2 and $|0\rangle_{E_2}$.

2.5.2 Pauli transfer matrix representation

In this section, I introduce the Pauli transfer matrix representation of states, observables, and channels as a preliminary. By using Pauli transfer representation, a state, and an observable are mapped to a real column and row vectors respectively, as follows

$$\begin{aligned} |\rho\rangle\rangle &= [\dots \rho_k \dots] \\ \rho_k &= \mathrm{Tr}(P_k \rho), \end{aligned} \tag{2.78}$$

and

$$\begin{aligned} \langle\langle Q| &= [\dots Q_k \dots] \\ Q_k &= \frac{1}{d}\mathrm{Tr}(Q P_k), \end{aligned} \tag{2.79}$$

where $P_k \in \{I, \sigma_x, \sigma_y, \sigma_z\}^{\otimes n}$, n is the number of qubits, and $d = 2^n$. Furthermore, for a process, i.e., $\mathcal{E}(\rho) = \sum_k K_k \rho K_k^\dagger$, the Pauli transfer matrix representation is

$$E_{k,j} = \frac{1}{d}\mathrm{Tr}\big(P_k \mathcal{E}(P_j)\big), \tag{2.80}$$

and the measurement in the Pauli transfer representation is

$$\mathrm{Tr}\big(Q\mathcal{E}(\rho)\big) = \langle\langle Q| E |\rho\rangle\rangle. \tag{2.81}$$

2.5.3 A complete basis operation set

Every single-qubit operation can be decomposed into a linear combination of 16 basis operations. This is because every single-qubit operation (including projective measurements) can be expressed with square matrices with $4 \times 4 = 16$ elements by using the Pauli transfer representation in Section 2.5.2. Therefore, 16 linearly independent operations are sufficient to emulate arbitrary single-qubit operations. Table 2.1 shows a set of single-qubit basis operations [149].

1	$[I]$ (no operation)	2	$[\sigma^{\mathrm{x}}]$
3	$[\sigma^{\mathrm{y}}]$	4	$[\sigma^{\mathrm{z}}]$
5	$[R_{\mathrm{x}}] = [\frac{1}{\sqrt{2}}(I + i\sigma^{\mathrm{x}})]$	6	$[R_{\mathrm{y}}] = [\frac{1}{\sqrt{2}}(I + i\sigma^{\mathrm{y}})]$
7	$[R_{\mathrm{z}}] = [\frac{1}{\sqrt{2}}(I + i\sigma^{\mathrm{z}})]$	8	$[R_{\mathrm{yz}}] = [\frac{1}{\sqrt{2}}(\sigma^{\mathrm{y}} + \sigma^{\mathrm{z}})]$
9	$[R_{\mathrm{zx}}] = [\frac{1}{\sqrt{2}}(\sigma^{\mathrm{z}} + \sigma^{\mathrm{x}})]$	10	$[R_{\mathrm{xy}}] = [\frac{1}{\sqrt{2}}(\sigma^{\mathrm{x}} + \sigma^{\mathrm{y}})]$
11	$[\pi_{\mathrm{x}}] = [\frac{1}{2}(I + \sigma^{\mathrm{x}})]$	12	$[\pi_{\mathrm{y}}] = [\frac{1}{2}(I + \sigma^{\mathrm{y}})]$
13	$[\pi_{\mathrm{z}}] = [\frac{1}{2}(I + \sigma^{\mathrm{z}})]$	14	$[\pi_{\mathrm{yz}}] = [\frac{1}{2}(\sigma^{\mathrm{y}} + i\sigma^{\mathrm{z}})]$
15	$[\pi_{\mathrm{zx}}] = [\frac{1}{2}(\sigma^{\mathrm{z}} + i\sigma^{\mathrm{x}})]$	16	$[\pi_{\mathrm{xy}}] = [\frac{1}{2}(\sigma^{\mathrm{x}} + i\sigma^{\mathrm{y}})]$

Table 2.1: Sixteen basis operations. These operations are composed of single-qubit rotations and measurements. $[I]$ denotes an identity operation (no operation), $[\sigma^i]$ $(i = x, y, x)$ corresponds to operations applying Pauli matrices. $[\pi]$ corresponds to projective measurements.

Here, the complete set of basis operations is denoted as $\{\mathcal{B}_i\}$. For multiple qubit systems, a tensor product of single-qubit operations, e.g., $\mathcal{B}_i \otimes \mathcal{B}_j$, also forms a complete basis set for composite systems. Therefore, we can decompose any n-qubit operation into the basis $\{\mathcal{B}_i\}^{\otimes n}$

$$\mathcal{E} = \sum_i \lambda_i \mathcal{B}_i. \tag{2.82}$$

Although the decomposition is universal, it may produce a large coefficient $\Lambda := \sum_i |\lambda_i|$, which results in a large sampling overhead. More details will be discussed in Chapter 6.

2.6 Quantum error mitigation

2.6.1 Error Model

Noise is inevitable in a quantum device due to interaction with the environment. Effective quantum error mitigation (QEM) schemes are crucial for suppressing errors in order to guarantee calculation accuracy. This section provides a review of QEM.

In a digital gate-based quantum computer, the effect of noise is simplified as a quantum channel appearing either before or after each gate. The output state is different from the ideal one, which can be described as

$$\begin{aligned} \rho_{\mathrm{out}}^{\mathrm{noisy}} &= \mathcal{N}_{N_g} \circ \mathcal{U}_{N_g} \circ \ldots \mathcal{N}_1 \circ \mathcal{U}_1(\rho_{\mathrm{in}}) \\ \rho_{\mathrm{out}}^{\mathrm{ideal}} &= \mathcal{U}_{N_g} \circ \cdots \circ \mathcal{U}_1(\rho_{\mathrm{in}}), \end{aligned} \tag{2.83}$$

where $\rho_{\text{out}}^{\text{noise}}$ is a noisy output and $\rho_{\text{out}}^{\text{ideal}}$ is a noise-free output from the quantum circuit, $\mathcal{U}_k$ and $\mathcal{N}_k$ are k^{th} quantum operation and the noise accompanying it, and N_g is the number of gates. Here, we assume the noise processes are Markovian for the sake of simplicity. Fault-tolerant error correction based on encoding of qubits can be used to compensate for the effect of noise and obtain correct computation results. However, in near-term quantum computing, the number of qubits and gate operations are restricted due to imperfections of quantum devices, which include physical noise and limited interactions among qubits. Therefore, fault-tolerant error correction necessitating encoding of qubits could be challenging for near-term quantum computing. Instead, QEM was introduced for mitigating errors in quantum circuits without using additional qubits. By using QEM, one cannot restore the quantum state itself, but can instead obtain an approximation of expectation values of observables corresponding to the ideal density matrix, i.e.,

$$\text{Tr}\left[\text{QEM}\left(\rho_{\text{out}}^{\text{noisy}}\right) O\right] \approx \text{Tr}\left[\rho_{\text{out}}^{\text{ideal}} O\right], \tag{2.84}$$

for any observable O. Here, QEM(ρ) denotes the process of error mitigation, which may not satisfy the requirements of a quantum channel. Therefore, we generally need classical post-processing to realise QEM(ρ), which may introduce a sampling overhead (cost) when measuring observables. The cost in general increases exponentially with respect to the error strength, as we will see below. Therefore, a constant small error strength is generally required in order to make QEM to work.

2.6.2 Quasi-probability method

Among different QEM schemes via different post-processing, the quasi-probability error mitigation method is one of the most effective approaches [97, 149, 150]. It recovers the ideal unitary processes by randomly generating noisy operations, with post-processing of measurement results. Suppose the ideal quantum operation is denoted as $\mathcal{U}$, then the key idea of the quasi-probability method is to express the ideal evolution $\mathcal{U}$ as a linear combination of noisy operations $\mathcal{K}_i$ as

$$\mathcal{U} \approx \sum_i q_i \mathcal{K}_i = C \sum_i p_i \text{sgn}(q_i) \mathcal{K}_i, \tag{2.85}$$

where $\mathcal{U}$ and $\mathcal{K}_i$ are superoperators, and $\sum_i q_i = 1$, $C = \sum_i |q_i|$, $p_i = |q_i|/C$. As q_i can be negative, we refer to q_i as the quasi-probability, and therefore the overhead coefficient $C \geq 1$ in general. To obtain the error free expectation value of an observable O, we randomly generate noisy operation $\mathcal{K}_i$ with probability p_i, multiply the measured result by the parity factor $\text{sgn}(q_i)$, and obtain the expectation value $\langle O \rangle_{\text{eff}}$ as follows,

$$\langle O \rangle_{\text{eff}} = \sum_i p_i \text{sgn}(q_i) \text{Tr}[O \mathcal{K}_i(\rho_{\text{in}})], \tag{2.86}$$

Finally, the error-free expectation value of $\langle O \rangle$ is approximated by $C \langle O \rangle_{\text{eff}}$. The variance is amplified C^2 times greater, and thus the number of measurements required to achieve the same

accuracy as that without QEM will be amplified C^2 times greater. C^2 can be interpreted as a resource cost for QEM.

By way of illustration, we shall take a case in which a single-qubit operation is affected by depolarising errors as $\mathcal{DU}$. The removal of the error $\mathcal{D}$ can be formally done by applying its inverse channel $\mathcal{D}^{-1}$. Now, the depolarising channel can be expressed as

$$\mathcal{D}(\rho) = \left(1 - \frac{3}{4}p\right)\rho + \frac{p}{4}(X\rho X + Y\rho Y + Z\rho Z), \tag{2.87}$$

with the inverse channel derived as

$$\mathcal{D}^{-1}(\rho) = C_{\mathcal{D}^{-1}}[p_1\rho - p_2(X\rho X + Y\rho Y + Z\rho Z)], \tag{2.88}$$

where $C_{\mathcal{D}^{-1}} = (p+2)/(2-2p) > 1$, $p_1 = (4-p)/(2p+4)$, and $p_2 = p/(2p+4)$. Here, it is assumed that $p < 1$.

Consequently, the ideal channel $\mathcal{U}$ can be expressed as

$$\begin{aligned} \mathcal{U} &= \mathcal{D}^{-1}\mathcal{DU} \\ &= C_{\mathcal{D}^{-1}}[p\mathcal{IDU} - p_2(\mathcal{XDU} + \mathcal{YDU} + \mathcal{ZDU})], \end{aligned} \tag{2.89}$$

where $\mathcal{I}$, $\mathcal{X}$, $\mathcal{Y}$ and $\mathcal{Z}$ correspond to an identity operation, and the superoperators for Pauli operators. Note that Eq. (2.89) is written in the same form as Eq. (2.85), and hence the quasi-probability method can be performed similarly.

To mitigate errors in a quantum circuit consisting of multiple gates, the quasi-probability operations are applied after each noisy gate. The parity is updated depending on the generated operations, and the final outcome of the parity is applied to measurement results in the same way as a single quantum operation, shown in Eq. (2.85). Suppose there are N gates in a circuit; the total overhead C_N can be expressed as

$$C_N = \prod_{i=1}^{N} C_i, \tag{2.90}$$

where C_i is the overhead for the ith gate. Assume the error ε_i for each gate is small, and thus the cost C_i is close to 1. We assume C_i takes the form of $C_i \approx 1 + \lambda_i\varepsilon_i$, which is the first-order expansion with respect to ε_i, where λ_i characterises the error strength. The total overhead C_N can be approximated by

$$C_N \approx \prod_i (1 + \lambda_i\varepsilon_i). \tag{2.91}$$

For simplicity, we assume $\lambda_i = \lambda$ and $\varepsilon_i = \varepsilon$ are independent of i. Then we have

$$C_N \approx (1+\lambda\varepsilon)^N = (1+\lambda\varepsilon)^{\frac{1}{\lambda\varepsilon}\lambda\varepsilon N} \approx e^{\lambda\varepsilon N} = e^{\lambda\varepsilon_N}. \tag{2.92}$$

Here $\varepsilon_N = \varepsilon N$ is denoted as the total error rate of all the N gates. The total cost C_N increases exponentially with the total error rate ε_N. Error mitigation methods are useful when the total error rate of a quantum circuit is small, $\varepsilon_N = \mathcal{O}(1)$.

2.6.3 Stochastic error mitigation

In the above discussion, the QEM method assumes that the noise appears either before or after each gate in a digital gate-based quantum computer, but realistic noise occurring in the experimental apparatus is more complicated. Specifically, every gate in digital circuits or every process in analogue simulation is physically realised via a continuous real-time evolution of a Hamiltonian, and thus errors can either inherently mix with the evolution making it strongly gate- or process-dependent, or act on multiple qubits leading to highly nonlocal correlated effects (crosstalk). Since conventional quantum error mitigation methods are restricted to gate-based digital quantum computers and over-simplified noise models, they fail to work when applied to realistic errors and general continuous quantum processes. In Chapter 10, I extend and apply the QEM method to a more practical scenario, demonstrating how errors can be mitigated for inherent dynamics-based and nonlocal noise in practical noisy quantum devices.

Part II

Estimating static and dynamic properties of quantum many-body systems

In Part I, I introduced a general recipe for describing and solving a quantum many-body problem. The first step is to encode the problem, which requires an effective and efficient description of a quantum many-body system. With this efficient representation in place, we prepare the initial state, evolve the state under the Hamiltonian, and measure the final state to get the observation of interest. In this process, the essential components are eigenstate preparation and dynamics simulation, which correspond to the static problems and dynamic problems introduced in Section 1.2. In this part, I will show a solution to such a general quantum many-body problem, covering specifically how to estimate static and dynamic properties of quantum systems.

In Chapter 3, I introduce a new quantum-classical architecture, the hybrid tensor network, which serves as a basis for a general representation of quantum systems and is applicable to a broad range of problems. This method incorporates the complementary strength of tensor networks and quantum computing, and is shown to be capable of unifying many existing classical and quantum algorithms.

In Chapter 4, I focus on eigenenergy estimation and eigenstate property estimation. A quantum algorithmic cooling scheme is proposed to find the energy spectrum and estimate eigenstate properties, which achieves exponential speed-up for the circuit complexity with the requirement of only one ancillary qubit, and shows clear advantages over quantum phase estimation and variational algorithms.

In Chapter 5 and Chapter 6, I focus on dynamic problems. The central objective is to approximate a unitary operator or a time-evolved state, formalised in Section 2.4. In Chapter 5, I propose an adaptive product formula for the latter problem. I further show how to enable large-scale quantum simulation by using the methods developed in Chapter 3. In Chapter 6, I introduce a perturbative approach for a generic interacting quantum system described by $H = H^{\mathrm{loc}} + V^{\mathrm{int}}$.

Chapter 3

Hybrid quantum-classical tensor networks: framework and application

The first step of quantum simulation is to develop a general framework for representing a quantum many-body system in an efficient way. There are several methods of describing a quantum system, one of which is tensor networks. Tensor network theory, originating from the density matrix renormalisation group (DMRG) for 1D Hamiltonians [151, 152], describes the quantum state with a network consisting of low-rank tensors. It has achieved notable success in its application to a variety of problems, particularly in 1D local gapped systems. However, the tensor network theory may become inadequate for application to general strongly correlated systems or those that do not admit an efficient tensor network description.

Drawing on tensor network theories and quantum computation, I introduce a new paradigm to represent quantum many-body states incorporating the power of quantum computing and classical tensor networks (overviewed in Section 2.1.4). A framework of the hybrid tensor network consisting of classical low-rank tensors and many-body quantum states is established. This method provides a basis for general hybrid quantum-classical representation of a many-body wave function that is applicable to a broad range of problems with concrete examples, such as quantum chemistry, quantum spin systems, searching for topological phase transitions, and quantum field theories.

By leveraging the ability of tensor networks in the efficient classical representation of quantum states, complex interacting quantum systems can be effectively represented using fewer quantum resources, which enables medium- or large-scale quantum simulation using small quantum processors. More importantly, this method demonstrates the ability to unify some typical algorithms in quantum simulation, which are commonly used in condensed matter physics and chemistry.

In this chapter, I will introduce the definition of classical and quantum tensors, the definition of tensor contraction and its meaning, the way to measure local observables, its properties, and applications in quantum many-body systems. This chapter is relevant to work published in

collaboration with Xiao Yuan, Junyu Liu, Qi Zhao, and You Zhou [90]. Author contributions are listed at the end of this chapter. The same format will be used in the following chapters.

3.1 Framework

3.1.1 Classical and quantum tensors

As introduced in Section 2.1.4, a hybrid tensor network consists of classical and quantum tensors, whose mathematical definition is consistent with that of a conventional tensor network. That is, tensor contractions are mathematically defined in the same way for classical and quantum tensors. Nevertheless, it is worth highlighting the difference between them. Operationally, classical tensors are contracted classically via tensor multiplication, while quantum tensors are contracted via measuring a quantum state on a quantum computer.

To motivate the establishment of hybrid tensor networks (TN), I begin with classical tensor network theory. A rank-n tensor, when regarded as a multi-dimensional array, can be represented as $T_{j_1,j_2,\ldots,j_n}$ with n indices. The amplitude of an n-partite quantum state in the computational basis corresponds to a rank-n tensor

$$|\psi\rangle = \sum_{j_1,j_2,\ldots,j_n} \psi_{j_1,j_2,\ldots,j_n} |j_1\rangle |j_2\rangle \ldots |j_n\rangle .$$

A classical TN typically consists of low-rank tensors, which could efficiently describe physical states that lie in a small subset of the whole Hilbert space.

A typical tensor network is a matrix product state (MPS) [153], which has achieved success in representing 1D local gapped systems with area law of entanglement. MPS admits the form of

$$|\psi\rangle = \sum_{j_1\cdots j_n} \mathrm{Tr}[A^{j_1} \ldots A^{j_n}] |j_1 \ldots j_n\rangle$$

which consists of rank-3 tensors with a small bond dimension κ of each matrix A^{j_k}, and compresses the state dimension from $O(2^n)$ to $O(n\kappa^2)$. As discussed in Section 2.1.4, tensor networks may not be able to represent general quantum many-body systems. This motivates the incorporation of quantum computing. A quantum state generated from applying a unitary circuit to a certain initial state generically forms an intrinsic large-rank quantum tensor and can be naturally stored and manipulated with a quantum computer.

In the literature, classical tensor network theory and quantum simulation are generally used as separate techniques in classical and quantum computing. Below I introduce quantum tensors as general n-partite quantum states prepared by a quantum computer, and classical tensors as low-rank tensors stored in a classical computer; I then demonstrate the combination of quantum and classical tensors as a hybrid tensor network.

Suppose an n-partite quantum state is generated by applying a unitary U_ψ to an initial state $|\bar{0}\rangle$ as $|\psi\rangle = U_\psi |\bar{0}\rangle$, as shown in Figure 3.1(a). As shown in Eq. (2.17), the quantum state can be

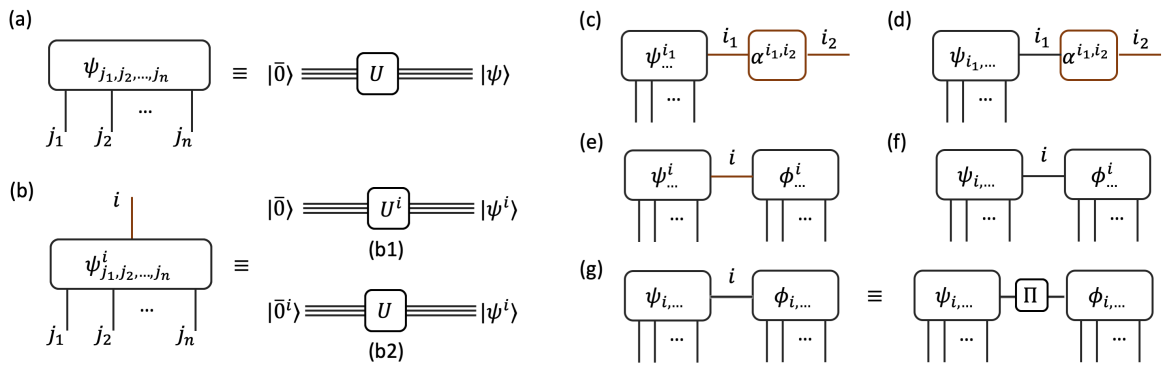


Figure 3.1: Tensor network representation of quantum states and tensor contractions. (a) A general n-partite quantum state can be regarded as a rank-n tensor (black line, subscript). (b) A classical index (red line, superscript) is added to an n-partite quantum state to generate a rank-$(n+1)$ tensor with n indices representing n quantum systems and 1 classical index. With a quantum circuit, it is equivalent to preparing different states $|\psi^i\rangle = U^i |i\rangle$ with (b1) different unitary operations as $|\psi^i\rangle = U^i |\bar{0}\rangle$ or (b2) simply the same unitary but different initial states as $|\psi^i\rangle = U |\bar{0}^i\rangle$. (c, d) Tensor contractions between a quantum tensor and a classical tensor. The contracted index could be (c) classical (red) or (d) quantum (black), which is contracted in different ways. (e, f, g) Tensor contractions between two quantum tensors. (c, e) The contracted index of both tensors corresponds to a classical index. (d, f) The contracted index corresponds to a classical index for one tensor and a quantum index for another tensor. (g) The contracted index of both tensors corresponds to a quantum index. The tensor Π is equivalent to a projective measurement $\sum_{i=i'} |i\rangle\langle i| \otimes |i'\rangle\langle i'|$.

regarded as a rank-n tensor in the computational basis. We can also introduce classical index to the quantum state by applying different unitary gates as

$$|\psi^i\rangle = U_{\psi^i} |\bar{0}\rangle = \sum_{j_1,j_2,\ldots,j_n} \psi^i_{j_1,j_2,\ldots,j_n} |j_1\rangle |j_2\rangle \ldots |j_n\rangle . \tag{3.1}$$

Alternatively, we can also apply the same unitary but to different initial states as

$$|\psi^i\rangle = U |\bar{0}^i\rangle = \sum_{j_1,j_2,\ldots,j_n} \psi^i_{j_1,j_2,\ldots,j_n} |j_1\rangle |j_2\rangle \ldots |j_n\rangle , \tag{3.2}$$

where the classical index i indicates the different unitaries or different initial states. As a result, $\{|\psi_i\rangle\}$ as a whole forms a rank-$(n+1)$ tensor. This is illustrated in Figure 3.1(b).

For simplicity, only one classical index is introduced here, but there is no restriction on introducing more classical indices. All these cases are regarded as quantum tensors, and a network connecting with quantum tensors and classical tensors is termed as a *hybrid tensor network*. Hereafter, we put indices corresponding to classical labels and quantum basis to the superscript and subscript of the tensor, respectively. This work focuses on qubits, and the results can be straightforwardly generalised to qudits in higher dimensions.

3.1.2 Hybrid tensor networks

In this section, I show how to connect quantum and classical tensors to form a hybrid tensor. When connecting two tensors, being either classical or quantum, we follow the conventional rule for tensor contraction.

Regarding low-rank tensors as classical tensors (superscript index, coloured red in figures) and quantum states as quantum tensors (subscript index, coloured black), we define hybrid TNs as networks constructed by connecting both classical and quantum tensors. For example, the tensor A^{i_1,i_2} represents a classical tensor with two classical indices and $\psi^{i}_{j_1,j_2,\ldots,j_n}$ represents a set of n-partite quantum states. Two tensors, being either classical or quantum, are connected by following the conventional contraction rule, i.e., $C^{i_1,i_3} = \sum_{i_2} A^{i_1,i_2} B^{i_2,i_3}$. As an example of this, the connection of a quantum and a classical tensor is shown in Figure 3.1(c, d).

While the mathematical definition of tensor contraction of a hybrid tensor network is consistent with the conventional definition, its practical meaning can be different. Specifically, classical and quantum tensors are contracted in two different ways via tensor contraction and quantum state measurement, respectively. Depending on whether the tensor and the index are quantum or classical, there are five different cases under contraction, as shown in Figure 3.1(c-g). For ease of explanation, this chapter considers the contraction of rank-2 classical tensors and rank-$(n+1)$ quantum tensors without loss of generality. Hereafter, "$\ldots$" is used to abbreviate the quantum indices (e.g., j_1, j_2, ..., j_n in Figure 3.1(a)) when they are not contracted.

Case 1: quantum tensor (contracted index: classical) & classical tensor (contracted index: classical) in Figure 3.1(c).

First, a quantum tensor $\psi^{i_1}_{\ldots}$ is connected with a classical tensor α^{i_1,i_2} to form a new rank-$(n+1)$ tensor,

$$\tilde{\psi}^{i_2}_{\ldots} = \sum_{i_1} \psi^{i_1}_{\ldots} \cdot \alpha^{i_1,i_2}, \tag{3.3}$$

where the contracted index from the quantum and classical tensors is a classical label. To understand the meaning of Eq. (3.3), the quantum tensor $\psi^{i_1}_{\ldots}$ is regarded as a set of independent quantum states $\{|\psi^{i_1}\rangle\}$ and the tensor $\tilde{\psi}^{i_2}_{\ldots}$ represents a new set of states,

$$|\tilde{\psi}^{i_2}\rangle = \sum_{i_1} \alpha^{i_1,i_2} |\psi^{i_1}\rangle, \tag{3.4}$$

where each one is now a superposition of the original states $\{|\psi^{i_1}\rangle\}$. As a special case, when the classical tensor is rank-1, α^{i_1}, the output tensor is

$$|\tilde{\psi}\rangle = \sum_{i_1} \alpha^{i_1} |\psi^{i_1}\rangle. \tag{3.5}$$

Therefore, we can connect a classical tensor to the classical index of a quantum tensor to effectively represent a superposition of quantum states.

Case 2: quantum tensor (contracted index: quantum) & classical tensor (contracted index: classical) in Figure 3.1(d).

When the contracted index i_1 of the quantum tensor $\psi_{i_1,\dots}$ corresponds to a quantum system, the tensor contraction is similarly defined as

$$\tilde{\psi}^{i_2}_{\dots} = \sum_{i_1} \psi_{i_1,\dots} \cdot \alpha^{i_1,i_2}. \tag{3.6}$$

When considering quantum states, the contraction transforms an input state $|\psi\rangle$ to a set of output states $\{|\tilde{\psi}^{i_2}\rangle\}$ as

$$|\tilde{\psi}^{i_2}\rangle = \sum_{i_1} \alpha^{i_1,i_2} \langle i_1|\psi\rangle \,, \tag{3.7}$$

which is equivalent to projecting the contracted system onto $\langle i_1|$ to form a set of un-normalised states $|\psi^{i_1}\rangle = \langle i_1|\psi\rangle$ and re-combining them with coefficients α^{i_1,i_2}. Actually, if we regard α as a unitary gate with i_2 representing a quantum system, it corresponds to a local unitary transformation of the state.

Case 3: quantum tensor (contracted index: classical) & quantum tensor (contracted index: classical) in Figure 3.1(e).

Next, consider the contraction of two quantum tensors, with the contracted index being classical for both tensors. Suppose the two quantum tensors are $\psi^i_{\dots}$ and $\phi^i_{\dots}$, the contraction of index i gives

$$\tilde{\psi}_{\dots} = \sum_i \psi^i_{\dots} \cdot \phi^i_{\dots}. \tag{3.8}$$

Considering quantum states, the contraction transforms two sets of states $\{|\psi^i\rangle\}$ and $\{|\phi^i\rangle\}$ to an un-normalised state

$$|\tilde{\psi}\rangle = \sum_i |\psi^i\rangle \otimes |\phi^i\rangle \,. \tag{3.9}$$

By contracting two quantum tensors, we can thus effectively entangle two quantum systems. A classical tensor can also be added in between so that the amplitude for each $|\psi^i\rangle \otimes |\phi^i\rangle$ is different.

Case 4: quantum tensor (contracted index: quantum) & quantum tensor (contracted index: classical) in Figure 3.1(f).

When one of the contracted indices corresponds to a quantum system, the contraction is similarly defined as

$$\tilde{\psi}_{\dots} = \sum_i \psi_{i,\dots} \cdot \phi^i_{\dots}. \tag{3.10}$$

Considering quantum states, the contraction converts $|\psi\rangle$ and $\{|\phi^i\rangle\}$ to

$$|\tilde{\psi}\rangle = \sum_i \langle i|\psi\rangle \otimes |\phi^i\rangle \,. \tag{3.11}$$

Again, this is equivalent to applying a projection to $|\psi\rangle$ to get a set of states $\{|\psi^i\rangle = \langle i|\psi\rangle\}$ and then connecting the classical indices of the two quantum tensors.

Case 5: quantum tensor (contracted index: quantum) & quantum tensor (contracted index: quantum) in Figure 3.1(g).

When both contracted indices represent quantum systems, two quantum tensors $\psi_{i,\ldots}$ and $\phi_{i,\ldots}$ are contracted as

$$\tilde{\psi}_{\ldots} = \sum_i \psi_{i,\ldots} \cdot \phi_{i,\ldots}. \tag{3.12}$$

It is equivalent to

$$|\tilde{\phi}\rangle = \sum_i \langle i|\psi\rangle \otimes \langle i|\phi\rangle = \sum_i \langle i,i|\,|\psi\rangle \otimes |\phi\rangle\,, \tag{3.13}$$

which accounts for a Bell state projection on the contracted systems. Note that since it is a joint measurement on both states, the success probability could be less than 1. If there are multiple contractions of quantum indices, the overall probability could be exponentially small. Therefore, we only allow a constant number of contractions of two quantum indices in the hybrid tensor network.

For a general hybrid tensor network consisting of classical and quantum tensors, its tensor contraction rule and operational meaning are similar to the above cases.

3.1.3 Calculation of expectation values of local observables

In this section, I show how to measure the expectation values of tensor products of local observables given a hybrid tensor network representation of a quantum state. The basic rule follows the same mathematics of tensor contraction. For classical tensors, the expectation value is calculated in the same way as conventional tensor networks. Conversely, for quantum tensors we can no longer calculate the expectation value via tensor contraction since it involves matrix multiplication of a rank-n tensor. Instead, the expectation value is obtained by preparing the state on a quantum computer and measuring the state.

Figure 3.2(a) shows how to get an expectation value of local observables for the quantum systems described by a rank-$(n+1)$ quantum tensor. This tensor can be either an n-partite quantum state with a classical index i or an $(n+1)$-partite quantum state with a quantum basis index i. By measuring the n parties, it gives a new rank 2 tensor $M^{i',i}$ with two open indices i and i',

$$M^{i',i} = \langle \psi^{i'}|\, O_1 \otimes O_2 \otimes \cdots \otimes O_n\, |\psi^i\rangle\,, \tag{3.14}$$

where O_j is a local observable acting on the jth bit. Here the definition is the same if we measure an $(n+1)$-partite quantum state.

The indices of $M^{i',i}$ are always put in the superscript, because the observable in Eq. (3.14) is always a classical low-rank tensor. Note that the matrix $M^{i',i}$ is always hermitian so that it can be measured when the indices i and i' are contracted to another quantum tensor. I now illustrate how to get $M^{i',i}$ under different circumstances.

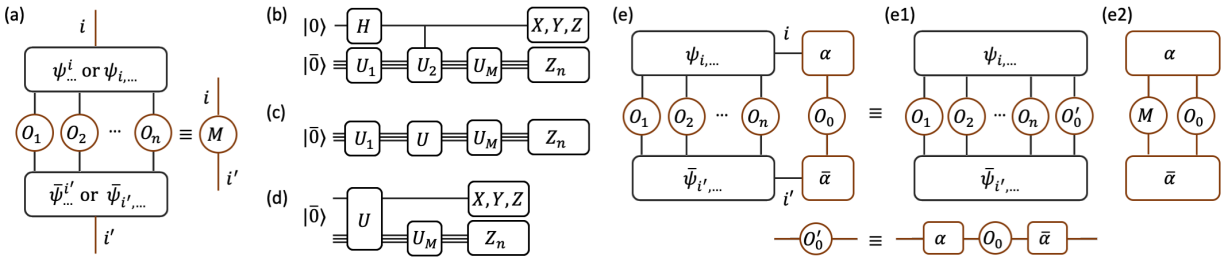


Figure 3.2: Measuring expectation values of a quantum tensor. (a) Consider a rank-$(n+1)$ quantum tensor, which could be either an n-partite quantum state with a classical index i or an $(n+1)$-partite quantum state with a quantum basis index i. The expectation value of the n quantum systems gives a hermitian observable $M^{i',i} = \langle\psi^{i'}|\, O_1 \otimes O_2 \otimes \cdots \otimes O_n \,|\psi^i\rangle$ on the open indices. Each element $M^{i',i}$ can be measured with a quantum circuit of (b), (c), or (d). (b) Suppose the index i is classical and $|\psi^i\rangle = U^i\,|\bar{0}\rangle$, we choose $U_1 = U^i$, $U_2 = U^{i'}(U^i)^\dagger$, and U_M to be the unitary that rotates the eigenstates of the observables to the computational basis. Each $M^{i',i}$ is obtained by measuring the ancillary qubit in the X, Y, Z bases and the other n qubits in the computational Z basis. (c) Suppose the index i is classical and $|\psi^i\rangle = U\,|\bar{0}^i\rangle$; U_1 is applied to prepare four input states $|\bar{0}^i\rangle, |\bar{0}^{i'}\rangle, (|\bar{0}^i\rangle + |\bar{0}^{i'}\rangle)/\sqrt{2}, (|\bar{0}^i\rangle + i\,|\bar{0}^{i'}\rangle)/\sqrt{2}$. Each $M^{i',i}$ corresponds to a linear combination of the measurement results. (d) Suppose the index i is quantum, after applying the unitary U for preparing the state $|\psi\rangle = U\,|\bar{0}\rangle$, we measure n qubits in the computational basis and the qubit with index i in the Pauli X, Y, and Z bases. (e) Tensor contraction can have different orders. With a rank-$(n+1)$ quantum tensor connected to a classical tensor, we can either (e1) first calculate the expectation value of the classical tensor and then measure the $(n+1)$-partite quantum state or (e2) first measure the n systems via (d) and then do classical tensor contraction.

Case 1: The rank-$(n+1)$ quantum tensor is an n-partite quantum state with a classical index i.

- Suppose $|\psi^i\rangle = U^i\,|\bar{0}\rangle$, $U_1 = U^i$. $M^{i',i}$ is measured with the quantum circuit in Figure 3.2(b). Consider $U_2 = U^{i'}(U^i)^\dagger$ and U_M to be the unitary that rotates the eigenstates of the observable to the computational basis. The output state before the U_M gate is

$$|\tilde{\psi}\rangle = \frac{1}{\sqrt{2}}\big(\,|0\rangle\,|\psi^i\rangle + |1\rangle\,|\psi^{i'}\rangle\,\big). \tag{3.15}$$

When the ancillary qubit measures the Pauli X, Y, Z operators, and the n-partite system measures $M = O_1 \otimes O_2 \otimes \cdots \otimes O_n$, the expectation values are

$$\begin{aligned}
\langle\tilde{\psi}|X \otimes M|\tilde{\psi}\rangle &= \frac{1}{2}\big(M^{2,1} + M^{1,2}\big),\\
\langle\tilde{\psi}|Y \otimes M|\tilde{\psi}\rangle &= \frac{1}{2}\big(iM^{2,1} - iM^{1,2}\big),\\
\langle\tilde{\psi}|Z \otimes M|\tilde{\psi}\rangle &= \frac{1}{2}\big(M^{1,1} - M^{2,2}\big).
\end{aligned} \tag{3.16}$$

Note that $M^{1,2}$ is the complex conjugate of $M^{2,1}$, and we have

$$\langle\tilde{\psi}|I \otimes M|\tilde{\psi}\rangle = \frac{1}{2}\big(M^{1,1} + M^{2,2}\big), \tag{3.17}$$

which can be obtained from the measurement of any Pauli basis. Therefore, we can exactly solve for each term $M^{i,j}$ $(i,j=1,2)$ and construct the measurement

$$\tilde{M} = \begin{Bmatrix} M^{1,1} & M^{1,2} \\ M^{2,1} & M^{2,2} \end{Bmatrix}.$$

- Suppose $|\psi^i\rangle = U|\bar{0}^i\rangle$. $M^{i',i}$ can be measured with the quantum circuit in Figure 3.2(c). Now we need to input $(|\bar{0}^i\rangle \pm |\bar{0}^{i'}\rangle)/\sqrt{2}$ and $(|\bar{0}^i\rangle \pm i|\bar{0}^{i'}\rangle)/\sqrt{2}$ and the matrix elements can be similarly obtained.

Case 2: The rank-$(n+1)$ quantum tensor is an $n+1$-partite quantum state with a quantum index I.

We need to measure

$$M^{i',i} = \langle\psi| |i'\rangle \otimes O_1 \otimes O_2 \otimes \cdots \otimes O_n \otimes \langle i| |\psi\rangle = \langle\psi| (|i'\rangle\langle i|) \otimes O_1 \otimes O_2 \otimes \cdots \otimes O_n |\psi\rangle. \quad (3.18)$$

Note that the matrix $|i'\rangle\langle i|$ can always be represented as a linear combination of the Pauli operators, we can thus instead measure the uncontracted qubit in the three X, Y, Z Pauli bases to equivalently get any $M^{i',i}$ as shown in Figure 3.2(d). Suppose $|\psi\rangle = U|\bar{0}\rangle$, denote

$$E(\sigma) = \langle\psi| \sigma \otimes O_1 \otimes O_2 \otimes \cdots \otimes O_n |\psi\rangle, \quad (3.19)$$

and we can reconstruct the measurement $\tilde{M}$ as

$$\tilde{M} = \frac{1}{2}\big(E(I)I + E(X)X - E(Y)Y + E(Z)Z\big), \quad (3.20)$$

where $E(X)$, $E(I)$, $E(Y)$, $E(Z)$ are the obtained expectation values with Pauli measurements I, X, Y, Z.

Calculating the expectation value of a general hybrid tensor network follows the above basic rules for classical and quantum tensors. Nevertheless, similar to conventional tensor networks, different orders of tensor contraction could have different procedures and complexities. For example, say that we are considering the hybrid tensor shown in Figure 3.2(e), which consists of a rank-$(n+1)$ quantum tensor and a classical tensor. We could first contract the right classical observable O_0 with the classical tensor α, and obtain a new observable $O_0^{'}$. Then we measure the $n+1$-partite quantum state to get the final expectation value. Here we need classical contraction and a single local measurement with repetition samples $\mathcal{M}$. This procedure is shown in Figure 3.2(e1). Alternatively, we can also use the circuit in Figure 3.2(d) to reconstruct observable M by measuring the $n+1$-partite quantum state and then contract the classical tensors. This procedure requires three local measurement settings (X, Y, and Z on the first qubit) with total repetition samples of $3\mathcal{M}$.

3.1.4 Application in quantum simulation

The hybrid tensor network provides a way to more efficiently represent quantum states with fewer quantum resources. The hybrid tensor network can be applied in variational quantum simulation for solving static energy spectra and simulating dynamics of a quantum system.

Consider a many-body Hamiltonian

$$H = \sum_i \lambda_i h_i \tag{3.21}$$

with coefficients λ_i and tensor products of Pauli matrices h_i. To find the ground state of the Hamiltonian, consider a parameterised hybrid quantum tensor network, which corresponds to a possibly un-normalised state $|\psi(\vec{x})\rangle$. Here $\vec{x}$ denotes the parameters in a hybrid tensor network, which include the parameterised rotation angles in the quantum circuit and parameters in the classical tensors. Then we can measure the average energy as

$$E(\vec{x}) = \frac{\langle\psi(\vec{x})|\, H\, |\psi(\vec{x})\rangle}{\langle\psi(\vec{x})|\psi(\vec{x})\rangle} = \frac{\sum_i \lambda_i \langle\psi(\vec{x})|\, h_i\, |\psi(\vec{x})\rangle}{\langle\psi(\vec{x})|\psi(\vec{x})\rangle}, \tag{3.22}$$

where each $\langle\psi(\vec{x})|\, h_i\, |\psi(\vec{x})\rangle$ and the normalisation $\langle\psi(\vec{x})|\psi(\vec{x})\rangle$ can be obtained by calculating the expectation value of the hybrid tensor network with the method I discussed in the previous section. Having measured $E(\vec{x})$ for any $\vec{x}$, we can then optimise the parameters via the classical algorithm to minimise $E(\vec{x})$ to search for the ground state. The whole optimisation procedure is identical to a conventional approach named variational quantum eigensolver (VQE) which is introduced in Section 2.2.2. The difference lies in the usage of the hybrid tensor network, which may enable quantum simulation of large systems with small quantum processors. We can also use hybrid tensor networks for simulating Hamiltonian dynamics. The circuit for the implementation of variational Hamiltonian simulation with hybrid tensor networks is slightly more complicated. I leave the discussion to Chapter 5. In the following sections, we mainly focus on using hybrid tensor networks for finding the ground state of a Hamiltonian.

3.2 Hybrid tree tensor networks

Calculating a general hybrid tensor network can be costly. Here we expand the discussion of the main text and focus on hybrid tensor networks with a tree structure. I first consider representative examples of hybrid tree tensor networks (TTN) and discuss its application in representing correlations of the multipartite quantum state. I then study the cost of calculating the expectation values of a general hybrid TTN.

I show several examples of hybrid tree tensor networks in Figure 3.3. Below, I discuss the application of each tensor network and its connection with existing results. For each n-partite state, we assume it is an n-qubit parameterised state $|\psi(\vec{\theta})\rangle$, obtained by applying a sequence of local gates as $|\psi(\vec{\theta})\rangle = \prod_j U_j(\theta_j)\,|\bar{0}\rangle$ with an initial state $|\bar{0}\rangle$ and parameters $\vec{\theta} = \{\theta_j\}$.

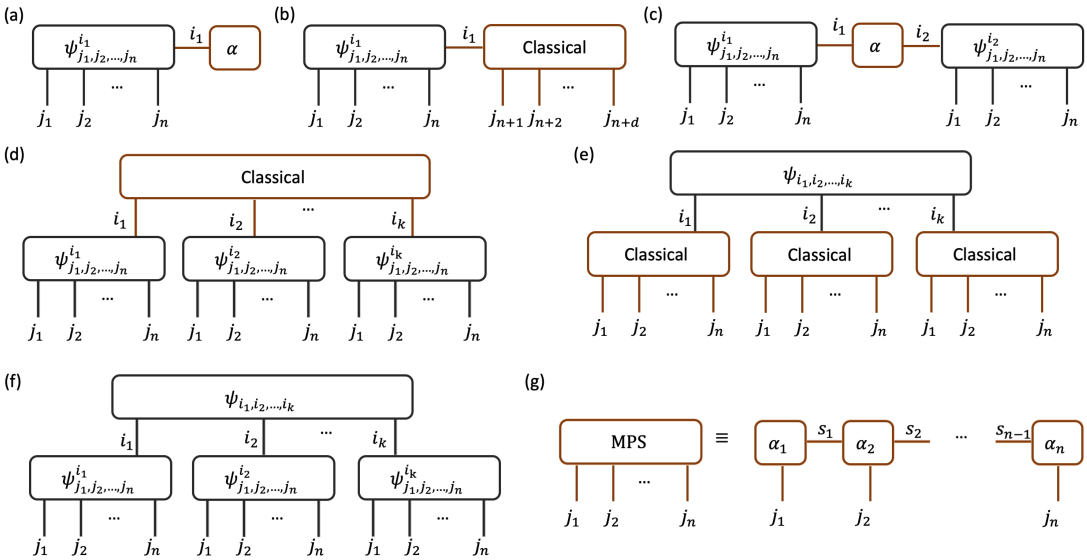


Figure 3.3: Hybrid quantum-classical tensor network. (a) We can extend the power of a quantum state by adding a classical tensor as in Eq. (3.23). (b) We can combine a quantum state and a classical tensor to represent a state in a larger Hilbert space as in Eq. (3.29). (c) We can use a classical tensor to connect two quantum states as in Eq. (3.32). (d) A quantum-classical hybrid tensor in Eq. (3.34). (e) A classical-quantum hybrid tensor in Eq. (3.37). (f) A quantum-quantum hybrid tensor in Eq. (3.40). (g) A commonly used classical tensor network MPS.

3.2.1 Extending the power of the quantum state

Suppose we use the quantum state $|\psi(\vec{\theta})\rangle$ as a potential solution to an n-qubit problem. We can regard the quantum state as a pure rank-n quantum tensor. A simple way to extend the capability of an rank-n quantum tensor is to concatenate a classical tensor α^i to it,

$$|\psi(\vec{x})\rangle = \sum_i \alpha^i \, |\psi^i(\vec{\theta}_i)\rangle \,, \tag{3.23}$$

where each $|\psi^i(\vec{\theta}_i)\rangle$ can be regarded as different rank-n quantum tensors and $\vec{x} = (\alpha^1, \dots, \vec{\theta}_1, \dots)$ are the total parameter setting. Such a concatenation corresponds to the hybrid tensor network in Figure 3.3(a). To find the ground state of Hamiltonian H, the energy can be obtained as

$$E(\vec{x}) = \frac{\langle\psi(\vec{x})|\, H \,|\psi(\vec{x})\rangle}{\langle\psi(\vec{x})|\psi(\vec{x})\rangle} = \frac{\sum_{i,j} \bar{\alpha}^i \alpha^j \, \langle\psi^i(\vec{\theta}_i)|\, H \,|\psi^j(\vec{\theta}_j)\rangle}{\sum_{i,j} \bar{\alpha}^i \alpha^j \, \langle\psi^i(\vec{\theta}_i)|\psi^j(\vec{\theta}_j)\rangle}, \tag{3.24}$$

and a minimisation over the parameter space could lead to the solution.

We can see that such a hybrid tensor network contains the subspace expansion method as a special case. In particular, suppose we fix the parameters of the quantum tensors $|\psi^i(\vec{\theta}_i)\rangle$ and denote $|\psi^i(\vec{\theta}_i)\rangle = |\psi^i(\vec{\theta})\rangle$, then we can analytically solve the minimisation of $E(\vec{x})$ as follows. Denote

$$H^{i,j} = \langle\psi^i(\vec{\theta})|\, H \,|\psi^j(\vec{\theta})\rangle \,,\; S^{i,j} = \langle\psi^i(\vec{\theta})|\psi^j(\vec{\theta})\rangle \,. \tag{3.25}$$

Suppose we consider the subspace with $\langle\psi(\vec{x})|\psi(\vec{x})\rangle = 1$, then it is equivalent to optimising $E(\vec{x}) = \langle\psi(\vec{x})| H |\psi(\vec{x})\rangle$, or the function $E'(\vec{x}) = E(\vec{x}) - \lambda \langle\psi(\vec{x})|\psi(\vec{x})\rangle$ with a Lagrangian multiplier λ. Variation of the new function $E'(\vec{x})$ gives

$$\delta E'(\vec{x}) = \sum_{i,j}(\alpha^j \delta\bar{\alpha}^i + \bar{\alpha}^i \delta\alpha^j)H^{i,j} - \lambda \sum_{i,j}(\alpha^j \delta\bar{\alpha}^i + \bar{\alpha}^i \delta\alpha^j)S^{i,j}, \tag{3.26}$$

and a local minimum solution requires $\delta E'(\vec{x}) = 0$, which is equivalent to

$$H^{i,j}\alpha^j = \lambda S^{i,j}\alpha^j. \tag{3.27}$$

Writing the equation in the matrix form, it is equivalent to

$$H\vec{\alpha} = \lambda S\vec{\alpha}, \tag{3.28}$$

which coincides with the subspace expansion method.

In practice, we can optimise all the parameters in both quantum and classical tensors. We can simultaneously optimise them by treating $E(\vec{x})$ as a black box cost function. Alternatively, we can first optimise the parameters of the quantum tensor and fix them, and then optimise parameters of the classical tensor. Since the parameters are not simultaneously optimised, we may need to repeat the procedure several times until convergence of energy.

3.2.2 Virtual qubits via classical tensors

In addition to extending the power of the quantum circuits, we can also use the classical tensor to represent physical quantum systems, similar to a classical tensor network. As shown in Figure 3.3(b), one can connect a rank-$(n+1)$ quantum tensor to a rank $d+1$ classical tensor network to represent a system of $n+d$ qubits. Here we assume the classical tensor network consists of low-rank classical tensors and admits efficient contraction, such as the matrix product state (MPS) as defined in Eq. (2.18). In the remainder of the section, I consider MPS as an example of the classical tensor network.

Suppose the rank-$(n+1)$ quantum tensor $\psi^{i_1}_{\cdots}$ represents a set of n-qubit quantum states $\{|\psi^{i_1}\rangle\}$ and the classical tensor is given by $\alpha^{i_1,j_{n+1},\ldots j_{n+d}} = \mathrm{Tr}[\alpha_1^{i_1,j_{n+1}}\alpha_2^{j_{n+2}}\ldots\alpha_n^{j_{n+d}}]$, then the hybrid tensor of Figure 3.3(b) represents a quantum state

$$|\tilde{\psi}\rangle = \sum_{i_1,j_1,\ldots,j_{n+d}} \alpha^{i_1,j_{n+1},\ldots j_{n+d}} |\psi^{i_1}\rangle |j_{n+1}\rangle \ldots |j_{n+d}\rangle . \tag{3.29}$$

For any tensor products of local observables $M = O_1 \otimes \cdots \otimes O_{n+d}$, we have

$$\langle\tilde{\psi}|M|\tilde{\psi}\rangle = \sum_{i_1,i_1'} \langle\psi^{i_1'}|O_1 \otimes \ldots O_n|\psi^{i_1}\rangle M^{i_1',i_1}, \tag{3.30}$$

with

$$M^{i'_1,i_1} = \sum_{j'_{n+1},\dots j'_{n+d},j_{n+1},\dots j_{n+d}} \bar{\alpha}^{i'_1,j'_{n+1},\dots j'_{n+d}} \alpha^{i_1,j_{n+1},\dots j_{n+d}} \langle j'_{n+1}|O_{n+1}|j_{n+1}\rangle \dots \langle j'_{n+d}|O_{n+d}|j_{n+d}\rangle \,. \tag{3.31}$$

Here each $\langle \psi^{i'_1}|O_1 \otimes \dots O_n|\psi^{i_1}\rangle$ is obtained with a quantum computer and each element $M^{i'_1,i_1}$ is obtained by an efficient tensor contraction of the MPS ansatz. Note that the dimension of i_1 can be chosen to be a small number, similar to how we decide the bond dimension of the MPS ansatz. The definition also holds when the quantum tensor is an $(n+1)$-partite state, where we can assign multiple qubits to the system that the i_1 label represents.

3.2.3 Local quantum correlation and non-local classical correlation

Quantum tensors can be used to represent quantum correlations of local subsystems and classical tensors to represent correlations between the subsystems[1]. For example, consider two subsystems A and B with Hamiltonian $H = H_A + H_B + \lambda h_A \otimes h_B$ with a small coupling constant λ. We can use the hybrid tensor network in Figure 3.3(c) to represent its ground state,

$$|\tilde{\psi}\rangle_{AB} = \sum_{i_1,i_2} \alpha^{i_1,i_2} |\psi^{i_1}\rangle_A \otimes |\psi^{i_2}\rangle_B \,. \tag{3.32}$$

Here $|\psi^{i_1}\rangle_A$ and $|\psi^{i_2}\rangle_B$ represent the state of subsystem A and B, respectively, and $\alpha^{i,j}$ is the classical tensor representing the correlation between A and B. If the quantum correlation is not too strong, we can set the rank of $\alpha_{i,j}$ to be a small number. The average energy of the Hamiltonian is

$$E = \frac{\langle\tilde{\psi}|H|\tilde{\psi}\rangle_{AB}}{\langle\tilde{\psi}|\tilde{\psi}\rangle_{AB}} = \frac{\sum_{i_1,i_2,i'_1,i'_2} \bar{\alpha}^{i'_1,i'_2}\alpha^{i_1,i_2}\left(H_A^{i'_1,i_1}S_B^{i'_2,i_2} + S_A^{i'_1,i_1}H_B^{i'_2,i_2} + \lambda h_A^{i'_1,i_1}h_B^{i'_2,i_2}\right)}{\sum_{i_1,i_2,i'_1,i'_2} \bar{\alpha}^{i'_1,i'_2}\alpha^{i_1,i_2}S_A^{i'_1,i_1}S_B^{i'_2,i_2}}, \tag{3.33}$$

where the matrices $H_A^{i'_1,i_1}$, $S_B^{i'_2,i_2}$, $S_A^{i'_1,i_1}$, $H_B^{i'_2,i_2}$, $h_A^{i'_1,i_1}$, $h_B^{i'_2,i_2}$ are defined in a general way as in Eq. (3.25), that is, $M^{i,j}_{A(B)} = \langle\psi^i|M|\psi^j\rangle_{A(B)}$. We can then obtain the energy by measuring the matrices with a quantum computer and contracting the classical tensors classically. Suppose that each system A and B consists of n qubits, so that the total system size is $2n$ qubits. The energy can be obtained by only manipulating states of n qubits instead of $2n$ qubits.

In a similar way, we can extend the hybrid tensor network for two subsystems to k subsystems, as shown in Figure 3.3(d). The matrix product state $\alpha^{i_1,i_2\dots i_k} = \mathrm{Tr}[\alpha_1^{i_1}\alpha_2^{i_2}\dots\alpha_k^{i_k}]$ is used as the description of the correlation between subsystems. Suppose each subsystem is represented by quantum states $\{|\psi^{i_s}\rangle_s\}$, such that the hybrid tensor network of Figure 3.3(d) represents a quantum state

$$|\tilde{\psi}\rangle = \sum_{i_1,i_2,\dots i_k} \alpha^{i_1,i_2,\dots i_k} |\psi^{i_1}\rangle_1 \otimes |\psi^{i_2}\rangle_2 \dots |\psi^{i_k}\rangle_k \,. \tag{3.34}$$

[1] In this context, 'local' concerns correlations in each subsystem, whilst 'non-local' concerns correlations between different subsystems.

To measure the expectation value of $M = O_1 \otimes \cdots \otimes O_k$ with each O_s representing a local observable on the s-th subsystem, we have

$$\langle\tilde{\psi}|M|\tilde{\psi}\rangle = \sum_{i_1\ldots i_k, i'_1,\ldots i'_k} \bar{\alpha}^{i'_1,\ldots i'_k} \alpha^{i_1,\ldots i_k} M_1^{i'_1,i_1} \ldots M_k^{i'_k,i_k}, \tag{3.35}$$

with

$$M_s^{i'_s,i_s} = \langle\psi^{i'_s}|O_s|\psi^{i_s}\rangle_s \,. \tag{3.36}$$

As a result, we can just use an n-qubit system to represent a kn-qubit system, and the bipartite version corresponds to $k = 2$. The dimension of each index $i_1, \ldots i_k$ should be a small number, similar to the bond dimension of MPS. This is a general form of the hybrid quantum-MPS tensor network, and one can also consider other classical tensor networks (see Appendix A). Indices involved here in the contraction between the quantum and classical tensors are all classical ones. Alternatively, one can also make a hybrid contraction, where the index of the quantum tensor is a quantum one, as shown in Eq. (3.6).

In this instance, the quantum tensors are used to represent the local n-qubit correlation, and the classical rank-k tensor is used to represent the global correlation among these k clusters of qubits. This kind of quantum-classical tensor can be used as a heuristic ansatz for representing systems where local correlation dominates over global correlation, such as weakly coupled cluster systems. However, it is worth noting at this point that it may still be inadequate for representing a general quantum system.

3.2.4 Local classical correlation and non-local quantum correlation

Instead of using a quantum processor to represent local correlations within each subsystem, one can also consider a classical-quantum two-depth tree structure in Figure 3.3(e). In this structure, classical tensors are used for representing local correlations of each subsystem, and a quantum tensor is used for representing the non-local correlation between the subsystems.

The idea underpinning this approach is that after the quantum circuit is applied to prepare a k-qubit state $|\psi\rangle$, we additionally connect a classical tensor network to each qubit to transform it to n qubits. Suppose we use the MPS for representing each subsystem as $\alpha^{i_s,j_1^s\ldots j_n^s} = \mathrm{Tr}[\alpha_1^{i_s,j_1^s} \ldots \alpha_n^{j_n^s}]$, the state corresponding to Figure 3.3(e) is

$$|\tilde{\psi}\rangle = \sum_{\vec{i},\vec{j}^1,\ldots\vec{j}^k} \alpha^{i_1,\vec{j}^1} \ldots \alpha^{i_k,\vec{j}^k} \psi_{i_1,\ldots,i_k} |\vec{j}^1\rangle \otimes \cdots \otimes |\vec{j}^k\rangle \,, \tag{3.37}$$

where we denote $\vec{i} = (i_1, \ldots i_k)$, $\vec{j}^s = (j_1^s, \ldots, j_n^s)$, and $\psi_{i_1,\ldots,i_k} = \langle\vec{i}|\psi\rangle$. To measure $M = O_1 \otimes \cdots \otimes O_k$ with each O_s representing tensor products of local observables, we have

$$\langle\tilde{\psi}|M|\tilde{\psi}\rangle = \langle\psi|\tilde{O}_1 \otimes \cdots \otimes \tilde{O}_k|\psi\rangle \,, \tag{3.38}$$

with each observable $\tilde{O}_s$ obtained by classical tensor contraction as

$$\tilde{O}_s^{i'_s,i_s} = \sum_{\vec{j}^{s'},\vec{j}^s} \tilde{\alpha}^{i'_s,\vec{j}^{s'}} \alpha^{i_s,\vec{j}^s} O_s^{\vec{j}^{s'},\vec{j}^s}. \tag{3.39}$$

Again, only k qubits are used to represent a system of nk qubits. Note that each subsystem may have a different number of qubits, and we can use multiple qubits to represent each index i_s to increase the bond dimension. When $n \gg 1$, this kind of hybrid tensor network can be used as a heuristic ansatz to represent long-range correlation due to the effect of quantum tensor, and it may be applied to an exotic topological state. When n is a small number, it also represents a normalisation of local correlations with the classical tensor.

3.2.5 Local and non-local quantum correlations

In the previous two cases, a classical tensor network is used to represent either local or non-local correlation, and a quantum tensor is used to represent the other part. I will now demonstrate how quantum tensors can be used to represent both the local and non-local correlations. Considering a tree structure with depth 2 in Figure 3.3(f), it represents a state

$$|\tilde{\psi}\rangle = \sum_{i_1,\ldots,i_k} \psi_{i_1,\ldots,i_k} |\psi_1^{i_1}\rangle \ldots |\psi_k^{i_k}\rangle, \tag{3.40}$$

where $\psi_{i_1,\ldots,i_k} = \langle i_1| \ldots \langle i_k| \psi\rangle$ denotes the quantum tensor of the correlation between the subsystems and $\{|\psi_s^{i_s}\rangle\}$ denotes the quantum states for each subsystem s. The expectation values of local observables can be measured using an approach similar to that outlined in the cases above. To measure $M = O_1 \otimes \cdots \otimes O_k$ with each O_s representing tensor products of local observables on the subsystem s, we have

$$\langle\tilde{\psi}|M|\tilde{\psi}\rangle = \langle\psi|\tilde{O}_1 \otimes \cdots \otimes \tilde{O}_k|\psi\rangle, \tag{3.41}$$

with each observable $\tilde{O}_s$ being

$$\tilde{O}_s^{i'_s,i_s} = \langle\psi_s^{i'_s}|O_s|\psi_k^{i_s}\rangle \tag{3.42}$$

obtained via the method discussed in Section 3.1.3. Here we represent a system of nk qubits by controlling a quantum device with up to $\max\{n,k\}$ qubits. We can also use multiple qubits for each index i_s to increase the bond dimension. Suppose the quantum states are generated as $|\psi\rangle = U|0\rangle_0^{\otimes k}$ and $|\psi_s^{i_s}\rangle = U_s|i_s\rangle|0\rangle_s^{\otimes(n-1)}$, the hybrid tensor network of Figure 3.3(f) can be obtained via a quantum circuit

$$|\tilde{\psi}\rangle = U_k\left(\ldots U_2\left(U_1\left(U|0\rangle_0^{\otimes k} \otimes |0\rangle_1^{\otimes(n-1)}\right) \otimes |0\rangle_2^{\otimes(n-1)}\right) \cdots \otimes |0\rangle_k^{\otimes(n-1)}\right), \tag{3.43}$$

where each U_s applies to the sth qubit of the first k qubits with subscript 0 and the new $n-1$ qubits with subscript s. While such a quantum circuit requires joint control of nk qubits, our hybrid tensor network allows us to represent the same state by controlling up to $\max\{n,k\}$ qubits.

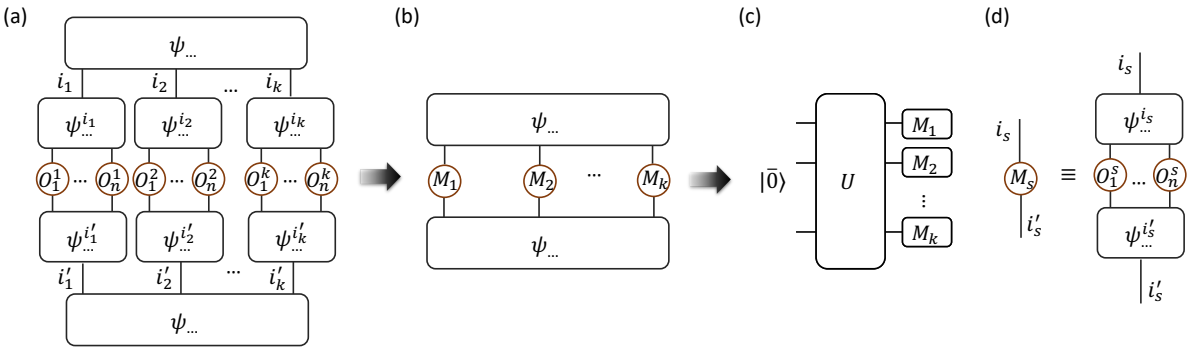


Figure 3.4: An example for calculating expectation values of a hybrid TTN. Considering a hybrid TTN of Figure 3.3(f), the expectation value of local observables $\otimes_{i=1}^{k} \otimes_{j=1}^{n} O_j^i$ corresponds to tensor contraction of (a). We first calculate the observable $M^{i'_s,i_s}$ for each tensor on the second layer as (d) with quantum circuits shown in Figure 3.2(b,d). Then the tensor contraction of (a) reduces to the contraction of (b), which corresponds to a quantum circuit representation in (c) that prepares the state $|\psi\rangle$ with $|\psi\rangle = U\,|\bar{0}\rangle$ and measures the observables $M_1 \otimes \cdots \otimes M_k$.

The 2-layer MPS structure is used by way of an example; other classical tensor networks such as projected entangled-pair state (PEPS) can be similarly used to represent the full quantum state. In the quantum-quantum network, the contracted indices of the second-layer quantum tensors are classical. Figure 3.4 illustrates a quantum circuit representation of this network. The expectation values of arbitrary local observables can be efficiently obtained from measurements using quantum circuits.

3.3 Construction of D-depth tree tensor networks

For quantum systems that have local renormalisation properties, such as a d-dimensional lattice problem, we can construct a D-depth tree tensor network to represent the many-body quantum state. The procedure can be summarised as follows.

1. Divide the quantum system into local units (subsystems).
2. Perform a real-space renormalisation group transformation to produce a coarse-grained system, where a tensor is attached in order to connect the original local units.
3. Repeat procedure 2 until the full quantum system is represented.
4. Contract the D-depth tree tensor network using the contraction rules.

The tensor of the upper layer[2], which connects the subsystems in procedure 2, can be constructed according to the type of interactions between the subsystems. Recall that we can

[2]For example, the "upper" tensor in Figure 3.3 refers to $\psi_{i_1,i_2,\dots,i_k}$.

introduce a classical index to the quantum state of the subsystem by applying unitary operations as $|\psi^i\rangle = U_{\psi^i}|\bar{0}\rangle$. This indeed forms a representation of the target quantum state in the basis given by $\{U_{\psi^i}|\bar{0}\rangle\}$. In the above procedure, each node of the tree can be regarded as a coarse-grained quantum state and the degree of connections among all the tree nodes can be bounded by a constant number t. Denote $C_j = \max\{n_j, t\}$ with n_j representing the maximum number of qubits of the subsystems in the jth layer. Even if we construct the tensors of the upper layer ($(j$-1)th layer) by considering all operators in the interacting Hamiltonian of the subsystems, the maximum bond dimension of tensors scales as $\mathcal{O}(\text{poly}(n_j))$ and at most $\mathcal{O}(\log n_j)$ qubits are required to encode the interactions for the subsystem. The number of controllable qubits to represent the quantum state with system size $\mathcal{O}(t^{D-1})$ is less than $\max_j(C_j \log n_j)$ qubits, which in practice can be greatly reduced by considering boundary conditions or renormalisation properties.

In Section 3.6, I provide an example of 2-layer tree structures for 1D and 2D systems by using different initial states $|\bar{0}\rangle$. We can also specify different unitaries depending on the type of interactions of the target quantum systems. In addition, we may also consider the case where contracted indices of quantum tensors are quantum.

3.4 Resource estimation

Since conventional tensor network theory is a special case of our hybrid TN framework, contracting an arbitrary hybrid TN is at least a #P-hard problem. In this section, I consider special tree networks and give a resource estimation for the cost of calculating expectation values of tensor products of local observables on a hybrid tree tensor network.

Starting from a chosen node referred to as a *root* in a tree, we separate other vertices into different layers according to the distance to the root. Each node can be either a classical tensor or a quantum tensor. In order to efficiently contract the whole tensor network, only classical tensor networks that can be efficiently contracted are considered, for instance, matrix product states (MPS), in Figure 3.3(g). Suppose the tree has a maximum of D layers, and each node has at most g connections, it corresponds to a tree with depth D and maximal degree g, which we call a (D, g)-tree. The total number of nodes in the tree is upper bounded by $\mathcal{O}(g^D)$. The D-th layer has at most g^{D-1} tensors, and each one is at most a g-rank classical tensor network or quantum tensors, with the open index representing at most g qubits. Thus, a (D, g)-tree represents about $\mathcal{O}(g^D)$ physical qubits.

Now, suppose our aim is to measure the expectation value of tensor products of local physical observables. Without loss of generality, consider a node with degree g and denote the cost to be C_q or C_c for a quantum or a classical tensor, respectively. In the ith layer, denote n_i^q and n_i^c to be the numbers of quantum and classical tensors, respectively, which satisfy $n_i^q + n_i^c \leq g^{i-1}$. The cost of contracting the ith layer is thus $C_i = C_q n_i^q + C_c n_i^c \leq g^{i-1}(C_q + C_c)$ and the total cost of

contracting the whole tree is at most $\mathcal{O}(g^D(C_q + C_c))$ by summing all the layers. Note that the number of qubits N represented by the (D, g)-hybrid tree tensor network is $N = \mathcal{O}(g^D)$, so the cost is also linear in the number of qubits. Here we take the sum of the classical and quantum cost. The cost for quantum tensors is the measurement of quantum circuits, while the cost for classical tensors is pure classical computation. We can thus separately use $\mathcal{O}(NC_q)$ and $\mathcal{O}(NC_c)$ to denote the quantum and classical cost. The result can be summarised as follows.

Proposition 1. *The cost for evaluating the expectation values of local observables of a (D, g)-hybrid tree tensor network is at most $\mathcal{O}(g^D(C_q + C_c))$ or $\mathcal{O}(N(C_q + C_c))$. In particular, we need $\mathcal{O}(NC_q)$ quantum circuits and classical cost $\mathcal{O}(NC_c)$. Here $N = \mathcal{O}(g^D)$ is the number of qubits represented by the (D, g)-hybrid tree tensor network.*

The proposition focuses on local observables that have a tensor product form, $O = O_1 \otimes O_2 \otimes \cdot \otimes O_N$ with O_j acting locally on the jth qubit. An arbitrary observable can be expressed as a sum of local observables. For example, it can be decomposed in the Pauli bases as $O = \sum_{i=1}^{M} P_i$ with $P_i \in \{I, X, Y, Z\}^{\otimes N}$ being the tensor product of single-qubit Pauli operators. Here, we require the number of terms M is polynomial in the number of qubits in order to make the method efficient. In the following, we further discuss the magnitude of C_q and C_c. The value of C_q depends on the bond dimension of the index connecting the node to its parent, which quantifies how many measurement settings one needs to contract the quantum tensor to get the effective observable, as illustrated in Figure 3.4(a). For each observable element, C_q also depends on the number of samples required for suppressing shot noise to a desired accuracy ε. Meanwhile, the value of C_c depends on the choice of classical tensor network and the bond dimension of the connecting index as well. Suppose the bond dimension of each contracted index is upper bounded by κ, then we have $C_q = \mathcal{O}(\kappa^2/\varepsilon^2)$ and $C_c = \mathcal{O}(g\kappa^4)$ for MPS (the cost could be improved with more dedicated contraction methods). Therefore, one has the following detailed contraction cost.

Proposition 2. *Consider a (D, g)-hybrid tree tensor network with quantum tensors and classical MPS tensor networks as the building block with bond dimension at most κ. The cost for evaluating the expectation values of local observables is $\mathcal{O}(N(\kappa^2/\varepsilon^2 + g\kappa^4))$. In particular, we need to evaluate $\mathcal{O}(N\kappa^2/\varepsilon^2)$ quantum circuits with $\mathcal{O}(Ng\kappa^4)$ classical computation cost.*

Quantum tensor (C_q)	Classical tensor (C_c)		
	MPS	(D', g')-Tree	PEPS
$\mathcal{O}(\kappa^2/\varepsilon^2)$	$\mathcal{O}(g\kappa^4)$	$\mathcal{O}(g\kappa^{(g'+2)})$	$\mathcal{O}(g\kappa^8\tilde{\kappa}^2)$ (approx.)

Table 3.1: Contraction cost for g-qubit tensors

There are a few additional remarks worth making at this juncture. It appears that the cost reduction when using quantum tensors is not very significant when compared to using classical tensors, that is, κ^2 compared to κ^4. However, this is not actually the case. The quantum tensor

in our hybrid TTN can express more complicated quantum correlations. The entanglement structure of hybrid TTNs is discussed in Appendix A. If one substitutes this quantum tensor with a classical tensor, say MPS, the new bond dimension of the classical tensor denoted by κ' could be much larger than κ, hence leading to a significantly larger cost. This is the major quantum advantage of our hybrid TN framework. Note that here we regard the cost of classical and quantum tensor contractions to be the same, and add them together to be the total cost. In practice, classical and quantum computation are run independently on a classical and a quantum processor, respectively, and so are totally parallel. If we only focus on the resource cost for the quantum processor, the cost scales as $\mathcal{O}(N(\kappa^2/\varepsilon^2))$, which is linear to the number of qubits and polynomial in the bound dimension and inverse polynomial in the simulation accuracy.

I summarise the contraction cost for different types of g-qubit tensors in Table 3.1. Hybrid tensor networks in a generalised tree structure with a small number of loops can be found in [90].

3.5 Applications

Hybrid tensor networks could have a wide array of applications in solving a variety pf physics problems. The key benefit of a hybrid tensor network is to be capable of representing a multipartite quantum state efficiently, so that the required quantum resource can be reduced with the help of classical computers. The hybrid tensor network could extend the power of near-term quantum computers so that the limitation on the number of controllable qubits and the circuit depth could be greatly alleviated.

The most promising application of the hybrid TN is for clustered subsystems with weak subsystem-wise interactions. Consider that the whole quantum system is divided into several subsystems, where particles in the same subsystem are strongly correlated, and the particles from different subsystems are weakly correlated. Since each subsystem has strong interaction, the whole system, in general, could be classically hard to solve. On the contrary, our hybrid TN uses a classical TN solely to represent the cluster-wise interactions and a quantum TN to represent the strongly interacting subsystems. Therefore, hybrid TNs integrate the advantages of classical tensor networks and quantum computing. Here, each subsystem could have a general topology structure of interactions.

In this section, I focus on potential applications of the hybrid tensor network in chemistry problems. Real-world physical systems admit the interaction form considered in our framework. Several chemical molecules can be described by an interacting Hamiltonian that has a similar form as that in our simulation, providing a natural application of our method. The examples include $(Cr_7Ni)_2$ dimers, consisting of two purple-(Cr_7Ni) antiferromagnetic rings. The two Cr_7Ni rings are linked through a pyrazine unit as shown in Figure 3.5, which provides two donor atoms binding to Ni centres in the nearest-neighbour rings. This leads to a weak exchange coupling

between the Ni ions [154,155]. The subsystems have strong correlations, and the two subsystems interact weakly via the boundary spins.

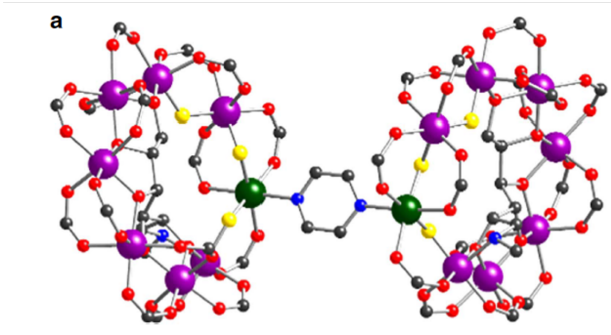


Figure 3.5: Molecular structure of the molecular dimer, consisting of two purple-(Cr_7Ni) antiferromagnetic rings. The two Cr_7Ni rings are linked through a pyrazine unit, which provides two N-donor atoms binding to Ni centres in different rings. This leads to a weak exchange coupling between the Ni ions. The figure is adapted from Ref. [154].

Aside from physical systems featuring these cluster properties, we can use the hybrid TNs to represent different degrees of freedom. The following is an illustration of their application in chemistry for finding molecular vibrational spectra. Methods of calculating the vibrational and electronic structure of a molecule generally assume the Born-Oppenheimer approximation by treating the electrons and nuclei separately. Here, I show how to go beyond the Born-Oppenheimer approximation by using the hybrid tensor network method. Consider the molecular Hamiltonian in atomic units as

$$H_{\text{mol}} = -\sum_i \frac{\nabla_i^2}{2} - \sum_I \frac{\nabla_I^2}{2M_I} - \sum_{i,I} \frac{Z_I}{|\mathbf{r}_i - \mathbf{R}_I|} + \frac{1}{2}\sum_{i\neq j} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} + \frac{1}{2}\sum_{I\neq J} \frac{Z_I Z_J}{|\mathbf{R}_I - \mathbf{R}_J|}. \tag{3.44}$$

with M_I, $\mathbf{R}_I$, and Z_I being the mass, position, and charge of the nucleus I, respectively, and $\mathbf{r}_i$ being the position of electron i. Given the location of the nuclei, the electronic Hamiltonian is

$$H_e(\mathbf{R}_I) = -\sum_i \frac{\nabla_i^2}{2} + \sum_{i,I} \frac{Z_I}{|\mathbf{r}_i - \mathbf{R}_I|} + \frac{1}{2}\sum_{i\neq j} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|}, \tag{3.45}$$

and the total Hamiltonian can be represented as

$$H_{\text{mol}} = -\sum_I \frac{\nabla_I^2}{2M_I} + \frac{1}{2}\sum_{I\neq J} \frac{Z_I Z_J}{|\mathbf{R}_I - \mathbf{R}_J|} + H_e(\mathbf{R}_I). \tag{3.46}$$

Under the Born-Oppenheimer approximation, we assume the electrons and nuclei are in a product state,

$$|\psi\rangle = |\psi\rangle_n |\psi\rangle_e, \tag{3.47}$$

and the ground state energy in the Born-Oppenheimer approximation is given by

$$E_0 = \min_{|\psi\rangle_n} \min_{|\psi\rangle_e} \langle\psi|_n \langle\psi|_e H_{\text{mol}} |\psi\rangle_n |\psi\rangle_e. \tag{3.48}$$

Because only the electronic Hamiltonian $H_e(\mathbf{R}_I)$ depends on electronic state $|\psi\rangle_e$, minimisation over the electronic state $|\psi\rangle_e$ is equivalent to finding the ground state of the electronic Hamiltonian $H_e(\mathbf{R}_I)$. Suppose we solve the electronic structure for any $H_e(\mathbf{R}_I)$ by finding

$$V_0^e(\mathbf{R}_I) = \min_{|\psi\rangle_e} \langle\psi|_e H_e(\mathbf{R}_I) |\psi\rangle_e , \tag{3.49}$$

then the ground state of H_{mol} can be found by solving the ground state of H_0,

$$H_0 = -\sum_I \frac{\nabla_I^2}{2M_I} + \frac{1}{2}\sum_{I\neq J} \frac{Z_I Z_J}{|\mathbf{R}_I - \mathbf{R}_J|} + V_0^e(\mathbf{R}_I). \tag{3.50}$$

The Born-Oppenheimer approximation enables us to solve the molecular Hamiltonian by separately solving the electronic Hamiltonian and the nuclei Hamiltonian. We thus only need to operate a quantum system, either for the electronic Hamiltonian or the nuclei Hamiltonian.

The conventional approach beyond the Born-Oppenheimer approximation is to consider the electrons and nuclei together as a whole system and directly solve the Hamiltonian H_{mol}. However, this requires storing the joint entangled state of electrons and nuclei, making it harder to simulate with near-term quantum computers. Since the nuclei are much heavier than the electrons, beyond the Born-Oppenheimer approximation breaks, the entanglement between electrons and nuclei may still be weak. Therefore, we can use the hybrid tree tensor network to represent the whole state. Suppose the tensor for the electrons and nuclei are $\{|\psi_e^i(\vec{\theta}_e)\rangle\}$ and $\{|\phi_n^i(\vec{\theta}_n)\rangle\}$, respectively. Then a hybrid tensor network representation of the joint state is

$$|\tilde{\psi}\rangle = \sum_i \alpha_i |\psi_e^i(\vec{\theta}_e)\rangle |\phi_n^i(\vec{\theta}_n)\rangle , \tag{3.51}$$

and this is used to represent the ground state of the molecule by only controlling states of either the electrons or the nuclei. We can also apply the hybrid tensor network for representing the electrons or the nuclei to further reduce the size of the quantum system we need to control.

Our method could also be applied to represent virtual qubits. In Ref. [156], the authors demonstrated that they could effectively represent the state in the full space[3] by applying operations on the active space only. They started by preparing the quantum state in the active space as a reference state $|\psi_{\text{Ref}}\rangle$. They chose a set of expansion operators $\{O_i\}$, applied it on the reference, and got $O_i |\psi_{\text{Ref}}\rangle$. This forms a representation of $\{O_i\}$ in the basis given by $\{O_i |\psi_{\text{Ref}}\rangle\}$. The ground state and low-lying eigenstates can be obtained by solving the generalised eigenvalue problem in the well-conditioned subspace as

$$HC = SCE \tag{3.52}$$

[3]Here, the full space includes the active space and the virtual space. In quantum computational chemistry, the active space usually refers to partially occupied spatial orbitals, in which dominant mechanisms of interest lie [44], while the virtual space usually refers to unoccupied spatial orbitals.

with the matrix elements of H and S given by

$$H^{ij} = \langle\psi_{\text{ref}}|O_i^\dagger H O_j|\psi_{\text{ref}}\rangle \quad S^{ij} = \langle\psi_{\text{ref}}\,|O_i^\dagger O_j|\psi_{\text{ref}}\,\rangle. \tag{3.53}$$

The subspace expansion method was demonstrated analytically in Section 3.2 to be one special case in our framework. More specifically, we can choose the quantum tensor as $|\psi_i\rangle = U_i\,|\psi_{\text{Ref}}\rangle$, where U_i is generated by a quantum circuit, and it reduces to the original method when U_i is selected as the single and double excitation operators. We can add a classical tensor to further increase the representation capability in chemistry problems. As the subspace expansion method could potentially improve the accuracy of the ground state and provide approximations to excited states, we expect our method to be applicable to these problems.

3.6 Numerical simulation

In this section, I test the effectiveness of hybrid TNs in finding ground states of 1D and 2D spin lattice systems with nearest-neighbour interactions and external fields.

3.6.1 Model Hamiltonians

I consider the Hamiltonians for 1D and 2D spin systems that admit a general form

$$H = \sum_{j=1}^{k} H_j + \lambda H_{\text{int}}, \tag{3.54}$$

where H_j and H_{int} represents the local Hamiltonian of the jth subsystem and their interactions, with interaction strength of subsystems λ. The topology of the spin systems can be found in Figure 3.6.

For 1D spin clusters, each adjacent $n = 8$ qubits is taken as a subsystem and consider $k = 2, 3, \ldots, 8$ subsystems with $n \times k$ qubits. Consider the local Hamiltonian of the jth subsystem as

$$H_j = \sum_{i=1}^{7} f\hat{Z}_{8j+i}\hat{Z}_{8j+i+1} + \sum_{i=1}^{8}\left(g\hat{X}_{8j+i} + h\hat{Z}_{8j+i}\right)$$

and their interactions as

$$H_{\text{int}} = \sum_{j=1}^{k-1} f_j\hat{Z}_{8j}\hat{Z}_{8j+1}.$$

Here, the interaction strength is characterised by λ, $\hat{X}_i$ and $\hat{Z}_i$ are Pauli operators acting on the ith qubit.

For the 2D $n \times k$ spin lattice, each $n = 3 \times 3$ qubits on a small square lattice are grouped as a subsystem and consider $k = N_x \times N_y$ subsystems with N_x (N_y) subsystems along the x (y) direction. The 2D Hamiltonian is

$$H = \sum_{\langle i,j\rangle} f_{ij}\hat{Z}_i\hat{Z}_j + \sum_i\left(g\hat{X}_i + h\hat{Z}_i\right), \tag{3.55}$$

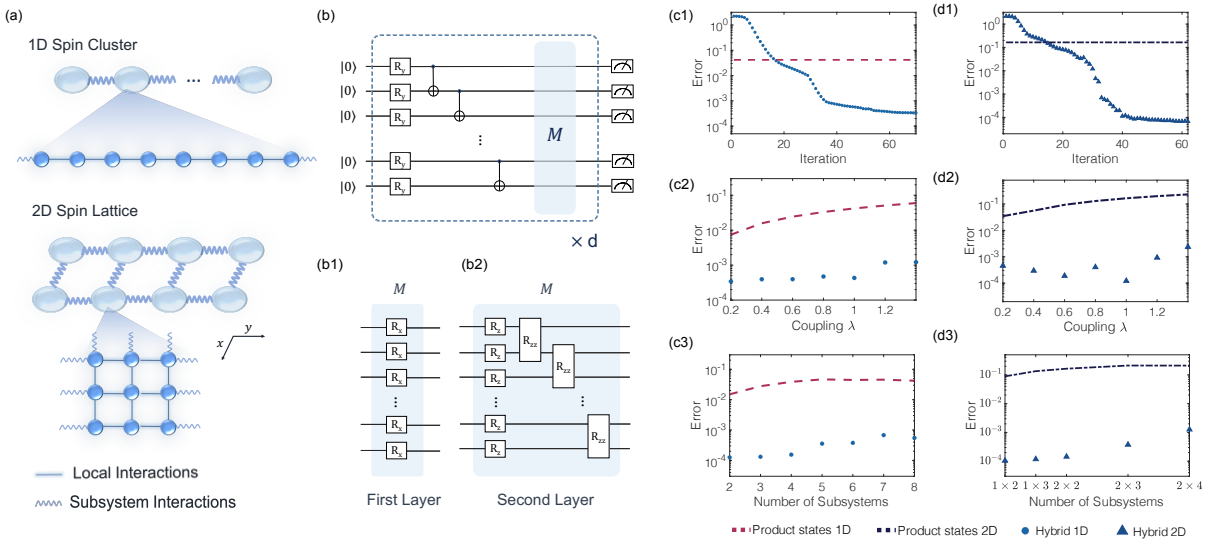


Figure 3.6: Numerical simulation for 1D and 2D quantum systems with hybrid TTN. (a) 1D spin cluster and 2D spin lattice with interactions (thin lines) on the boundary. The interactions of subsystems are represented by thick lines. 8 adjacent qubits and 3×3 qubits on a square sublattice are grouped as subsystems for the 1D and 2D systems, respectively. (b) The ansatz circuit for the quantum tensors in Figure 3.3(f). The circuits of both layers share similar structures, with d repetitions of circuits in the dashed box. Here, R_α ($\alpha \in \{\hat{X}, \hat{Y}, \hat{Z}\}$) represents single-qubit rotation around α-axis and the two-qubit gate is $R_{ZZ}(\theta_i) = e^{-i\theta_i \hat{Z}\otimes\hat{Z}}$. The rotation angle (parameter) for each gate is initialised from a small random value and updated in each variational cycle. The circuit depths for V (first layer tensor) and U (second layer tensor) are $d(V) = 6$ and $d(U) = 8$. The additional unitary M is inserted at the 1st and $[d/2+1]$th block of the first layer (b1) and the second layer (b2). (c)-(d) Simulation results of the ground state energy. For the 1D and 2D cases, we compare E to the reference results $E_0 = E_{\text{MPS}}$ and $E_0 = E_{\text{PEPS}}$ obtained from a standard DMRG with bond dimension $\kappa = 32$ and from PEPS. The relative error $1 - E/E_0$ is used to characterise the accuracy. The red dashed line (1D) and blue dash-dotted line (2D) correspond to the energy using tensor products of the ground state of local subsystems. The cyan dot (1D) and blue triangle (2D) are results obtained with hybrid TNs. (c1, d1) Convergence towards the ground state for the 1D 8×8 and 2D 9×4 systems with $\lambda = 1$, respectively. (c2, d2) Error versus different subsystem coupling strength λ for the 1D 8×8 and 2D 9×4 systems, respectively. (c3, d3) Errors with different numbers of local subsystems with $\lambda = 1$, respectively.

where $\langle i, j\rangle$ represents all the nearest-neighbour pairs on a square lattice. Consider that the interactions in each subsystem are identical $f = 1$, while interactions on the boundary of nearest-neighbour subsystem $\{f_j\}$ or $\{f_{i,j}\}$ are generated randomly from $[0, 1]$, as shown in Figure 3.6(a). The parameters of the external fields for both the 1D and 2D cases are set as $h = 1/\pi = 0.32$ and $g = 0.5$.

The interactions between subsystems $\{f_j\}$ or $\{f_{j,i}\}$ are first generated randomly from $[0, 1]$ and then fixed. The lattice models, which have local interactions in the subsystems and non-local interactions between subsystems of random strengths, have been investigated to describe

the phenomena in high energy physics [157].

3.6.2 Results

Considering the hybrid TTN of Figure 3.3(f), the first layer state and the jth subsystem of the second layer are generated as

$$|\psi\rangle = V(\vec{\theta}_0)\,|\bar{0}^0\rangle = \sum \alpha_{i_1,\ldots,i_k}\,|i_1,\ldots,i_k\rangle \tag{3.56}$$

and

$$|\psi_j^{i_j}(\vec{\theta}_j)\rangle = U(\vec{\theta}_j)\,|\bar{0}^{i_j}\rangle\,, \tag{3.57}$$

respectively, with V and U shown in Figure 3.6(b) and initial states $|\bar{0}^{i_j}\rangle = |i_j\rangle^{\otimes n}$, $i_j \in \{0,1\}$. The hybrid TTN represents a quantum state

$$|\tilde{\psi}(\vec{\theta})\rangle = \sum_{i_1\ldots i_k} \alpha_{i_1,\ldots,i_k}(\vec{\theta}_0)\,|\psi_1^{i_1}(\vec{\theta}_1)\rangle \otimes \cdots \otimes |\psi_k^{i_k}(\vec{\theta}_k)\rangle \tag{3.58}$$

with $\vec{\theta} = (\vec{\theta}_0, \vec{\theta}_1, \ldots, \vec{\theta}_k)$ representing all the parameters. The state is automatically normalised since

$$\langle\psi_j^{i'_j}|\psi_j^{i_j}\rangle = \delta_{i'_j,i'_j}.$$

For parameters $\vec{\theta}$, we obtain the energy expectation value $E(\vec{\theta}) = \langle\tilde{\psi}(\vec{\theta})|H|\tilde{\psi}(\vec{\theta})\rangle$ by following the contraction rule of Figure 3.4. Specifically, we first measure the observables of the second layer states and then measure the effective observables of the first layer state.

I choose to use variational imaginary time evolution to minimise the energy $E(\vec{\theta})$, as shown in Section 2.2.3.3. As demonstrated in Eq. (2.32), we can variationally simulate imaginary time evolution by tracking the evolution of the parameters with one ancillary qubit. This requires the calculation of A and C for given parameters by Eq. (2.34). When $|\psi(\vec{\theta})\rangle$ is directly prepared by a quantum circuit, we can obtain the matrix elements by a modified quantum circuit by introducing an ancillary qubit [120, 131]. When considering trial states represented by a hybrid tensor network, we can calculate A and C by making use of a similar method for calculating the expectation values of hybrid tensor networks. The main idea is to generalise the circuit to implement the contraction of two quantum tensors.

I then show how to calculate the matrix elements by the finite difference method. For example, to calculate each $A_{i,j}$, we can approximate it as

$$\begin{aligned} A_{i,j} &= \mathrm{Re}\left(\frac{(\langle\psi(\vec{\theta}+\delta\theta_i)| - \langle\psi(\vec{\theta})|)}{\delta\theta_i}\frac{(|\psi(\vec{\theta}+\delta\theta_j)\rangle - \partial|\psi(\vec{\theta})\rangle)}{\delta\theta_j}\right), \\ &= \frac{1}{\delta\theta_i\delta\theta_j}\,\mathrm{Re}\left(\langle\psi(\vec{\theta}+\delta\theta_i)|\psi(\vec{\theta}+\delta\theta_j)\rangle - \langle\psi(\vec{\theta})|\psi(\vec{\theta}+\delta\theta_j)\rangle - \langle\psi(\vec{\theta}+\delta\theta_i)|\psi(\vec{\theta})\rangle + \langle\psi(\vec{\theta})|\psi(\vec{\theta})\rangle\right). \end{aligned} \tag{3.59}$$

The last term corresponds to the normalisation of the hybrid tensor, which is 1 for the hybrid TTN considered in the simulation. The second two terms are the overlaps of two different hybrid tensor networks. Again, calculating such overlaps requires quantum circuits similar to calculating expectation values. In our simulation, for simplicity, the overlaps are obtained by directly contracting the quantum tensors by summing over the state vector array. Each C_i element can be obtained via the difference of the energy gradient,

$$C_i = \frac{\langle\psi(\vec{\theta}+\delta\theta_i)|H|\psi(\vec{\theta}+\delta\theta_i)\rangle - \langle\psi(\vec{\theta})|H|\psi(\vec{\theta})\rangle}{2\delta\theta_i}. \tag{3.60}$$

Therefore, the C vector can be obtained from the finite difference of energy changes. We can optimise the total energy by either minimising all the parameters as $\min_{\theta_0,\theta_1,\ldots,\theta_k} \langle\tilde{\psi}(\vec{\theta})|H|\tilde{\psi}(\vec{\theta})\rangle$ or minimising local subsystem of each layer as $\min_i \min_{\{\theta_i\}} \langle\tilde{\psi}(\vec{\theta})|H|\tilde{\psi}(\vec{\theta})\rangle$.

We can readily find that the quantum systems needed for simulating the $8\times k$-qubit 1D and $9\times k$-qubit 2D systems need $8+1$ and $9+1$ qubits, respectively. The calculation is benchmarked by comparison with open-boundary MPS for 1D systems and imaginary time evolution projected entangled-pair state (PEPS) for 2D systems. I choose to consider the relative error $1-E/E_0$ with the ground state energy E and the reference energy E_0. For the 1D case, E is compared to the reference result $E_0 = E_{\mathrm{MPS}}$ obtained from a standard DMRG implementation with bond dimension $\kappa = 32$. For the 2D case, the full quantum state is represented using PEPS. We can start from a random tensor product state, and use a standard imaginary time evolution scheme to find the ground state of the 2D Hamiltonian. To reduce the computational cost, I use the local update method, i.e. the so-called 'simple update' method [158], and set the bond dimension $\kappa = 5$ and the maximum allowed boundary bond dimension, which approximates the original tensor during the contraction, $\tilde{\kappa} = 64$.

Figure 3.6(c1, d1) study the convergence of ground state energy of 1D (c1) and 2D (d1) systems with coupling strength $\lambda = 1$ on 8×8 and 9×4 qubits respectively, and show a relative error below 10^{-3}. Next, I examine how the coupling strength or the number of subsystems affect the efficacy of hybrid TTN. We present the calculation error with respect to different λ for the 8×8-qubit 1D and 9×4-qubit 2D systems in Figure 3.6(c2) and (d2), respectively. We can find that although the error fluctuates with different coupling strengths, which might owe to instability from the optimisation, the error remains consistent around 10^{-3}. In Figure 3.6(c3, d3), I show the calculation error for the 1D with k subsystems (c3) and 2D with $N_x\times N_y$ subsystems (d3) for $\lambda = 1$, and we can achieve a desired simulation accuracy.

The simulation results demonstrate that we can decrease the error to a relatively low level, which indicates the effectiveness of our method in a proof-of-principle way. In practice, we can use different types of optimisation method and circuit ansatz to further reduce simulation error. The results with different coupling strengths and numbers of subsystems verify the robustness of

the hybrid TTN method. The spin lattice problem with uniform nearest-neighbour interactions on the boundary is shown in Appendix A.

3.7 Discussion

The Hybrid tensor networks established in this chapter provide a natural way of performing qubit-time trade-off in quantum computing. Quantum resources (such as qubit number and circuit depth) can be reduced by replacing some of the quantum tensors with their classical counterparts, at the cost of increasing running time. Our protocol also lends itself to studying the physics of many-body systems, with an approach that combines quantum and classical computation in an appropriate form. This is particularly useful when the quantum resource is limited and we have some prior knowledge of the target problem.

In this chapter, I have chosen to exemplify our hybrid TTN using a quantum-quantum tensor network, which can be implemented on a near-term quantum computer using variational quantum algorithms. Other hybrid TTN structures with classical tensors can be implemented in a similar way. We may also simulate other systems to explore interesting physics behind these models, such as the models discussed in Chapter 8.

Finally, I would like to discuss the limitations of hybrid TTNs. As described, a hybrid TTN can be used as a heuristic ansatz for systems, and enables the simulation of large quantum systems with fewer qubits. However, it may not guarantee simulation accuracy for a general quantum system. With a very small bond dimension κ, hybrid TTNs may be inadequate to represent a target quantum system. This is similar to classical tensor networks. It is expected that the representation capability of hybrid TTNs could be increased by increasing the bond dimension. A rigorous analysis of the representation capability of hybrid TTNs, and its advantage over classical tensor networks, could be an interesting avenue of future research.

Author contributions. X.Y. and J.S. conceived the idea of the hybrid tensor network ansatz for representing quantum many-body states. X.Y., J.S., and J.L initiated the project. X.Y. and J.S developed the theoretical framework of hybrid tensor networks. J.S. conceived the implementation of hybrid tree tensor networks (TTN) and carried out the numerical simulation. J.S., J.L., and X.Y. conceived the applications of hybrid TTN. Q.Z. and Y.Z. analysed the resource cost of hybrid TTN. X.Y. supervised the project.

Chapter 4

Eigenenergy estimation and eigenstate property estimation

Estimating properties of the ground state and excited states of a quantum many-body system is a long-standing problem of fundamental interest in condensed matter physics, quantum chemistry, and material science. This is a challenging problem for both classical computing and quantum computing, and indeed estimating the ground state energy of a general Hamiltonian is a QMA-complete problem. Despite this, quantum computing promises to address this problem under certain assumptions. The common assumptions include a nonvanishing energy gap and a nonvanishing overlap of the initial state and the target state, as in phase estimation [29] and qubitisation [31].

The problems considered in this chapter (eigenenergy estimation and eigenstate property estimation) were initially formalised in Section 1.2.1, followed by a discussion on the intuition in Section 2.3. In this chapter, I discuss a solution to the two problems by realising generalised quantum imaginary time evolution with a concrete theoretical guarantee. Gate complexity and compilation to standard quantum gates will be discussed in Chapter 9.

This chapter is relevant to a preprint in collaboration with Pei Zeng and Xiao Yuan [93] and a manuscript under preparation [92].

4.1 Motivation

Among the various task-specific classical algorithms — such as perturbation theories [95,159,160], variational approaches [161, 162], self-consistent embedding methods [163, 164], machine learning [165–167], etc, — imaginary time evolution [168] defines a natural and universal cooling procedure, which can also be regarded as a projection process. Consider a Hamiltonian H, imaginary time evolution $e^{-\tau H}$ with a real-valued time τ monotonically lowers the average energy and deterministically drives an arbitrary pure state to the lowest eigenstate that has a nonzero overlap. Despite being mathematically universal, realising imaginary time evolution for an arbitrary Hamiltonian is by no means an easy problem. The notorious sign problem [169, 170] in

the Monte Carlo implementation of imaginary time evolution has limited its usage for solving general strongly correlated problems [171, 172].

As motivated in Section 2.3, a new paradigm is proposed for realising general quantum projection procedures on a quantum computer. For a general decaying function $g(h)$, such as $g = e^{-|h|}$ or $g = e^{-h^2}$, I demonstrate how to realise a projection procedure $g(\tau H)$ for a Hamiltonian H using solely real-time evolution and proper classical post-processing. The algorithm essentially implements general non-Hermitian dynamics, and can efficiently find energy spectra and the corresponding eigenstates, with any initial states that have nonvanishing overlaps with the target state. It is worth pointing out that the term 'projection' is commonly referred to as a low-rank projector in the context of quantum information and quantum computing. However, $g(\tau H)$ regarded as a spectral projector of H since it can be applied to prepare the eigenstate by projecting out the contributions from other eigenstates. Hereafter, we refer to $g(\tau H)$ as a projection operator.

Compared to the variational approaches in Section 2.2, this method works deterministically without any assumption on parameterised circuits. Compared to other universal quantum algorithms, such as quantum phase estimation, our method only needs an exponentially shallower circuit with only one ancillary qubit, and it reaches the Heisenberg limit of eigenenergy estimation.

Ground state preparation can be understood as a cooling procedure, a deterministic process that drives any state to the lowest eigenstate. In this process, we do not experimentally cool down the system, but instead implement the cooling by applying a few basis operations e^{-iHt} with different time lengths. Therefore, in the paper [93], we also term this state preparation method as *algorithmic cooling*.

The essential component of the ground state property estimation is to realise a cooling operation (or projection operation), which projects out the contributions from the other eigenstates. In Section 4.2, I introduce the basics of cooling operations. I demonstrate how this can be applied to estimate the eigenstate property and the eigenenergy in Section 4.3. The complexity analysis is shown in Section 4.4. Finally, numerical tests are presented in Section 4.5.

4.2 Cooling operation and the dual phase representation

First, a general matrix function acting on the Hamiltonian is defined, which is expressed as

$$g(H) := \sum_{i=0}^{N-1} g(E_i) \left|u_i\right\rangle \left\langle u_i\right| . \tag{4.1}$$

where $g(h) : \mathbb{R} \to \mathbb{C}$ is a generic continuous-variable function determining the transformation of the spectrum of the Hamiltonian. To realise the cooling process, $g(h)$ is required satisfying the following property.

Definition 1. *A real-valued function $g(h) : \mathbb{R} \to \mathbb{C}$ is called a cooling function if it satisfies,*

1. *The value of $g(0)$ is non-zero: $g(0) \neq 0$;*
2. *Strictly non-increasing absolute value: $|g(h')| < |g(h)|,\ \forall h' > h > 0$ or $h' < h < 0$;*
3. *Vanishing asymptotic value: $\lim_{\tau\to\infty} |g(\tau h')/g(\tau h)| = 0,\ \forall h' > h > 0$ or $h' < h < 0$ or alternatively $\lim_{\tau\to\infty} |g(\tau h)/g(0)| = 0,\ \forall |h| > 0$.*

A typical example is the exponential decreasing function $g(h) = e^{-|h|}$, corresponding to imaginary time evolution $g(\tau H) = e^{-\tau H}$ for positive H.

Suppose we want to prepare the jth eigenstate $|u_j\rangle$ with the eigenenergy E_j. With the definition in Eq. (E.15), a cooling operator is introduced as follows,

$$g(\tau(H - E_j)) = \sum_{i=0}^{N-1} g(\tau(E_i - E_j)) \, |u_i\rangle \langle u_i| \, . \tag{4.2}$$

For an arbitrary input state $|\psi_0\rangle = \sum_i c_i |u_i\rangle$, the state after the cooling $g(\tau(H - E_j))$ at a finite τ is given by

$$|\psi(\tau)\rangle = \frac{g(\tau(H - E_j)) \, |\psi_0\rangle}{\| g(\tau(H - E_j)) \, |\psi_0\rangle \, \|} = \frac{\sum_i g(\tau(E_i - E_j)) c_i \, |u_i\rangle}{\sum_i p_i g(\tau(E_i - E_j))^2}, \tag{4.3}$$

with $p_i = |c_i|^2$. Since for the eigenenergies $E_i \neq E_j$, the function $g(\tau E_i)$ decreases faster with τ, the amplitudes of the normalised state $|\psi(\tau)\rangle$ concentrate to the eigenstates with energy E_j, and the evolved state asymptotically approximates the eigenstate $|u_j\rangle$ with nonzero $|\langle\psi_0|u_j\rangle|^2 \neq 0$ for sufficiently large τ as

$$\lim_{\tau\to\infty} g(\tau(H - E_j)) \, |\psi_0\rangle \propto |u_j\rangle \, . \tag{4.4}$$

There are two important parameters, τ and E_j, in the Hamiltonian function $g(h)$ in the cooling procedure, in which τ is a time scaling factor, and E_j indicates a shift of the original function. In later discussion, we shall see that τ indicates the timescale for the cooling procedure, in which larger τ will cool the state closer to the target eigenstate. The shift E_j plays an important role in searching for the eigenenergies and, additionally, observable estimation on the eigenstate. From the above equation, readers may wonder if a concentration around the eigenstate $|u_j\rangle$ with energy E_j can be obtained if the exact value of E_j is not known a priori. It is therefore worth noting that E_j only appears in classical post-processing and will not be involved in quantum measurements. In the implementation, E_j is tuned as a classical variable to find the true eigenenergy, and thus does not increase any quantum resource cost.

The next step is to realise the cooling operator, which is nonunitary by construction. I show how to implement it by using a combination of real-time dynamics. Consider its dual realisation based on a Fourier transform, $g(h) = \frac{1}{2\pi}\int_{-\infty}^{\infty} f(x) e^{ixh} dx$. and its inverse form $f(x) =$

$\int_{-\infty}^{\infty} g(h)\, e^{-ixh} dh$. Given the norm of $f(x)$ defined as $\|f\| = \int_{-\infty}^{\infty} |f(x)| \mathrm{d}x$, consider the normalised function $p(x) = |f(x)|/\|f\|$ and we have

$$g(h) = c \int_{-\infty}^{\infty} e^{i\theta_x} p(x) e^{ixh} \mathrm{d}x, \tag{4.5}$$

where $c := \|f\|/2\pi$ and $e^{i\theta_x} := f(x)/|f(x)|$. Here, the function $g(h)$ is required to be *Fouirer-realisable*, which is defined as follows.

Definition 2. *A function $g(h)$ in Definition 1 is called Fouirer-realisable if the following requirements hold.*

1. *The Fourier transform of $g(h)$, $f(x) = \int_{-\infty}^{\infty} g(h) e^{-ixh} dh$ exists.*
2. *The norm of the dual function $\|f\|$, $\|f\| := \int_{-\infty}^{\infty} |f(x)| dx$, is finite.*
3. *$\exists L(\varepsilon) = \mathcal{O}(\mathrm{poly}(\frac{1}{\varepsilon}))$, such that $\left|1 - \int_{-L(\varepsilon)}^{L(\varepsilon)} p(x) dx\right| \leq \varepsilon,\ \forall \varepsilon \geq 0$.*

The third condition implies that a finite frequency in $[-L(\varepsilon), L(\varepsilon)]$ with $L(\varepsilon) = \mathcal{O}(\mathrm{poly}(1/\varepsilon))$ is sufficient for approximating the normalised function $g(h)/c$, i.e., $\left|g(h)/c - \int_{-L(\varepsilon)}^{L(\varepsilon)} p(x) e^{ixh+\theta_x} \mathrm{d}x\right| \leq \varepsilon$. Table 4.1 gives several examples of $g(h)$. Since our examples have zero phase, for a simpler presentation, the phase terms are set to be 0 in the following discussion, although the discussion also naturally generalises to cases with nonzero phases.

I introduce several typical cooling functions $g(h)$ and discuss their properties. The most important property is the cutoff error $L(\varepsilon)$ defined in Definition 2 which shows the hardness of realising $g(h)$ with real-time sampling.

For the exponential function we have $L(\varepsilon) = 2/(\pi\varepsilon)$. For the Gaussian function, we have $L(\varepsilon) = 2\sqrt{\ln(\varepsilon^{-1})}$. Therefore, the Gaussian function is a realisable cooling function. Note that unlike the exponential function, the tail of the Gaussian function decays quickly with respect to x_m. This implies that Gaussian cooling is more experimentally-friendly. The properties of the representative cooling functions are summarised in Table 4.1.

Table 4.1: Representative cooling functions $g(h)$. $f(x)$ corresponds to the Fourier transform of $g(h)$.

Function	$g(h)$	$f(x)$	$\|f\|/2\pi$	$L(\varepsilon)$
Exponential	$e^{-\|h\|}$	$\frac{2}{x^2+1}$	1	$2/(\pi\varepsilon)$
Gaussian	e^{-h^2}	$\sqrt{\pi} e^{-x^2/4}$	1	$2\sqrt{\ln(1/\varepsilon)}$
Hyperbolic	$\mathrm{sech}(h)$	$\pi\mathrm{sech}(\pi x/2)$	1	$\frac{2}{\pi}\ln(1/\varepsilon)$

In this thesis, I mainly focus on the analysis of the Gaussian function $g(\tau H) = e^{-\tau^2 H^2}$ in the application. It is easy to verify that the three condition hold. We have the dual form of the

Gaussian function $f(x) = \sqrt{\pi}e^{-x^2/4}$, the norm is finite $\|f\| = 2\pi$, and the truncation at the finite frequency $L(\varepsilon) = 2\sqrt{\ln(1/\varepsilon)}$ is sufficient to ensure the truncation error,

$$\left|g(h) - \int_{-L(\varepsilon)}^{L(\varepsilon)} p(x)e^{ixh}\mathrm{d}x\right| \leq \varepsilon,$$

is less than ε.

With the normalised dual phase representation $p(x)$ of $g(h)$, the cooling operator is

$$g(\tau(H-E)) = c\int_{-\infty}^{\infty} \mathrm{d}p(x)e^{i\theta_x}e^{ix(H-E)}. \tag{4.6}$$

The quantum state in Eq. (4.3) becomes,

$$\begin{aligned} |\psi(\tau)\rangle &\propto g(\tau(H-E))\,|\psi_0\rangle \\ &\propto c\int_{-\infty}^{\infty} \mathrm{d}p(x)e^{i(\theta_x-\tau xE)}\,|\phi(x\tau)\rangle\,, \end{aligned} \tag{4.7}$$

which is now a superposition of real-time evolved states $|\phi(x\tau)\rangle = e^{ix\tau H}\,|\psi_0\rangle$ with probability distribution $\mathrm{d}p(x) = p(x)\mathrm{d}x$. In Section 4.3, I will show how to effectively prepare[1] the quantum state described by Eq. (4.7) through quantum algorithmic cooling.

4.3 Quantum algorithmic cooling

Suppose we want to estimate the expectation value of an observable O with respect to an eigenstate $|u_j\rangle$, $\langle O\rangle := \langle u_j|O|u_j\rangle$. For a given finite evolution time τ, the estimation value of $\langle O\rangle_\tau$ is given by

$$\langle O\rangle_\tau = \frac{N_\tau(O)}{D_\tau}, \tag{4.8}$$

where

$$\begin{aligned} D_\tau &= \langle\psi_0|g^2(\tau(H-E_j))|\psi_0\rangle\,, \\ N_\tau(O) &= \langle\psi_0|g(\tau(H-E_j))Og(\tau(H-E_j))|\psi_0\rangle\,. \end{aligned} \tag{4.9}$$

are, respectively, the normalisation factor and the unnormalised expectation value.

The basic idea is that the numerator $N_\tau(O)$ can be efficiently obtained by sampling x and x' from the distribution $p(x)$ and $p(x')$ (defined in Eq. (4.6)), respectively, and then estimating the mean value $\mathbb{E}_{x,x'}\langle\phi(x'\tau)|O|\phi(x\tau)\rangle$, where each term can be evaluated on a quantum computer with a Hadamard test circuit. The denominator D_τ can be similarly obtained by estimating $\mathbb{E}_y\langle\psi_0|e^{iy\tau H}|\psi_0\rangle$ with probability $\mathrm{d}\tilde{p}(y)$ (see Eq. (4.13) below). A quantum computer is only required to estimate the expectation values of a state overlap like $\langle\psi_0|U|\psi_0\rangle$ with U being either $e^{-ix'\tau H}Oe^{ix\tau H}$ or $e^{iy\tau H}$, and then we can effectively obtain the time-dependent expectation value of any observable by post-processing the measurement outcomes. In practice, we also need to consider a cutoff of the integral from $[-\infty,\infty]$ to $[-x_m, x_m]$ to avoid infinite integration.

[1] This chapter is focused on eigenenergy estimation and eigenstate property estimation. The latter corresponds to obtaining the expectation value of an observable measured on the target eigenstate. This task will be less demanding when compared to directly preparing the target eigenstate.

I will now demonstrate how to estimate state overlaps by a Hadamard test circuit. The target observable O can be decomposed by the Pauli operators,

$$O = \sum_{l\in\mathcal{P}_n} o_l P_l = \|O\|_1 \sum_{l\in\mathcal{P}_n} \mathrm{Pr}_O(l) P_l \tag{4.10}$$

where $\mathcal{P}_n$ denotes the n-qubit Pauli group, o_l is the coefficient of the Pauli component P_l, and here the coefficients $\{o_l\}$ are all set to be positive. The signs of the coefficients are put into the corresponding Pauli matrices $\{P_l\}$. In Eq. (4.10), $\|O\|_1$ is the l_1-norm of the Pauli coefficients of O,

$$\|O\|_1 = \sum_{l\in\mathcal{P}_n} o_l. \tag{4.11}$$

The probability distribution $\mathrm{Pr}_O(l)$ is defined to be,

$$\mathrm{Pr}_O(l) = \frac{o_l}{\sum_{l\in\mathcal{P}_n} o_l} = \frac{1}{\|O\|_1} o_l. \tag{4.12}$$

Using Eq. (4.6), one can expand the cooling operator $g(\tau(H-E_j))$ and receive the following estimation formulae,

$$\begin{aligned} D_\tau &= c^2 \int_{-\infty}^{\infty} dy \tilde{p}(y) e^{-i\tau y E_j} \langle\psi_0| e^{i\tau y H} |\psi_0\rangle\,, \\ N_\tau(O) &= c^2 \|O\|_1 \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dx' p(x) p(x') \sum_{l\in\mathcal{P}_n} \mathrm{Pr}_O(l) e^{-i\tau(x-x')E_j} \langle\psi_0| e^{-i\tau x' H} P_l e^{i\tau x' H} |\psi_0\rangle\,, \end{aligned} \tag{4.13}$$

where $\tilde{p}(y) := \frac{1}{2}\int_{-\infty}^{\infty} p(\frac{z+y}{2}) p(\frac{z-y}{2}) dz$ is the self-correlation of the probability $p(x)$. The normalisation factor c^2 can be removed, since it is the same for both D_τ and $N_\tau(O)$, hence, unimportant for the estimation of $\langle O\rangle_\tau$. In the following discussion, this normalisation factor is ignored during the estimation procedure.

In the quantum cooling algorithm, the (normalised) evolution time x (or y) is generated based on the probability $p(x)$ (or $\tilde{p}(y)$). For the sake of practicality, when the (normalised) evolution time x (or y) is larger than a cutoff value x_m, we denote the estimation of this round to be 0 without performing a quantum experiment. In this way, the estimation formula of $N_\tau(O)$ and D_τ, originally given by Eq. (4.13), now becomes

$$\begin{aligned} D_\tau^{(x_m)} &= \int_{-x_m}^{x_m} dy \tilde{p}(y) e^{-i\tau y E_j} \langle\psi_0| e^{i\tau y H} |\psi_0\rangle\,, \\ N_\tau^{(x_m)}(O) &= \|O\|_1 \int_{-x_m}^{x_m} dx \int_{-x_m}^{x_m} dx' p(x) p(x') \sum_{l\in\mathcal{P}_n} \mathrm{Pr}_O(l) e^{-i\tau(x-x')E_j} \langle\psi_0| e^{-i\tau x' H} P_l e^{i\tau x' H} |\psi_0\rangle\,. \end{aligned} \tag{4.14}$$

To estimate the values of $N_\tau^{(x_m)}(O)$ and $D_\tau^{(x_m)}$ in Eq. (4.14), the core issue is to realise an unbiased estimation of the following quantities,

$$\langle\psi_0| e^{i\tau y H} |\psi_0\rangle\,, \qquad \langle\psi_0| e^{-i\tau x' H} P_l e^{i\tau x' H} |\psi_0\rangle\,. \tag{4.15}$$

These $\langle\psi|U|\psi\rangle$ can be estimated using the Hadamard test, shown in Figure 4.1.

To measure $\langle\psi|U|\psi\rangle$, the state $|\psi\rangle$ is first prepared and an extra ancillary qubit is prepared in the computational basis state $|+\rangle$. Afterward, a C-U gate from ancillary to $|\psi\rangle$ is performed. If we directly measure the ancillary qubit in the X-basis, the outcome a will be 0 with a probability of $\frac{1}{2}(1+\mathrm{Re}(\langle\psi|U|\psi\rangle))$ and 1 with a probability of $\frac{1}{2}(1-\mathrm{Re}(\langle\psi|U|\psi\rangle))$. Alternatively, if we perform an extra S gate before the X-basis measurement, the outcome a will be 0 with a probability of $\frac{1}{2}(1+\mathrm{Im}(\langle\psi|U|\psi\rangle))$ and 1 with a probability of $\frac{1}{2}(1-\mathrm{Im}(\langle\psi|U|\psi\rangle))$.

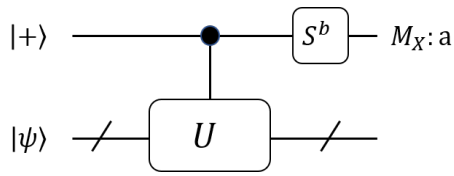


Figure 4.1: The diagram of the Hadamard test. If $b=0$, $\Pr(a=0)=\frac{1}{2}(1+\mathrm{Re}(\langle\psi|U|\psi\rangle))$; if $b=1$, $\Pr(a=0)=\frac{1}{2}(1+\mathrm{Im}(\langle\psi|U|\psi\rangle))$.

To simplify the theoretical analysis, I now introduce a combined estimation of the real part and imaginary parts in a single round. In each round of experiments, we first randomly decide the binary value b uniformly from $\{0,1\}$ in Figure 4.1. Denote the binary uniform distribution of b as $\Pr_U(b)$. Based on the measurement outcome a, the following complex-valued estimator is constructed,

$$\hat{r}=\begin{cases}2, & (b,a)=(0,0),\\ -2, & (b,a)=(0,1),\\ 2i, & (b,a)=(1,0),\\ -2i, & (b,a)=(1,1).\end{cases} \tag{4.16}$$

That is, $\hat{r}=2i^b(-1)^a$. In this case, it is easy to verify that $\hat{r}$ is an unbiased estimator for $\langle\psi|U|\psi\rangle$ as $\mathbb{E}(\hat{r})=\langle\psi|U|\psi\rangle$.

In the estimation of D_τ and $N_\tau(O)$, the phase term and normalisation factor $\|O\|_1$ in Eq. (4.13) can also be put into the estimator. That is to define

$$\begin{aligned}\hat{d}&=e^{-i\tau y E_j}\hat{r},\\ \hat{n}&=\|O\|_1 e^{-i\tau(x-x')E_j}\hat{r}.\end{aligned} \tag{4.17}$$

As a result, if y is randomly sampled from $\tilde{p}(y)$, based on Eq. (4.13), we have

$$\mathbb{E}_{y,b,a}(\hat{d})=D_\tau. \tag{4.18}$$

On the other hand, if we randomly sample x,x' from $p(x)$ and l from $\Pr_O(l)$, based on Eq. (4.13) we have

$$\mathbb{E}_{x,x',l,b,a}(\hat{n})=N_\tau(O). \tag{4.19}$$

In the experiments, variables x, x', a, b and l are generated independently, randomly in each round of experiment. After performing single-shot Hadamard test experiments, we will obtain a

group of unbiased estimators $\{\hat{d}_p\}_{p=1}^{N_M}$ and $\{\hat{n}_q\}_{p=1}^{N_M}$ for D_τ and $N_\tau(O)$, respectively. We then use the average value of the estimators as an accurate estimate of D_τ and $N_\tau(O)$, whose tightness is given by the concentration bound analysed in Section 4.4.2, as well as in Section B.1.3 and Section B.3.3.

There are two main applications of the universal algorithmic cooling: (1) The estimation of eigenenergies; (2) The estimation of the properties of the jth eigenstate with or without exact knowledge of the energy.

This can be divided into three elementary tasks:

1. Estimate the eigenenergy E_j of the jth eigenstate given an initial guess interval $[E_j^L, E_j^U]$;
2. Given a known eigenenergy E_j, estimate the normalisation factor D_τ;
3. Given a known eigenenergy E_j and an observable O, estimate the normalisation factor $N_\tau(O)$.

The algorithm for these tasks is introduced below.

4.3.1 Estimation of the eigenenergy

To start with, I first show how to estimate the normalisation factor D_τ based on Eq. (4.14) and the Hadamard test, when the eigenenergy E_j is known. Recall that the unbiased estimator for the normalisation factor is,

$$\hat{d} = e^{-i\tau y E_j}\hat{r} = 2i^b(-1)^a e^{-i\tau y E_j}. \tag{4.20}$$

If we randomly sample y from $\tilde{p}(y)$, based on Eq. (4.13) we have

$$\mathbb{E}_{y,b,a}(\hat{d}) = D_\tau. \tag{4.21}$$

I now show how to estimate the eigenenergy using the normalisation factor. Suppose the initial state $|\psi_0\rangle$ has a constant overlap with the eigenstate $|u_j\rangle$. We can then sweep the parameter $E_j^{(e)}$ in a range to maximise D_τ. The estimation accuracy depends on the finite imaginary time τ, finite truncation x_m, sampling number N_M and the overlap p_j.

To clarify this, consider a simple case where the jth eigenenergy is known to be in a range $E_j \in [E_j^L, E_j^U]$. Moreover, suppose other eigenenergies lie far beyond this range. Expand the initial state in the eigenstate basis as

$$|\psi_0\rangle = \sum_{j=0}^{N-1} c_j |u_j\rangle, \tag{4.22}$$

and denote the square overlap of $|\psi\rangle$ and $|u_i\rangle$ to be

$$p_i = |\langle u_i|\psi\rangle|^2 = |c_i|^2. \tag{4.23}$$

For an energy value E in the range $[E_j^L, E_j^U]$, the ideal normalisation factor D_τ is calculated as,

$$\begin{aligned} D_\tau(E) &= \langle\psi|g^2(\tau(H-E))|\psi\rangle \\ &= g^2(\tau(E_j-E))p_j + \sum_{i\neq j} g^2(\tau(E_i-E))p_i \\ &\approx g^2(\tau(E_j-E))p_j. \end{aligned} \tag{4.24}$$

The approximation holds when $g^2(\tau(E_j-E))p_j \gg g^2(\tau(E_i-E))p_i$. This naturally holds when the eigenenergies $\{E_i\}_{i\neq j}$ are far from the range $[E_j^L, E_j^U]$. From Eq. (4.24) we can see that, the normalisation factor takes a local maximum value close to p_j when the energy value $E = E_j$,

$$E_j = \underset{E\in[E_j^L, E_j^U]}{\operatorname{argmax}} \; D_\tau(E). \tag{4.25}$$

Therefore, we can sweep the values of E in a range $[E_j^L, E_j^U]$ to search the largest value of $D_\tau(E)$. Here, a free usage of classical computational resources is assumed, as we mainly focus on the quantum resources. This may be refined by the introduction of a better classical peak-value searching algorithm.

4.3.2 Estimation of the unnormalised observable expectation value

In this section, I discuss the estimation $N_\tau(O)$ of a given observable,

$$O = \sum_{l\in\mathcal{P}_n} o_l P_l = \|O\|_1 \sum_{l\in\mathcal{P}_n} \mathrm{Pr}_O(l) P_l \tag{4.26}$$

which is defined in Eq. (4.10).

Recall the unbiased estimator of $N_\tau(O)$ is

$$\hat{n} = \|O\|_1 e^{-i\tau(x-x')E_j}\hat{r} = 2\|O\|_1 e^{-i\tau(x-x')E_j} i^b (-1)^a. \tag{4.27}$$

If we randomly sample x, x' from $p(x)$ and l from $\mathrm{Pr}_O(l)$, based on Eq. (4.13) we have

$$\mathbb{E}_{x,x',l,b,a}(\hat{n}) = N_\tau(O). \tag{4.28}$$

$N_\tau(O)$ given in Eq. (4.14) can be similarly measured by the Hadamard test. With the above results, we can solve two problems, namely, eigenstate property estimation with known eigenstate energy, and eigenstate property estimation with unknown eigenstate energy.

4.4 Complexity analysis

In this section, I show the error and resource requirement for eigenenergy estimation and eigenstate property estimation in Section 4.4.1 and Section 4.4.2, respectively.

4.4.1 Error and resource requirement analysis for eigenenergy estimation

In this section, the estimation error of the jth eigenenergy is analysed. Based on the error dependence, I estimate the resource requirements (i.e., circuit depth and sampling number) of the energy estimation.

For simplicity, I focus on a case when the following assumptions hold. The eigenvalue searching range $[E_j^L, E_j^U]$ is a subset of $[E_j - \frac{\Delta}{2}, E_j + \frac{\Delta}{2}]$, where Δ is a known lower bound of $\min\{E_j - E_{j-1}, E_{j+1} - E_j\}$. Under this assumption, the normalisation factor can then be expressed as,

$$D_\tau(E) \approx p_j g(\tau(E - E_j))^2. \tag{4.29}$$

Here, the approximation holds when the values $g(\tau(E - E_i))^2$ contributed by other eigenenergies E_i are negligible. The location of the peak of $D_\tau(E)$ then indicates the eigenenergy E_j.

In practice, however, we can only obtain the estimation $\hat{D}_\tau^{(x_m)}(E)$ of $D_\tau(E)$, considering the finite cutoff time x_m and finite sample number N_M. We are going to prove that, the solution of the following maximisation problem

$$\hat{E}_j := \mathrm{argmax}_{E \in [E_j^L, E_j^U]} \hat{D}_{\tau(E)}^{(x_m)}, \tag{4.30}$$

will be close to the real solution E_j.

The sources of error include the finite imaginary time, finite cutoff of real-time evolution, discretisation error, and statistical error due to a finite sampling number,

$$D_\tau^{x_m,r}(\hat{E}_j) - D_\tau(E_j) \leq \varepsilon_\tau + \varepsilon_x + \varepsilon_r. \tag{4.31}$$

Based on the error dependence, the resource requirement (i.e., circuit depth and sample number) of the energy estimation can be estimated. The workflow of error analysis is summarised in Figure 4.2.

As a complete analysis involves many technical details and the proof is lengthy, I will show the main results here to streamline the discussion.

Theorem 1 (Accuracy of the eigenenergy estimation)**.** *Given constant $K > 0$, error $\varepsilon \in (0, 1)$, finite imaginary time $\tau \geq g^{-1}(\alpha p_j/6)/\kappa$, cutoff time $x_m \geq \sqrt{2}L\,(\alpha p_j/6)$, and sample number $N_M \geq 9Kp_j^{-2}/\alpha^2$ where $\alpha := 1 - g(1)$, the error between the estimated eigenenergy $\hat{E}_j$ and the one ideal E_j is $|\hat{E}_j - E_j| \leq \kappa$, with a failure probability of $2\delta = 4\exp(-K/8)$.*

Since only the imaginary time τ depends on the inverse accuracy $1/\kappa$, the total cost $\tau \cdot x_m \cdot N_M$, which is circuit complexity $\times$ sample complexity, also scales as $1/\kappa$, indicating Heisenberg's limit for eigenenergy estimation.

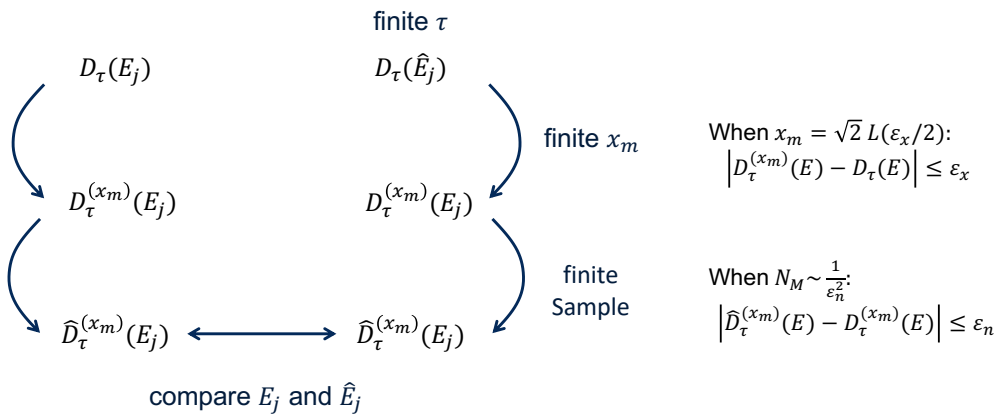


Figure 4.2: Summary of the error analysis of eigenenergy estimation. First study the effect of the finite imaginary time τ, finite normalised cutoff time x_m and finite sampling number N_M on the normalisation estimation, and then bound the difference between the estimation value $\hat{E}_j$ from true value E_j.

4.4.2 Eigenstate property estimation

In this section, I analyse the estimation error of the observable value $\langle O \rangle$ under a finite circuit depth and sampling numbers. Based on the error dependence, the resource requirements (i.e. circuit depth and sample number) of the cooling algorithm are estimated. For simplicity, it is assumed that we have already obtained a precise eigenenergy estimation E_j of the jth eigenstate. Similar to Section 4.4.1, the sources of error include the finite imaginary time, finite cutoff of real-time evolution, discretisation error, and statistical error due to a finite sampling number.

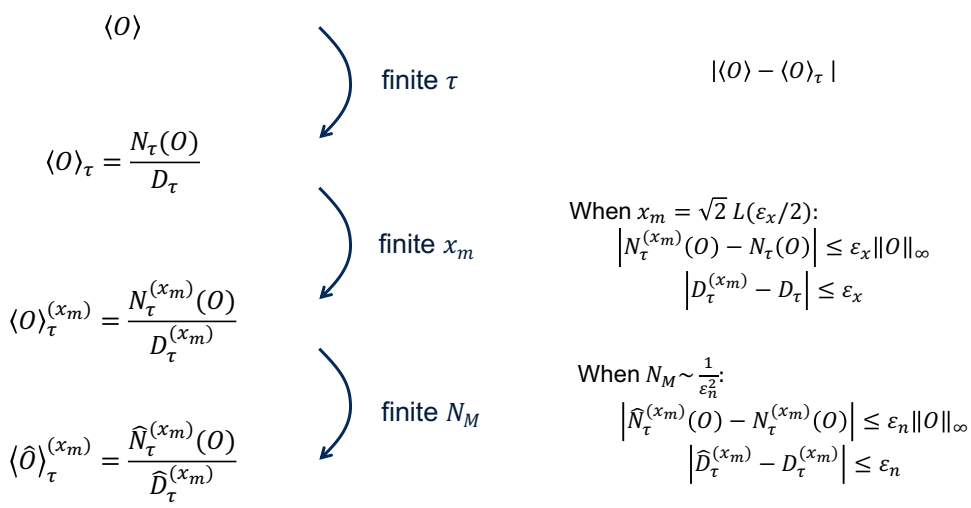


Figure 4.3: Summary of the error analysis of the observable estimation. Started from the ideal observation value $\langle O \rangle$, we sequentially study the effect of finite imaginary time τ, finite cutoff time x_m and finite sampling number N_M.

For the numerator and denominator, we have

$$\begin{aligned}\left|\hat{D}_\tau^{(x_m)} - D\right| &\leq \varepsilon_\tau^{(1)} + \varepsilon_x^{(1)} + \varepsilon_n^{(1)},\\ \left|\hat{N}_\tau^{(x_m)}(O) - N(O)\right| &\leq (\varepsilon_\tau^{(2)}\|O\|_\infty + \varepsilon_x^{(2)}\|O\|_\infty + \varepsilon_n^{(2)}\|O\|_1),\end{aligned} \tag{4.32}$$

In the appendix, the error is shown to have a bound as

$$|\langle\hat{O}\rangle_\tau^{(x_m)} - \langle O\rangle| \leq p_j^{-1}(\langle O\rangle + 1)(\varepsilon_\tau^{(1)} + \varepsilon_x^{(1)} + \varepsilon_n^{(1)}) + p_j^{-1}(\varepsilon_\tau^{(2)}\|O\|_\infty + \varepsilon_x^{(2)}\|O\|_\infty + \varepsilon_n^{(2)}\|O\|_1). \tag{4.33}$$

Based on the error dependence, I analyse the resource requirements (i.e., circuit depth and sample number) of the eigenstate property estimation. The workflow of the error analysis is summarised in Figure 4.2.

In order to simplify the discussion, I discuss the error of observable estimation with a known target eigenenergy. Denote the ideal (estimated) measurements with infinite (finite) τ, x_m, and N_M as $\langle O\rangle$ ($\langle\hat{O}\rangle_\tau^{(x_m)}$), the effect of the three factors on observable estimation is given as follows.

Theorem 2 (Accuracy of the observable estimation)**.** *Given constant $K > 0$, error $\varepsilon \in (0,1)$, finite imaginary time $\tau \geq \frac{2}{\Delta} g^{-1}(\varepsilon p_j/12)$, normalised cutoff time $x_m \geq \sqrt{2}L(\varepsilon p_j/12)$, and sample number $N_M \geq K/(\varepsilon p_j/6)^2$, the error between the expectation value estimation $\langle\hat{O}\rangle_\tau^{(x_m)}$ and the ideal expectation value $\langle O\rangle$ is bounded by $|\langle\hat{O}\rangle_\tau^{(x_m)} - \langle O\rangle| \leq \varepsilon(\|O\|_1 + 1)$ with a failure probability of $\delta = 4\exp(-K/8)$. Here $\|O\|_1 = \sum_l |o_l|$ where $\{o_l\}_l$ are the Pauli coefficients $O = \sum_l o_l P_l$, $\Delta = \min\{|E_{j-1} - E_j|, |E_{j+1} - E_j|\}$, and p_j is the overlap with the target eigenstate.*

Theorem 2 states that the asymptotic time complexity of the cooling algorithm is determined by $g^{-1}(\varepsilon)L(\varepsilon)$ of the cooling function. The valid cooling functions in Table 4.1 all satisfy $g^{-1}(\varepsilon) = \mathcal{O}(\log(1/\varepsilon))$. Therefore, we can choose τ to be $\mathcal{O}(\log(1/(p_j\varepsilon_\tau))\Delta^{-1})$, which is logarithmic in the inverse state overlap $1/p_j$ and inverse error $1/\varepsilon_\tau$, and linear to the inverse gap Δ^{-1}. For the cutoff x_m, it is shown that we have $L(\varepsilon) = \mathcal{O}(\mathrm{poly}(1/\varepsilon))$ for the triangle and exponential cooling functions, and we can achieve even better results $\mathcal{O}(\sqrt{\log(1/\varepsilon)})$ or $L(\varepsilon) = \mathcal{O}(\log(1/\varepsilon))$ for the Gaussian and secant hyperbolic cooling functions. Taking the Gaussian function as an example, we have $x_m = \mathcal{O}(\sqrt{\log(1/(\varepsilon p_j))})$. Note that the maximum real-time evolution is given by $t_m = \tau x_m$. As a result, we arrive at the following Proposition.

Proposition 3. *The time (circuit) t_m and sample N_M complexity for the Gaussian cooling function is*

$$t_m \sim \mathcal{O}\big(\Delta^{-1}\log(1/(\varepsilon p_j))\big), \quad N_M \sim \mathcal{O}(1/(\varepsilon p_j)^2). \tag{4.34}$$

The time or circuit complexity is logarithmic to $1/(p_j\varepsilon)$, which is exponentially better than that of quantum phase estimation or adiabatic state preparation, which is generally polynomial to $1/(p_j\varepsilon)$. The sample complexity in terms of p_j is slightly worse than that of quantum phase estimation, which is $\mathcal{O}(1/(p_j\varepsilon^2))$.

One can find that from Theorem 2, the efficiency of the cooling algorithm relies on two assumptions: a nonvanishing initial state overlap and a nonvanishing energy gap. These two assumptions are formalised in the following.

Assumption 1 (Nonvanishing state overlap assumption [29, 31, 141])**.** *For an n-qubit gapped Hamiltonian H with the ground state $|u_0\rangle$, we assume that it is feasible to prepare an initial state $|\psi_0\rangle$ satisfying a nonvanishing overlap of the target eigenstate $p_0 := |\langle\psi_0|u_0\rangle|^2$. We assume that the lower bound of the overlap satisfies $p_0 \geq \mathcal{O}(\mathrm{Poly}(\frac{1}{\mathrm{n}}))$.*

Assumption 2 (Gapped Hamiltonian assumption [29,31,141])**.** *For an n-qubit gapped Hamiltonian H with the ground state energy E_0, we assume that the energy gap $E_1 - E_0$ is lower-bounded by a known value Δ.*

Assumption 1 is reasonable in many practical scenarios, since even in many strongly-correlated quantum systems or chemical molecules Hamiltonians, the mean-field solution, like the Hartree-Fock state, still has a considerable overlap with the ground state. In practice, people usually apply the variational methods [130] or adiabatic methods [173] as a heuristic way to prepare an initial state with a large overlap p_0 with the ground state. Assumption 2 is also of practical relevance for the study of quantum many-body systems. For example, the ferromagnetic XXZ Heisenberg chain $H = J(\sum_i \sigma_i^x\sigma_{i+1}^x + \sigma_i^y\sigma_{i+1}^y + \Delta\sigma_i^z\sigma_{i+1}^z)$ with $J < 0$ has a finite energy gap if $\Delta > 1$ [174].

4.4.3 Comparison to prior works

Compared to prior works on eigenstate preparation and eigenenergy estimation, our method has the following advantages.

1. Only one ancillary qubit is needed in the whole algorithm. Furthermore, we do not require any quantum oracle, such as select-H for LCU or qubitisation methods, which could be difficult to implement[2].

2. The time complexity for ground-state energy estimation reaches the Heisenberg limit.

3. The time complexity for ground-state property estimation is logarithmic in the inverse of the precision requirement when we adopt the Gaussian cooling function.

To our knowledge, this is the first algorithm that encompasses all of the above advantages. In Table B.1 and Table B.2, our cooling method is compared with several typical quantum algorithms in the two primary tasks considered in this thesis, i.e., eigenenergy estimation and observable estimation. Here, we mainly focus on the algorithms without the usage of oracles

[2]The implementation of qubitisation is introduced in Section 2.4.3. The resource cost for implementing these quantum oracles (block encoding of a given Hamiltonian H) is discussed in Section 9.3.

for the block-encoding of a Hamiltonian. The latter will be discussed in Chapter 9. The Tables compare the maximal evolution time which characterises the complexity of quantum circuits for realising Hamiltonian simulation e^{-iHt}, the number of repetitions required to run the quantum circuit, and the total evolution time needed for estimating the ground state energy to within error ϵ. The complexity of estimating properties of excited states can be analysed in a similar fashion. Note that we simply compare with the conventional phase estimation method. Here we also use asymptotic notations, besides the usual $\mathcal{O}$ notation, to denote the complexity up to a polylogarithmic factors, similarly to that in Ref. [29].

Note that we can generalise Assumption 1 and Assumption 2 to the case of the estimation of jth eigenenergy E_j and the properties of jth eigenstate $|u_j\rangle$: it is also assumed that the initial state has a large overlap p_j with the jth eigenstate of the target Hamiltonian H; we also assume a known lower bound Δ on the energy gap $\min\{|E_j - E_{j-1}|, |E_j - E_{j+1}|\}$. In this case, our proposed algorithm can then be used for eigenenergy and eigenstate property estimation.

4.5 Numerical simulation

In this section, I show numerical implementation of the algorithmic cooling method for the anisotropic Heisenberg Hamiltonian,

$$H = J \sum_{i=1}^{7} \left(\sigma_i^x \sigma_{i+1}^x + \sigma_i^y \sigma_{i+1}^y + 2\sigma_i^z \sigma_{i+1}^z\right) + h \sum_{i=1}^{8} \sigma_i^z,$$

where $J = 1$ is the exchange coupling, σ_i^α $(\alpha = x, y, z)$ is the Pauli operator on the ith site, $h = 1$ is the strength of a uniform magnetic field in the z direction, and periodic boundary condition is imposed. The initial state is in the computation basis as $|01010101\rangle$, which is close to the ground state and has nonzero overlap with a relatively small number of low-energy eigenstates.

Under the dual phase representation of the unphysical Gaussian function, the eigenenergy E_i maximises the denominator $D_\tau(E)$ if the initial state has a nonvanishing spectral weight $|\langle\psi_0|u_i\rangle|^2$ on the eigenstate $|u_i\rangle$. We can thus determine the eigenenergy by searching for the peaks of $D_\tau(E)$ with the Gaussian cooling function. Both the imaginary time τ and the cutoff x_m are consistently adjusted by the target precision. I elaborate on how to determine the imaginary time τ and the cutoff x_m by taking the Gaussian function as an example. For a certain simulation accuracy ε, the imaginary time τ and the cutoff x_m satisfy $\tau \propto \sqrt{\log(2\Delta/\epsilon)}$ and $x_m = 2\sqrt{\log(2/\epsilon)}$, respectively. The reference line is calculated by the error due to a finite imaginary-time τ plus the theoretical maximum cutoff error $L(\epsilon)$. Using a finite imaginary time $\tau = 1.7$ and cutoff $x_m = 4.4$, and 10^5 number of samples for the integral, I show the energy spectra in Figure 4.4. The maximum error introduced is below 0.01, which aligns with our error analysis.

I then show the spectrum search with different total evolution times $T = \tau x_m$ with cutoff x_m and imaginary evolution time τ. Here, τ and x_m are calculated according to the target precision,

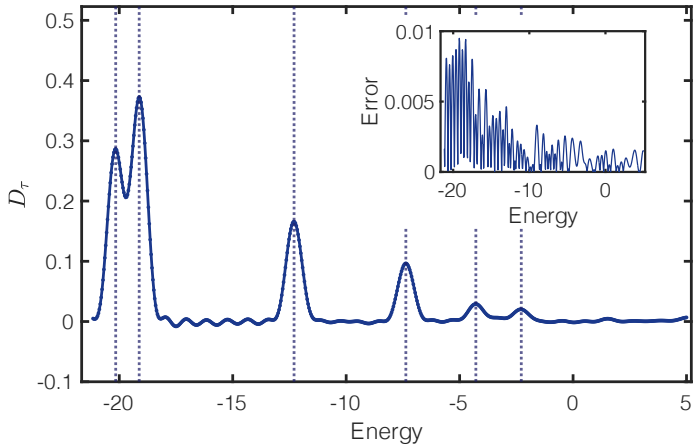


Figure 4.4: Energy spectra search of the 8-site anisotropic Heisenberg model using the Gaussian cooling function. The solid line shows the denominator of D_τ with $\tau = 1.7$ and cutoff $x_m = 4.4$. The dashed line shows the exact eigenenergy of the Heisenberg model. The figure inset shows the error of D_τ with respect to that under the exact cooling. $N_s = 10^5$ is set in the Monte Carlo sampling of the integral. Measurement shot noise is ignored.

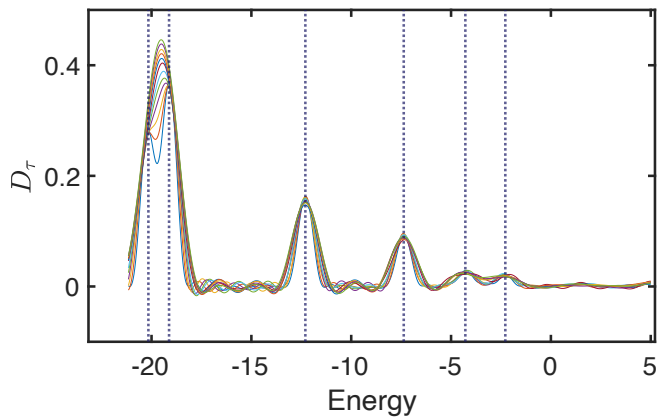


Figure 4.5: The spectrum search of the 8-site anisotropic Heisenberg model using the Gaussian cooling function with different evolution time. The solid line shows the denominator of D_τ with different τ from $\tau = 0.9$ to 1.7. The cutoff is calculated according to the target precision. The sampling number is set as $N_s = 10^5$ in the Monte Carlo sampling of the integral to keep consistent with Figure 4.4.

as discussed in the above sections. The maximum imaginary evolution time τ ranges from $\tau = 0.9$ to 1.7. As shown in Figure 4.5, with increasing imaginary time, we can distinguish the peaks that are close to each other, aligning with our theoretical analysis.

Finally, I show the error dependence of the eigenstate preparation, with a particular emphasis on the time complexity. Suppose we aim to find the second eigenstate $|u_2\rangle$, which has the largest overlap with the initial state. Given the associated eigenenergy E_2 found by the above method, one can analyse the error introduced from finite imaginary time and cutoff. Here, I focus on the state infidelity of the normalised state after cooling and the target ideal eigenstate, which can be

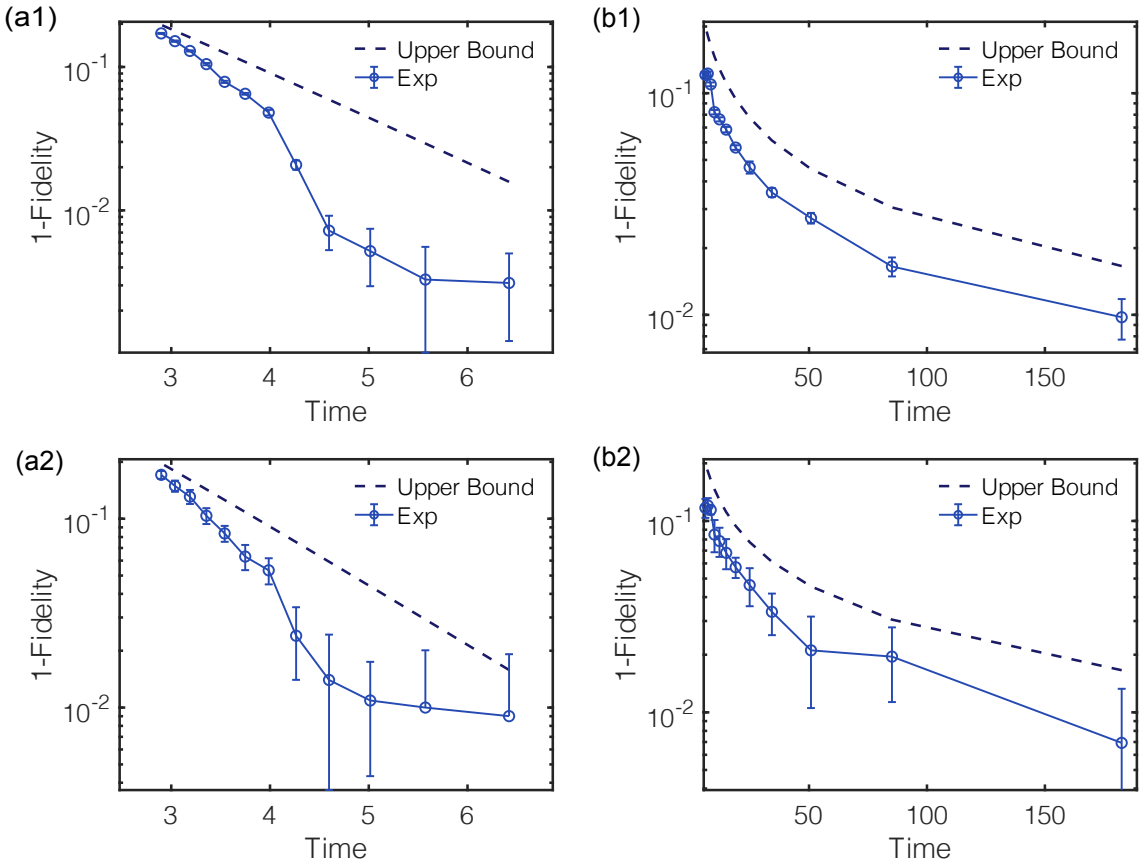


Figure 4.6: Error dependence of the total evolution time with (a1-a2) Gaussian and (b1-b2) exponential cooling functions. In (a1) and (b1), we study the infidelity versus maximal total time $t_m = \tau x_m$ and both the imaginary time τ and the cutoff x_m are adjusted. The measurement shot noise is ignored in (a1, b1). In (a2) and (b2), we further consider measurement shot noise, which introduces more instability of the infidelity. Nevertheless, we still observe the exponential and linear inverse decay of the infidelity for the Gaussian and exponential functions, respectively. The error bar is the standard deviation of D_τ error over 10 independent repetitions of the entire setup. 10^5 samples are used in the Monte Carlo sampling of the integral. Dashed line: theoretical upper bound of the infidelity assuming infinite samples; Dots with error bar: actual infidelity with a constant number of samples.

expressed as $\varepsilon = 1 - \langle O \rangle_\tau$ with $O = |u_2\rangle \langle u_2|$. I illustrate the error dependence on the imaginary time τ and the maximal total evolution time $t_m = \tau x_m$ with Gaussian and exponential cooling functions in Figure 4.6(a1-a2) and (b1-b2), respectively. I plot the theoretical upper bound of the infidelity in the dashed line in the figures (assuming finite τ, infinite x_m, and infinite samples N_M) and realistic infidelity with constant resources (a finite circuit depth and a finite number of measurements) using dots. I set the imaginary time τ and the cutoff according to precision in the worst case, similar to the spectrum analysis, and the pre-determined precision is plotted in the dashed line in the figures. In Figure 4.6(a1) and (b1), I consider finite τ and x_m with 10^5 samples for calculating the integral, but ignore statistical errors due to a finite number of

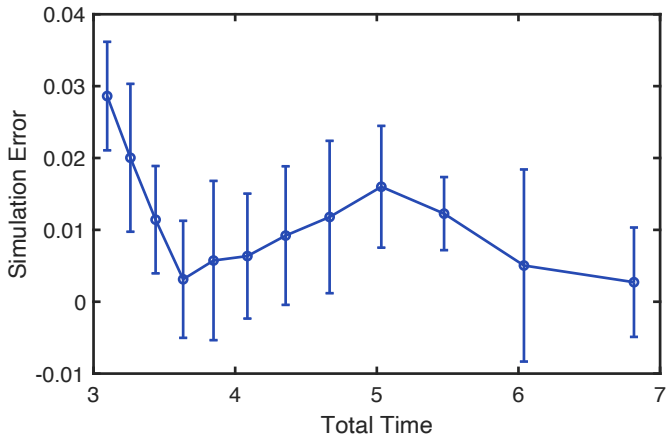


Figure 4.7: Simulation error during the cooling process. The error is compared with the ideal expectation value given by $\langle\psi_0|g(\tau(H-E_2))^\dagger Og(\tau(H-E_2))|\psi_0\rangle$. The simulation error is below 0.05 under evolution.

measurements. We can see the exponential scaling with respect to the total evolution time t_m for the Gaussian cooling function, while it asymptotically becomes $1/\varepsilon$ for the exponential cooling function. Figure 4.6(a2) and (b2) further show the error dependence when considering the statistical errors, where single-shot measurement outcomes are collected, and 10^5 samples are used.

Additionally, the simulation accuracy under the cooling process compared to the ideal cooling process is shown in Figure 4.7. We can see that the error is kept to a relatively small value during the cooling process, which verifies the effectiveness of our random sampling scheme. We can improve the simulation accuracy by increasing the Monte Carlo sampling of the integral.

Author contributions. This chapter is relevant to a preprint [93] and a manuscript under preparation [92]. Z.P., J.S., and X.Y. conceived the project. Z.P. and J.S. developed the theoretical framework of algorithmic cooling. J.S. carried out numerical simulation and resource estimation. X.Y. supervised the project.

Chapter 5

Dynamics simulation by adaptive product formulae

As discussed in Section 2.4, one major application of quantum computing is to simulate the time evolution of quantum many-body systems [6]. This is an essential component of quantum simulation, which not only allows us to study dynamical behaviours of quantum systems [7, 9, 12, 95, 175–178], but also supplements other quantum algorithms as a common subroutine (e.g. quantum phase estimation [28, 179]) for general tasks (e.g. finding energy spectra [80, 177, 180–184] or particle scattering [41, 42]).

In Section 2.4, I have introduced several solutions to Problem 1, from the initial simple Trotter-Suzuki product formula [139, 185] to the latest advanced approaches [37, 38, 144, 186–189], such as ones based on Taylor series [38] and quantum signal processing with qubitisation [32, 33]. In this chapter, I will focus on Problem 2 introduced in Section 2.4 and show the solution with adaptive product formula, as an extension to product formula introduced in Problem 2. Moreover, I will further show how this is built on the hybrid tensor networks established in Chapter 3, which extends its application to dynamics simulation.

This chapter is relevant to a published paper in collaboration with Zi-Jian Zhang, Xiao Yuan, and Man-Hong Yung [94] and a manuscript under preparation [91].

5.1 Introduction

Consider a Hamiltonian $H = \sum_{j=1}^{L} a_j P_j$ that is decomposed as a sum of tensor products of Pauli operators P_j with real coefficients a_j. The key idea of the product formula (PF) methods is to approximate the time evolution operator $e^{-iH\delta t}$ via a sequence of operators selected from the Hamiltonian evolution operation pool $\{e^{-iP_j\delta t'}\}$ such as the first-order Trotter-Suzuki formula

$$e^{-iH\delta t} \approx \prod_{j=1}^{L} e^{-ia_j P_j \delta t} + \mathcal{O}(\delta t^2).$$

Higher-order Trotter-Suzuki formulae are shown in Section 2.4. Theoretically, tighter bounds have been derived for special types of systems (e.g. lattice Hamiltonians with nearest-neighbour

interactions) [144, 189], verifying the empirical error estimates of the product formula methods. Numerically, it has been shown that product formula methods require entangling gates and T gates to orders of magnitude fewer than those more advanced methods (e.g. qubitisation) for a system of 100 qubits [37]. Since product formula methods are ancilla-free and are much easier to implement in experiments, they are favoured for simulating dynamics with near-term quantum devices.

Product formula methods approximate the time evolution operator (quantum channel), which works for arbitrary input states, and thus they may require an unnecessarily large number of gates when we only evolve a specific quantum state in practice. Recent numerical and theoretical studies have shown that product formulae with fixed or random input states only require a much shorter circuit [147, 148, 190]. On the one hand, this indicates an even greater degree of practicality of the product formula methods for fixed input states. On the other hand, since existing product formula methods do not exploit the input state information, whether we can further exploit such information to reduce the implementation complexity merits further study. As near-term quantum hardware has limited gate fidelity [8, 11, 25, 71, 74, 191–193], it is crucial to design Hamiltonian simulation algorithms with a shorter circuit, and hence higher calculation accuracy.

In this chapter, I will discuss a construction-based method to adaptively find the optimal product formula for evolving an unknown but fixed input quantum state. The task description is formulated in Problem 2 in Section 2.4. Instead of using the theoretical worst-case error bound of conventional product formula methods, I will introduce a measurable quantifier to describe the simulation error. This work considers different evolution operators at different times and constructs the optimal one by minimising the error quantifier. Since the error quantifier focuses on the quantum state instead of the entire evolution channel, we can thus adaptively obtain the state evolution circuit with a significantly reduced circuit depth than conventional approaches for the unitary evolution.

This chapter is organised as follows. In Section 5.2, I introduce the framework of adaptive product formulae. To stimulate the discussion, I begin with a single-step adaptive product formula. Then, I show a jointly-optimised adaptive product formula. In Section 5.3, I show how dynamics simulations can be scaled up using hybrid tensor networks. In Section 5.4, I discuss the relation between our work and prior works in this field, and provide some perspectives on dynamics simulation methods. The error analysis is presented in Section C.1.1.

5.2 Adaptive product formula

This section discusses how to adaptively find a short product formula for the Hamiltonian simulation from a fixed input quantum state. Consider the above time-independent Hamiltonian

$H = \sum_{j=1}^{L} a_j P_j$. Recall in Section 2.4 that conventional product formula approaches approximate the time evolution operator $e^{-iH\delta t}$ within a small time step δt as

$$\varepsilon = \|e^{-iH\delta t} - \prod_j e^{-iO_j\lambda_j\delta t}\|, \tag{5.1}$$

as stated in Problem 1. Here, ε is the approximation error, $O_j \in \{P_j\}$ are chosen from Pauli words in the Hamiltonian, λ_js are real coefficients, and $\|\cdot\|$ is the operator norm. By properly choosing $\{O_j\}$ with a given order, the approximation error ε can be suppressed to higher orders of δt (e.g. $\varepsilon = \mathcal{O}(\delta t^2)$) and we can accordingly simulate the whole evolution operator e^{-iHT} with a small error in the order of $\mathcal{O}(T\delta t)$. The error bound is a pessimistic estimation of the worst-case scenario, and several numerical and theoretical studies [147, 148, 190] have further shown that for Problem 2 the complexity could be much smaller. In the following, I show a solution to Problem 2 by adaptive product formulae.

5.2.1 Single-step adaptive product formula

In the single-step protocol, the exact time-evolved state $e^{-iH\delta t}|\Psi(t)\rangle$ is approximated by

$$|\Psi(t+\delta t)\rangle = \prod_j e^{-iO_j\lambda_j\delta t}|\Psi(t)\rangle, \tag{5.2}$$

where the tuple $\vec{O} = \{O_j\}$ at each step is chosen from the Hamiltonian and the coefficients $\vec{\lambda} = \{\lambda_j\}$ are real. The error for the approximation within each finite time step can be described by the Euclidean distance between the exact time-evolved state and the approximated state

$$\tilde{\varepsilon} = \|e^{-iH\delta t}|\Psi(t)\rangle - \prod_j e^{-iO_j\lambda_j\delta t}|\Psi(t)\rangle\|, \tag{5.3}$$

where $\||\psi\rangle\| = \sqrt{\langle\psi|\psi\rangle}$ represents the state vector norm or Frobenius norm.

Using Taylor expansion

$$e^{-iH\delta t} = \sum_{k=0}^{\infty} \frac{\delta t^k}{k!}(-iH)^k$$

and

$$\prod_j e^{-iO_j\lambda_j\delta t} = \sum_{k=0}^{\infty} \sum_{j_1,j_2,\ldots,j_k} \frac{(-i)^k\delta t^k}{k!} \prod_k (\lambda_{j_k} O_{j_k}),$$

the algorithmic error can be expressed by

$$\tilde{\varepsilon} = \sqrt{\Delta^2\delta t^2 + \mathcal{O}(\delta t^3)} \leq \Delta\delta t + \mathcal{O}(\delta t^{3/2}) \tag{5.4}$$

with the first-order error Δ being

$$\Delta^2 = \langle H^2\rangle + \sum_{jj'} A_{jj'}\lambda_j\lambda_{j'} - 2\sum_j C_j\lambda_j \tag{5.5}$$

Algorithm 1 Adaptive product formula (single step)

1. Set Δ_{cut} and $n = 1$. Input the initial state $|\Psi(t)\rangle$ and Hamiltonian $H = \sum_{j=1}^{L} a_j P_j$;
2. In the nth iteration, calculate $A, C, \vec{\lambda}$ and Δ of the new circuit $e^{-iP_k\lambda_k\delta t}\prod_{j=1}^{n-1} e^{-iO_j\lambda_j\delta t}|\Psi(t)\rangle$ for each Pauli word P_k in the Hamiltonian;
3. Set P_k that gives the lowest Δ in step 2 to be O_n;
4. If $\Delta > \Delta_{\text{cut}}$ for the new circuit, go to step 2 with $n = n + 1$. Else, stop and output $\vec{O}$, $\vec{\lambda}$ for approximation in Eq. (5.2).

Here, we have defined $A_{jj'} = \text{Re}\left(\langle\Psi(t)|O_jO_{j'}|\Psi(t)\rangle\right)$, $C_j = \text{Re}\left(\langle\Psi(t)|HO_j|\Psi(t)\rangle\right)$ and $\langle H^2\rangle = \langle\Psi(t)|H^2|\Psi(t)\rangle$. These values can be evaluated by quantum circuits [116, 131].

With a given $\vec{O}$, Δ^2 is a quadratic function of $\vec{\lambda}$, whose minimum can be obtained at the stationary point where $\frac{\partial}{\partial\lambda_j}\Delta = 0$ for all j. This is equivalent to $\sum_{j'} A_{jj'}\lambda_{j'} = C_j$, or $A\vec{\lambda} = C$ (in a vector form). Therefore, the coefficients $\vec{\lambda}$ that give the minimum of Δ^2 can be determined by solving a linear equation, either by applying the inverse matrix A^{-1} or by an iterative algorithm. The accuracy of time evolution at each time step can be bounded by error $\mathcal{O}(\Delta\delta t)$, and Δ (or Δ^2) serves as a handy measurable quantifier to estimate the quality of the time evolution with a choice of $\vec{O}$. Therefore, to construct a circuit (which could be represented by $\vec{O}$) that approximates the exact evolution with better quality, we can just add a new Pauli word to $\vec{O}$ and see how Δ will change. In the single-step strategy, we iterate over all the Pauli words in the Hamiltonian and add the Pauli word that gives the lowest Δ to $\vec{O}$. As the total error can be bounded by $\varepsilon_{\text{total}} \leq T\max\Delta$ when δt is small enough, we can set a threshold $\Delta_{\text{cut}} = \varepsilon_{\text{total}}/T$, and if $\Delta > \Delta_{\text{cut}}$ for the new circuit, we repeat the addition of new operators until $\Delta < \Delta_{\text{cut}}$. This adaptive circuit construction process can suppress the effect of the algorithmic error in the time evolution to any given $\varepsilon_{\text{total}}$ with proper Δ_{cut} and δt. The above method is summarised in Algorithm 1.

The performance of adaptive product formulae is guaranteed by Theorem 3

Theorem 3. *Algorithm 1 satisfies the following properties.*

(1) The error Δ strictly decreases at each iteration;

(2) Each Pauli word only needs to appear once;

(3) We can achieve an error $\Delta \leq \Delta_{\text{cut}}$ in at most L iterations for any $\Delta_{\text{cut}} \geq 0$.

The first property indicates that the strategy is always effective; The second property indicates that the strategy only requires applying each Pauli word once [1]; The third property states that Algorithm 1 will terminate in a finite number of steps. The above theorem thus guarantees the effectiveness and efficiency of the adaptive approach in the approximation of the time-evolved state in Eq. (5.2).

The proof of Theorem 3 was mainly carried out by my collaborator Zi-Jian Zhang. Here, I will merely give a rough idea of the proof and refer to a complete proof in [94]. The proof is

[1] This property only holds when we consider the first-order error Δ, which may fail for higher-order errors.

based on analysing the tangent space of product formulas. The first-order error is found to be the distance from the objective evolution direction to the tangent space of the current product formula. The next step is to find the condition for the error Δ to decrease when adding the new Pauli word to the product formula. The condition is proven to be fulfilled in our method.

5.2.2 Jointly optimised adaptive product formula

I first overview the jointly optimised adaptive product formula before moving on to the technical details. To evolve an initial state $|\Psi_0\rangle$ to $e^{-iHT}|\Psi_0\rangle = |\Psi(T)\rangle$, we can sequentially apply the single-step adaptive strategy, in a similar vein to conventional product formula methods. However, this might not be optimal since the circuit constructed independently at each step may still be redundant for the state evolution. To further reduce the circuit depth, a jointly-optimised adaptive strategy is introduced to produce a more compact quantum circuit.

Using an iterative description, it is assumed that the approximation of the time-evolved state at time t to be $|\Psi(t)\rangle = G(\vec{O}, \vec{\Lambda})|\Psi_0\rangle$, where $G(\vec{O}, \vec{\Lambda}) = \prod_j e^{-iO_j\Lambda_j}$ is the circuit that has been constructed at time t. Let us represent the list of Pauli words by $\vec{O} = \{O_j\}$ and the adjustable real coefficients (parameters) by $\vec{\Lambda} = \{\Lambda_j\}$. The single-step adaptive strategy considers the approximation of Eq. (5.2) as $G(\vec{O}', \vec{\lambda}'\delta t)G(\vec{O}, \vec{\Lambda})|\Psi_0\rangle$, in which the original circuit block $G(\vec{O}, \vec{\Lambda})$ is fixed and only the new added circuit block $G(\vec{O}', \vec{\lambda}'\delta t)$ (specifically $\vec{O}'$ and the associated $\vec{\lambda}'$) is optimised. However, since our objective is to obtain the time-evolved state from $|\Psi_0\rangle$, rather than a general unitary for the time evolution, we may only need a much shorter circuit to effectively implement the state evolution to any time t. In this section, I demonstrate how to optimise the parameters jointly and implement the state evolution up to a given simulation error.

A refined strategy (Algorithm 2) is to jointly optimise $G(\vec{O}', \vec{\lambda}'\delta t)$ and $G(\vec{O}, \vec{\Lambda})$. That is to consider the approximation

$$e^{-iH\delta t}|\Psi(t)\rangle \approx G(\vec{O}', \vec{\lambda}'\delta t)G(\vec{O}, \vec{\Lambda} + \vec{\lambda}\delta t)|\Psi_0\rangle\,, \tag{5.6}$$

where we also $\vec{\lambda}$, which corresponds to the variation of the parameters found in the previous steps. Specifically, consider the simulation error $\tilde{\varepsilon} = \|e^{-iH\delta t}|\Psi(t)\rangle - G(\vec{O}', \vec{\lambda}'\delta t)G(\vec{O}, \vec{\Lambda} + \vec{\lambda}\delta t)|\Psi_0\rangle\|$. The first-order error $\Delta = \lim_{\delta t\to 0}\tilde{\varepsilon}/\delta t$ has a similar form of Eq. (5.5). Given $\vec{O}', \vec{O}$ and $\vec{\Lambda}$, the optimal $\vec{\lambda}$ and $\vec{\lambda}'$ that minimise Δ could then be analytically obtained by solving a linear equation, and hence Δ can also be obtained. Below, I will show the analytical formula of Δ and detailed calculation of the optimal $\vec{\lambda}$ and $\vec{\lambda}'$).

Methods.— I will now expand on the methods employed in detail. In the jointly-optimised strategy, the exact time-evolved state $e^{-iH\delta t}G(\vec{O}_t, \vec{\Lambda}_t)|\Psi_0\rangle$ at time t is approximated by

$$G(\vec{O}'_t, \vec{\lambda}'_t\delta t)G(\vec{O}_t, \vec{\Lambda}_t + \vec{\lambda}_t\delta t)|\Psi_0\rangle) = G(\vec{O}_t \oplus \vec{O}'_t, (\vec{\Lambda}_t + \vec{\lambda}_t\delta t) \oplus (\vec{\lambda}'_t\delta t))|\Psi_0\rangle\,. \tag{5.7}$$

After this step at t, we will have $\vec{O}_{t+\delta t} = \vec{O}_t \oplus \vec{O}'_t$ and $\vec{\Lambda}_{t+\delta t} = (\vec{\Lambda}_t + \vec{\lambda}\delta t) \oplus (\vec{\lambda}'_t \delta t)$. The total error at time $t+\delta t$ for the approximation can be described by the Euclidean distance of the exact time-evolved state and the approximated state as $\varepsilon_{\text{total}} = D(e^{-iH(t+\delta t)} |\Psi_0\rangle, G(\vec{O}'_t, \vec{\lambda}'_t \delta t) G(\vec{O}_t, \vec{\Lambda}_t + \vec{\lambda}_t \delta t) |\Psi_0\rangle)$ where $D(\rho, \sigma) = \|\rho - \sigma\|$ represents the Euclidean distance. Using the triangle inequality, the total error is upper bounded by

$$\begin{aligned}\varepsilon_{\text{total}} \leq & D(e^{-iH(t+\delta t)} |\Psi_0\rangle, e^{-iH\delta t} G(\vec{O}_t, \vec{\Lambda}_t) |\Psi_0\rangle) \\ & + D(e^{-iH\delta t} G(\vec{O}_t, \vec{\Lambda}_t) |\Psi_0\rangle, G(\vec{O}'_t, \vec{\lambda}'_t \delta t) G(\vec{O}_t, \vec{\Lambda}_t + \vec{\lambda}_t \delta t) |\Psi_0\rangle) \\ \leq & \sum_{t \in \text{time steps}} D(e^{-iH\delta t} G(\vec{O}_t, \vec{\Lambda}_t) |\Psi_0\rangle, G(\vec{O}'_t, \vec{\lambda}'_t \delta t) G(\vec{O}_t, \vec{\Lambda}_t + \vec{\lambda}_t \delta t) |\Psi_0\rangle).\end{aligned} \tag{5.8}$$

Here, we used the invariance of distances under unitary transformation and the summation in the second line is made over all the t at which a time step is made. Denote the approximation error as

$$\tilde{\varepsilon}_t = D(e^{-iH\delta t} G(\vec{O}_t, \vec{\Lambda}_t) |\Psi_0\rangle, G(\vec{O}'_t, \vec{\lambda}'_t \delta t) G(\vec{O}_t, \vec{\Lambda}_t + \vec{\lambda}_t \delta t) |\Psi_0\rangle), \tag{5.9}$$

which characterises the error with the circuit $G(\vec{O}'_t, \vec{0}) G(\vec{O}_t, \vec{\Lambda}_t)$ at each time step.

In order to give a first-order approximation of the error $\tilde{\varepsilon}$ defined above, I will derive an equation for

$$\tilde{\varepsilon} = D(e^{-iH\delta t} G(\vec{O}, \vec{\Lambda}) |\Psi_0\rangle, G(\vec{O}, \vec{\Lambda} + \vec{\lambda}\delta t) |\Psi_0\rangle), \tag{5.10}$$

which will become $\tilde{\varepsilon}_t$ by setting $\vec{O} \leftarrow \vec{O}_t \oplus \vec{O}'_t$, $\vec{\Lambda} \leftarrow \vec{\Lambda}_t \oplus \vec{0}$ and $\vec{\lambda}\delta t \leftarrow \vec{\lambda}_t \delta t \oplus \vec{\lambda}'_t \delta t$.

The approximation error can be expanded as

$$\tilde{\varepsilon} = \sqrt{\Delta^2 \delta t^2 + \Delta_2^2 \delta t^3 + \mathcal{O}(\delta t^4)} \leq \Delta \delta t + \Delta_2 \delta t^{3/2} + \mathcal{O}(\delta t^2). \tag{5.11}$$

The total error can thus be bounded by

$$\varepsilon_{\text{total}} \leq \sum_{t \in \text{time steps}} \tilde{\varepsilon}_t \leq \Delta^{(\max)} T + \Delta_2^{(\max)} \sqrt{\delta t} T. \tag{5.12}$$

Proposition 4 shows the explicit form of the first-order error. An upper bound of Δ_2 is shown in the Section C.1.1.

Proposition 4 (First-order approximation error by adaptive product formula)**.** *The first-order approximation error of* $\tilde{\varepsilon} = D(e^{-iH\delta t} G(\vec{O}, \vec{\Lambda}) |\Psi_0\rangle, G(\vec{O}, \vec{\Lambda} + \vec{\lambda}\delta t) |\Psi_0\rangle)$ *is characterised by* Δ*, and we have*

$$\Delta^2 = \langle \psi_1 | \psi_1 \rangle = \langle H^2 \rangle + \sum_{jj'} A_{jj'} \lambda_j \lambda_{j'} - 2 \sum_j C_j \lambda_j, \tag{5.13}$$

where we denote

$$A_{jj'} = \text{Re}\left(\langle \Psi_0 | \partial_j G^\dagger(\vec{O}, \vec{\Lambda}) \partial_{j'} G(\vec{O}, \vec{\Lambda}) |\Psi_0\rangle\right), \quad C_j = \text{Im}\left(\langle \Psi_0 | \partial_j G^\dagger(\vec{O}, \vec{\Lambda}) H G(\vec{O}, \vec{\Lambda}) | \Psi_0 \rangle\right), \tag{5.14}$$

and

$$\langle H^2 \rangle = \langle \Psi_0 | G^\dagger(\vec{O}, \vec{\Lambda}) H^2 G(\vec{O}, \vec{\Lambda}) | \Psi_0 \rangle. \tag{5.15}$$

To minimise Δ, the coefficients $\vec{\lambda}$ can be determined by solving the linear equation as $\sum_{j'} A_{jj'}\lambda_{j'} = C_j$, similarly to that in the single-step strategy. Each term in Eq. (5.13) can be measured on a quantum circuit, as discussed in Section 2.2.3. In practice, due to a finite number of measurements or hardware noise, the estimates of A and C may deviate from the true values, which results in an error in the estimation of Δ. In this thesis, it is assumed that the true values of $A_{jj'}$ and C_j can be obtained. I refer to [94] for discussions on the shot noise due to a finite number of measurements.

One difference between the jointly-optimised strategy and the single-step strategy is that when old parameters are permitted to change, the Δ of the circuit may be lower than the cutoff Δ_{cut} by merely using the old circuit itself. In this case, adding new operators becomes unnecessary. Therefore, in the new jointly-optimised strategy, we only add new operators when $\Delta > \Delta_{\text{cut}}$, which largely reduces the quantum gates required for certain simulation accuracy.

Similar to the single-step strategy, the jointly-optimised strategy also constructs its circuit by iterating over all the Pauli words in the Hamiltonian. We calculate Δ for each Pauli word and add the Pauli word that gives the lowest Δ, $O^{(\text{min})}$, to the previous circuit. Δ is calculated by optimising over $\vec{\lambda} \oplus \vec{\lambda'}$ associated with the circuit $\vec{O} \oplus \vec{O'}$, instead of merely optimising over $\vec{\lambda'}$. $\oplus$ denotes the concatenation of the two vectors. This procedure is repeated until the first-order error Δ is less than a threshold.

A major advantage of the new strategy is that when the old parameters are allowed to change, we may only need to add much fewer new operators to ensure Δ is below the threshold. This is because we only add gates when $\Delta > \Delta_{\text{cut}}$; in the extreme case in which optimising over the existing parameters $\vec{\lambda}$ directly makes $\Delta \leq \Delta_{\text{cut}}$, no additional gates are needed to proceed the evolution. In addition, we arrange that every time the circuit is constructed, we add new operators until $\Delta \leq \Delta_{\text{cut}}/2$, so that we do not necessarily have to add new operators in a few following evolution steps until $\Delta > \Delta_{\text{cut}}$. The refined algorithm is summarised in Algorithm 2 and Figure 5.1.

Algorithm 2 Adaptive product formula

1. Set Δ_{cut}. Input the initial state $|\Psi_0\rangle$ and Hamiltonian $H = \sum_{j=1}^{L} a_j P_j$;
2. (Joint parameter optimisation) Calculate optimal $\vec{\lambda}$ and Δ of the current circuit $G(\vec{O}, \vec{\Lambda} + \vec{\lambda}\delta t)\,|\Psi_0\rangle$ at time t. If $\Delta > \Delta_{\text{cut}}$, go to the step 3. Otherwise, set $\vec{\Lambda} \to \vec{\Lambda} + \vec{\lambda}\delta t$ and continue step 2 with $t \to t + \delta t$; Terminate when $t = T$;
3. (Add new operators)

(a) For every Pauli word P_k in H, calculate Δ of the circuit $G(\vec{O} \oplus P_k, (\vec{\Lambda} + \vec{\lambda}\delta t) \oplus \lambda'\delta t)\,|\Psi_0\rangle$.

(b) With $O^{(\text{min})}$ being the Pauli word giving the smallest Δ in step (3a), add $O^{(\text{min})}$ to the end of the product formula as $\vec{O} \to \vec{O} \oplus O^{(\text{min})}$ and $\vec{\Lambda} \to \vec{\Lambda} \oplus 0$;

(c) If we have $\Delta \leq \Delta_{\text{cut}}/2$, stop adding operators and go to step 2. Otherwise, go to step (3a).

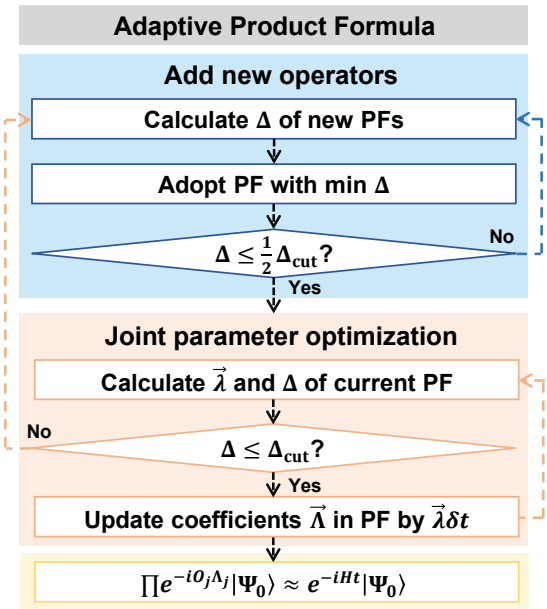


Figure 5.1: Scheme of the adaptive product formula (PF) method described in Algorithm 2. We adaptively construct the time evolution circuit of quantum states by learning the quantum gates and parameters that minimise the first-order error Δ at each step.

It is proven that the "Add new operator" procedure of Algorithm 2 also satisfies the same properties in Theorem 3. Compared to conventional product formula methods, which apply a deterministic sequence of gates, our method provides a circuit-growth strategy that optimises quantum circuits with a shallow depth.

5.3 Large-scale dynamics simulation with hybrid tensor networks

Simulation of quantum dynamical processes based on classical perturbative approaches or tensor networks in general scales exponentially with evolution time. Quantum computing provides an alternative solution, but the simulation task requires a number of physical or logical qubits no smaller than the problem size and a relatively deeper circuit. In this section, I propose a quantum dynamics simulation method with hybrid tensor networks that address these two challenges, making it compatible with the near-term quantum technology.

5.3.1 Stage setting

Consider a quantum many-body Hamiltonian

$$H = H^{\mathrm{loc}} + V^{\mathrm{int}}, \tag{5.16}$$

where $H^{\mathrm{loc}} = \sum_l H_l$ corresponds to a strong but local interaction with each H_l acting on the lth subsystem and V^{int} corresponds to a weak perturbation. In practice, one can always divide the

whole system into L subsystems, which will be discussed in Chapter 6 in detail. We can thus consider

$$V^{\text{int}} = \sum_j \alpha_j V_l^{(j)} \tag{5.17}$$

as the weak perturbation interaction between the subsystems with different interactions $V_l^{(j)}$ acting on the lth subsystem and the corresponding coefficients α_j. The local Hamiltonians and the perturbation interaction depend on the partitioning strategy of the subsystems. A time-independent Hamiltonian is assumed in the following discussion. However, our results also apply to general time-dependent Hamiltonians. Without loss of generality, we assume that every H_l could be decomposed as a linear combination of tensor product of Pauli operators and each V_j^{int} is a tensor product of Pauli operators.

The quantum state at time t is approximated by

$$|\Psi(t)\rangle = \sum_{i_1,\ldots,i_L} \alpha_{i_1,\ldots,i_L} |\psi_1^{i_1}(t)\rangle \otimes \cdots \otimes |\psi_L^{i_L}(t)\rangle \,, \tag{5.18}$$

where $\alpha_{i_1,\ldots,i_L}$ denotes the generalised tensor that characterises the correlation between the subsystems and $\{|\psi_l^{i_l}\rangle\}$ denotes the quantum states for each subsystem l. Each subsystem is represented by quantum states $\{|\psi_l^{i_l}\rangle\}$, and we may use either a quantum or classical tensor $\alpha^{i_1,i_2\ldots i_L}$ to describe the correlation between subsystems. If the entanglement of subsystems is not large, we can set the rank of $\alpha^{i_1,i_2\ldots i_L}$ to be a small number. For example, we can use the matrix product state $\alpha^{i_1,i_2\ldots i_L} = \text{Tr}[\alpha_1^{i_1}\alpha_2^{i_2}\ldots\alpha_L^{i_L}]$. Alternatively, we can use a quantum tensor to represent the correlation between subsystems as $\psi_{i_1,\ldots,i_L} = \langle i_1|\ldots\langle i_L|\tilde{\psi}\rangle$.

The time-evolved state from $|\Psi(t)\rangle$ based on the above hybrid quantum-classical tensor network can be approximated by

$$|\Psi(t+\delta t)\rangle = e^{-iH\delta t}|\Psi(t)\rangle \approx \sum_{i_1',\ldots,i_L'} \alpha_{i_1',\ldots,i_L'} |\psi_1^{i_1'}(t)\rangle \otimes \cdots \otimes |\psi_L^{i_L'}(t)\rangle \tag{5.19}$$

The evolution of state is thus mapped to the evolution of the tensors $\alpha_{\ldots}$ and $\psi_l^{i_l}$.

To measure the expectation value of $M = O_1 \otimes \cdots \otimes O_L$ with each O_l representing a local observable on the lth subsystem, we have

$$\langle\tilde{\psi}|M|\tilde{\psi}\rangle = \sum_{i_1\ldots i_L, i_1'\ldots i_L'} \bar{\alpha}^{i_1',\ldots i_L'} \alpha^{i_1,\ldots i_L} M_1^{i_1',i_1}\ldots M_L^{i_L',i_L}, \tag{5.20}$$

with

$$M_l^{i_l',i_l} = \langle\psi_l^{i_l'}(t)|O_l|\psi_l^{i_l}(t)\rangle \,. \tag{5.21}$$

The dimension of each index $i_1,\ldots i_L$ should be a small number similar to the bond dimension of MPS.

Here, quantum tensors are used to represent the local n-qubit correlation, and the classical rank-L tensor is used for global correlation among these L clusters of qubits. Consequently,

this kind of quantum-classical tensor could be suitable for systems, in which local correlation dominates over global correlation. Again, it is worth pointing out that the ansatz in Eq. (5.19) may not guarantee simulation accuracy for a general quantum system.

5.3.2 Results

In this section, I demonstrate how to update the parameters in the tensor networks to achieve certain simulation accuracy ε under time evolution. I further show a joint optimisation method that may avoid the exponential increase of the bond dimension.

5.3.3 Tensor evolution—single step

As a starting point, I first discuss the approximation error at single step with time step δt. Consider the evolution of a local subsystem by Trotterisation as

$$e^{-iH\delta t}\left|\Psi(t)\right\rangle = e^{-iV\delta t}\sum_{i_1,\ldots,i_L}\alpha_{i_1,\ldots,i_L}\bigotimes_l e^{-i\delta t H_l}\left|\psi_l^{i_l}(t)\right\rangle \equiv e^{-iV\delta t}\left|\Psi^{\mathrm{loc}}(t+\delta t)\right\rangle. \tag{5.22}$$

The error for the approximation of

$$e^{-iH\delta t}\left|\Psi(t)\right\rangle \approx \left|\Psi(t+\delta t)\right\rangle = \sum_{i'_1,\ldots,i'_L}\alpha_{i'_1,\ldots,i'_L}\left|\psi_1^{i'_1}(t)\right\rangle\otimes\cdots\otimes\left|\psi_L^{i'_L}(t)\right\rangle \tag{5.23}$$

within each finite time step can be described by the Euclidean distance of the exact time-evolved state and the approximated state as

$$\varepsilon = \|e^{-iV\delta t}\left|\Psi^{\mathrm{loc}}(t+\delta t)\right\rangle - \left|\Psi(t+\delta t)\right\rangle\|_2, \tag{5.24}$$

where $\|\left|\psi\right\rangle\|_2 = \sqrt{\langle\psi|\psi\rangle}$ represents the state vector norm.

The algorithmic error in a finite time step can be similarly expressed as

$$\delta\varepsilon = \sqrt{\Delta^2\delta t^2 + O(\delta t^3)} \leq \Delta\delta t + \mathcal{O}(\delta t^{3/2}) \tag{5.25}$$

with

$$\Delta^2 = \langle\delta\Psi_1(t+\delta t)|\delta\Psi_1(t+\delta t)\rangle, \quad \left|\delta\Psi_1(t+\delta t)\right\rangle = -\big(iV + \sum_j \dot{\alpha}_j\partial_j\big)\left|\Psi^{\mathrm{loc}}(t+\delta t)\right\rangle. \tag{5.26}$$

Here, the parameter set is denoted as $\vec{\alpha} = \{\alpha_{i_1,\ldots,i_L}\}$, the jth component of $\vec{\alpha}$ as α_j, and the differential operator on the jth component of $\vec{\alpha}$ as $\partial_j \equiv \partial/\partial\alpha_j$ in short. The variation of the tensor is represented by $\delta\alpha_{i_1,\ldots,i_L} = \sum_j \dot{\alpha}_j\delta t$. Note that this mathematically holds for the quantum tensor $\alpha_{\ldots}$. I remark that we can adaptively adjust the parameters in the tensors at each time t. For instance, the parameters can be adjusted as $\vec{\alpha} \leftarrow \vec{\alpha}\oplus\vec{\alpha}'$ from t to $t+\delta t$.

The first-order error Δ can be expressed as

$$\Delta^2 = \langle V^2\rangle + \sum_{jj'} A_{jj'}\dot{\alpha}_j\dot{\alpha}_{j'} - 2\sum_j C_j\dot{\alpha}_j, \tag{5.27}$$

where we denote $\langle V^2\rangle = \langle\Psi^{\text{loc}}(t+\delta t)|\, V^2\, |\Psi^{\text{loc}}(t+\delta t)\rangle$, $A_{jj'} = \text{Re}\left(\partial_j \langle\Psi^{\text{loc}}(t+\delta t)|\, \partial_{j'}\, |\Psi^{\text{loc}}(t+\delta t)\rangle\right)$, and $C_j = \text{Im}\left(\partial_j \langle\Psi^{\text{loc}}(t+\delta t)|\, H\, |\Psi^{\text{loc}}(t+\delta t)\rangle\right)$.

Note that the first-order algorithmic error is a quadratic function with respect to the derivative of the parameters $\dot{\alpha}_j$, whose minimum is obtained at $A(\dot{\vec{\alpha}}\oplus\dot{\vec{\alpha}}') = C$, and thus the coefficients can be similarly determined by solving the linear equation $\sum_{j'} A_{jj'}\alpha_{j'} = C_j$. Each term in Eq. (5.27) can be measured on a quantum circuit.

5.3.4 Tensor evolution—full dynamics

I will now discuss a jointly optimisation method that circumvents the exponential increase of bond dimension. I show how a circuit that realises the evolution from a given quantum state can be adaptively determined.

To initiate the discussion, I first revisit the single-step evolution from time t to $t+\delta t$. We find that the evolution under the non-local interaction operator $e^{-iH\delta t}$ is able to be realised by the evolution of the tensors. Mathematically, the evolution can be represented by

$$\begin{aligned} e^{-iH\delta t}\,|\Psi(t)\rangle &= \sum_{i_1,\ldots,i_L} e^{-iH\delta t}\alpha_{i_1,\ldots,i_L}(t)\bigotimes_l |\psi_l^{i_l}(t)\rangle \\ &\approx \sum_{i'_1,\ldots,i'_L} \alpha_{i'_1,\ldots,i'_L}(t+\delta t)\,|\psi_1^{i'_1}(t+\delta t)\rangle\otimes\cdots\otimes|\psi_L^{i'_L}(t+\delta t)\rangle\,. \end{aligned} \tag{5.28}$$

Note that compared to directly mimicking the local operator $e^{-iH_l t}$, the state is approximated by updating the tensors of the local quantum tensors $\psi_l^{i'_l}(t+\delta t) \leftarrow \psi_l^{i_l}(t)$ and the upper layer quantum tensors $\alpha_{i'_1,\ldots,i'_L}(t+\delta t) \leftarrow \alpha_{i_1,\ldots,i_L}(t)$ instead. Now, the evolution of the state is mapped to the evolution of the parameters in the tensors. Denote the total parameters as $\vec{\Lambda}$, the variation of parameters as $\delta\Lambda = \lambda\delta t$, and represent the state as $|\Psi(t)\rangle \equiv T(\vec{\Lambda}(t))\,|\Psi_0\rangle$. Note that we could adjust the tensors at each time and this could be described by $\vec{\Lambda}(t+\delta t) \leftarrow \vec{\Lambda}(t)\oplus\vec{\Lambda}'(t)$ where we add the additional tensors with parameters Λ' at time t.

The error for the approximation during the evolution can be described by the Euclidean distance of the exact time-evolved state and the approximated state as

$$\varepsilon = \|e^{-iHt}\,|\Psi_0\rangle - \sum_{i_1,\ldots,i_L}\alpha_{i_1,\ldots,i_L}\,|\psi_1^{i_1}(t)\rangle\otimes\cdots\otimes|\psi_L^{i_L}(t)\rangle\,\|_2 \tag{5.29}$$

where, $\|\,|\psi\rangle\,\|_2 = \sqrt{\langle\psi|\psi\rangle}$ represents the Frobenius norm. The exact time-evolved state $e^{-iH\delta t}\,|\Psi(t)\rangle$ is approximated by $T(\vec{\Lambda}(t+\delta t))\,|\Psi_0\rangle$. The total error at time $t+\delta t$ for the approximation can be described by the Euclidean distance of the exact time-evolved state and the approximated state as $\tilde{\varepsilon} = D(e^{-iH\delta t}\,|\Psi(t)\rangle\,, T(\vec{\Lambda}(t+\delta t))\,|\Psi_0\rangle)$ where $D(\rho,\sigma) = \|\rho-\sigma\|_2$ represents the Euclidean distance defined above. Using the triangle inequality, the total error is upper bounded by

$$\tilde{\varepsilon} \le \sum_t D(e^{-iH\delta t}T(\vec{\Lambda}(t))\,|\Psi_0\rangle\,, T(\vec{\Lambda}(t+\delta t))\,|\Psi_0\rangle). \tag{5.30}$$

Here distance invariance under the unitary transformation is used and the summation is made over all the time steps.

Denote $\delta\varepsilon = D(e^{-iH\delta t}T(\vec{\Lambda}(t))\,|\Psi_0\rangle\,, T(\vec{\Lambda}(t+\delta t))\,|\Psi_0\rangle)$, which characterises the approximation error with the tensor T at a time step. Denote the jth component in $\vec{\Lambda}\oplus\vec{\Lambda}'$ as Λ_j and denote the differential operator on the jth component of $\vec{\Lambda}\oplus\vec{\Lambda}'$ as $\partial_j \equiv \partial/\partial\Lambda_j$. The algorithmic error in a finite time step can be thus expressed as

$$\delta\varepsilon = \sqrt{\Delta^2\delta t^2 + \Delta_2^2\delta t^3 + O(\delta t^4)} \leq \Delta\delta t + \Delta_2\delta t^{3/2} + \mathcal{O}(\delta t^2) \tag{5.31}$$

with

$$\Delta^2 = \langle\delta\Psi_1(\vec{O},\vec{\lambda})|\delta\Psi_1(\vec{O},\vec{\lambda})\rangle\,, \Delta_2^2 = 2\mathrm{Re}\left(\langle\delta\Psi_2(\vec{O},\vec{\lambda})|\delta\Psi_1(\vec{O},\vec{\lambda})\rangle\right), \tag{5.32}$$

and

$$\begin{aligned} |\delta\Psi_1(\vec{O},\vec{\lambda})\rangle &= -\big(iH + \sum_j \lambda_j\partial_j\big)T(\vec{\Lambda})\,|\Psi_0\rangle \\ |\delta\Psi_2(\vec{O},\vec{\lambda})\rangle &= -\frac{1}{2}\big(H^2 + \sum_{jj'}\lambda_j\lambda_{j'}\partial^2_{jj'}\big)T(\vec{\Lambda})\,|\Psi_0\rangle\,. \end{aligned} \tag{5.33}$$

The first-order order error Δ can be expressed as

$$\Delta^2 = \langle H^2\rangle + \sum_{jj'} A_{jj'}\lambda_j\lambda_{j'} - 2\sum_j C_j\lambda_j, \tag{5.34}$$

where we denote $A_{jj'} = \mathrm{Re}\left(\langle\Psi_0|\,\partial_j T^\dagger\partial_{j'}T\,|\Psi_0\rangle\right)$,$C_j = \mathrm{Im}\left(\langle\Psi_0|\partial_j T^\dagger HT|\Psi_0\rangle\right)$, and $\langle H^2\rangle = \langle\Psi_0|T^\dagger H^2 T|\Psi_0\rangle$. Note that first-order algorithmic errors is a quadratic function with respect to $\vec{\lambda}\oplus\vec{\lambda}'$, whose minimum obtains at $A(\vec{\lambda}\oplus\vec{\lambda}') = C$, and thus the coefficients can be similarly determined by solving the linear equation as $\sum_{j'} A_{jj'}\lambda_{j'} = C_j$.

We can first choose the tensors α' that give the minimal Δ^2, and adaptively adjust the tensors by adding the corresponding tensor α' under evolution. The procedure is very similar to Algorithm 2.

With our method, one can represent a Ln-qubit system by using an $\mathcal{O}(n)$-qubit system, and emulate large-system dynamics by carrying out operations on a small quantum processor. The success of this method relies on the fact that the entanglement of the system does not grow drastically under time evolution.

5.4 Discussion

5.4.1 Comparison to other works

Compared to the first-order Trotter method, where O_j and Λ_j are set as the corresponding Pauli word P_j and weight a_j in the Hamiltonian, respectively, our method provides a systematic circuit-growth strategy that optimises the quantum circuits with much smaller gate counts. In the original paper [94], the adaptive product formula (PF) method is numerically benchmarked for molecular systems and spin systems in both the dynamic and static problems. Our numerical

simulation shows a significant improvement of the gate count saving in practical computation. Specifically, the number of CNOT gates are significantly reduced by two orders compared to the first-order Trotter-Suzuki product formula. I am indebted to my collaborator Zi-Jian Zhang for performing the numerical simulation, and I refer the reader to [94] for the numerical simulation results.

Meanwhile, variational methods (see Section 2.2.3) have been widely used to simulate quantum dynamics, which use a shallow quantum circuit than that in conventional Trotterisation algorithms. One major challenge in variational quantum simulation is to design an appropriate circuit ansatz to represent the manifold that the target states live in, and in general it is difficult to ensure simulation accuracy. In contrast to a pre-determined and fixed circuit ansatz in variational quantum simulation, the circuit under time evolution is optimised by tracking the accuracy of the circuit at each time, which in turn enables an adaptive circuit construction with guarantees on the simulation errors in the time evolution. In this chapter, I mainly focus on real-time dynamics, and it is left to future work on the subject to extend the results to imaginary-time evolution, which could potentially be used to find energy spectra of quantum many-body systems, similarly to Refs. [177, 178]. Several adaptive variational quantum algorithms have been proposed for finding the energy spectra of the quantum many-body problems [194, 195], and our results using the quantum Krylov algorithm can be compared against these (see [94]).

It is also worth comparing our results with a parallel work, termed AVQDS, proposed by Yao *et al.* [196]. AVQDS is a work of considerable interest, with points of commonality between the authors' approach and our own. AVQDS shares a similar idea to Algorithm 2 in this chapter, i.e., the jointly-optimised algorithm, which adaptively constructs a quantum circuit for real-time evolution. The approach outlined in this chapter is, however, based on the theory of product formula and has guaranteed simulation accuracy as its primary consideration. A direct result of this difference is that AVQDS permits a wider choice of operator pools, whereas our work only allows the use of operators from the decomposition of the Hamiltonian. This feature of our method guarantees that the circuit construction in adaptive PF is valid, i.e., Δ can be reduced to any $\Delta_{\text{cut}} > 0$ and thus objective simulation accuracy can be achieved. This is an important distinction as improper pools may result in errors that cannot be reduced by circuit constructions. It can be shown that when the length of the longest Pauli word in the pool is smaller than $\frac{2}{3}$ of the length of the longest Pauli word in the Hamiltonian, the algorithm cannot always suppress simulation error below an arbitrarily low level. Intuitively, this is because the short Pauli words cannot cover the evolution generated by long Pauli words. More rigorous and detailed discussions can be found in a future work [197].

Finally, I would like to remark that while two algorithms - the 'single-step' and 'jointly-optimised' algorithms - are introduced in order to help the readers comprehend the spirit of our

methods in a more natural way, the single-step adaptive PF has its own advantage. The single-step adaptive PF does not require computing the coefficient shift in the product formula that has been added to the circuit. This makes the number of circuits that are needed to evaluate in the single-step algorithm independent of the number of steps that have been completed. It is expected that in the long-term when applying quantum gates is both relatively easy and fast, the single-step adaptive PF may be more desirable than the jointly-optimised one.

5.4.2 Resource cost

Compared to variational quantum algorithms, our adaptive strategy requires measuring an additional expectation value of the operator H^2, which can be taken as the resource cost for the adaptive method. Our method could be inefficient when simulating arbitrary many-body dynamics. Nonetheless, our method may be practical in quantum systems with short-range interactions, such as the Ising or Heisenberg model. Moreover, recent works [74,87,98,198–201] have proposed methodologies to efficiently measure Hamiltonians and their higher-order moments, such as classical shadows, Pauli group approaches, etc, which can be directly incorporated into our method. More specifically, we find tensor product basis sets, and the elements in each set qubit-wisely commute. The number of measurements can be reduced significantly by exploiting the measurement compatibility of the Pauli terms of the observables. In the extreme case, we only need to measure once if all the Pauli terms of the observable are compatible. While finding the optimal Pauli sets is equivalent to the minimum clique cover problem, which is NP-hard, several heuristic methods have been put forward to solve the problem, such as largest degree first (LDF) grouping [74, 198]. An efficient measurement strategy will be discussed in detail in Chapter 10. Moreover, the measurements and the optimisations at each iteration are fully parallelable and hence our method could have further speedup with the addition of more quantum devices.

Author contributions. Section 5.2 is relevant to a published work [94]. In this work, Z.Z., and J.S. initiated the project of adaptive product formulae. X.Y. and M.Y. supervised the project. J.S. and X.Y. conceived the idea of adaptive product formulae. J.S. developed the framework of single-step and jointly-optimised adaptive product formulae with input from Z.Z. and X.Y.. Z.Z. carried out numerical simulation and verification of our algorithms. Z.Z. proved the main theorem in [94] with input from J.S.. J.S. and Z.Z. analysed the effect of shot noise in [94].

Section 5.3 is relevant to a manuscript under preparation [91]. J.S. is responsible for the results in Section 5.3.

Chapter 6

A perturbative approach to simulating interacting dynamics

In previous chapters, I initially demonstrated how the static and dynamic problems posed in Section 1.2 can be formulated in a general framework, and then show the solutions in the Chapter 4 and Chapter 5. In this chapter, I will focus on a generic interacting system, which can be divided by the local part and the interacting part as

$$H = H^{\text{loc}} + V^{\text{int}}. \tag{6.1}$$

Here H^{loc} corresponds to the strong but local interaction and V^{int} corresponds to weak perturbations. In this chapter, I introduce a perturbative approach to simulating quantum many-body dynamics, with a generic system Hamiltonian characterised by Eq. (6.1).

This chapter is relevant to a work published in collaboration with Suguru Endo, Huiping Lin, Patrick Hayden, Xiao Yuan, and Vlatko Vedral [95].

6.1 Motivation

I start the discussion with a generic quantum many-body Hamiltonian characterised by Eq. (6.1). In practice, we can always divide the full quantum system into L subsystems, according to their topological structures or degrees of freedom. For instance, a lattice Hamiltonian and a molecular system with virtual and active orbitals or in a clustered structure (see [154, 202, 203]) provide natural partitioning strategies for the subsystems, similarly to that in perturbation theory. As such, we could consider

$$H^{\text{loc}} = \sum_l H_l \tag{6.2}$$

as the local Hamiltonians with each H_l acting on the lth subsystem, and

$$V^{\text{int}} = \sum_j \lambda_j V_j^{\text{int}} \tag{6.3}$$

as the weak perturbation interaction between the subsystems with different interactions V_j^{int} and real coefficients λ_i. Without loss of generality, V_j^{int} may be taken to be a product of Pauli operators for efficient implementation on a quantum computer. It is worth mentioning that the local Hamiltonians and the perturbation interactions depend on the partitioning strategy of the subsystems. Partitioning strategies for several typical physical systems can be found in Section 8.1.

To simulate the dynamics of $U(t) = e^{-iHt}$, one possible solution is to leverage classical methods that have been developed to solve quantum many-body problems, wherein the most successful one is perturbation theory. This method divides the Hamiltonian into a major but easily solved component and a weak but potentially complicated counterpart, so that the full dynamics can be expressed as a series expansion [159, 202, 204–207].

A representative perturbation treatment via Dyson series expansion is expressed as

$$U(t) = 1 - i \int_{t_0}^{t} dt_1 e^{iH^{\text{loc}}(t_1 - t_0)} V^{\text{int}} e^{-iH^{\text{loc}}(t_1 - t_0)} + \dots \tag{6.4}$$

The time evolution becomes a linear combination of different trajectories, where each one undergoes local Hamiltonian evolution $e^{iH_l t}$ with interactions V^{int} uniformly inserted during the evolution time. When the local Hamiltonians $\{H_l\}$ are solvable, one can further represent the expansion graphically, such as via Feynman diagrams, and compute expectation values of the time evolved state with different graphs corresponding to different expansion terms.

A major limitation of classical perturbation theory is the assumption of the simple, and hence solvable, local Hamiltonians, which fails when the local Hamiltonians $\{H_l\}$ become strongly correlated, as can occur in some realistic systems. In this case, it becomes hard to realise the local evolution $e^{iH_l t}$. A suggested solution to this could be to consider a different partition, where the complex part of local Hamiltonians is assigned to the interaction. However, this in turn may result in strong interactions, causing the expansion to become inefficient or divergent. Additionally, computing the higher-order expansions of general quantum many-body problems could be challenging, which further limits the use of perturbation theory. Regardless, even when there is no interaction under certain partitioning strategy with $V^{\text{int}} = 0$, no classical methods exist that can efficiently simulate the dynamics of general Hamiltonian $H^{\text{loc}} = \sum_l H_l$, otherwise the computational class of bounded-error quantum polynomial time collapses.

To circumvent this conundrum, the leveraging of quantum computing power to simulate the local Hamiltonians is proposed. We have termed this method perturbative quantum simulation (PQS), which uses a quantum computer to directly simulate the major component while perturbatively approximates the weak interaction component. Since there is no assumption on the size or interactions of the major component, PQS potentially goes beyond the conventional perturbative approach, conceivably being able to simulate classically intractable systems, such as large systems with weak inter-subsystem interactions or intermediate systems with general

interactions. An explicit algorithm mimicking Dyson series expansion is proposed. Its optimality in relative to more general theories is proven. Although PQS may not be applicable to arbitrary problems, it possesses a major advantage over conventional quantum simulation algorithms, in that its perturbative treatment of the weak component reduces quantum resources. Remarkably, PQS runs a shallow circuit on a smaller number of qubits, making it more noise-robust and thus useful in benchmarking large quantum devices with smaller ones.

6.2 Perturbative quantum simulation

6.2.1 Method description — discrete time

A time-independent Hamiltonian is assumed in the following discussion; however, our results apply to general time-dependent Hamiltonians. Without loss of generality, I assume that every H_l could be decomposed as a linear combination of tensor product of Pauli operators and each V_j^{int} is a tensor product of Pauli operators.

To do so, I first introduce the concept of local generalised quantum operations,

$$\Phi(\rho) = \text{Tr}_E \left[\mathbf{U} \left(\rho \otimes |\mathbf{0}\rangle \langle \mathbf{0}|_E \right) \mathbf{V}^\dagger \right], \tag{6.5}$$

where we denote ancillary qubits $|\mathbf{0}\rangle \langle \mathbf{0}|_E = |0\rangle \langle 0|_{E_1} \otimes \cdots \otimes |0\rangle \langle 0|_{E_L}$ and unitary operators $\mathbf{U} = U_{1E_1} \otimes \cdots \otimes U_{LE_L}$ and $\mathbf{V} = V_{1E_1} \otimes \cdots \otimes V_{LE_L}$. The notation U_{jE_j} and V_{jE_j} indicates that the operators act only on subsystem j and the jth ancillary qubits. While the operation $\Phi(\rho)$ is non-physical in general, it can be realised effectively using local operations and post-processing as shown in Figure 6.1(b). Note that $\Phi(\rho)$ reduces to local quantum channels when $\mathbf{U} = \mathbf{V}$, as illustrated in Figure 6.1(b1). The properties of this generalised quantum operation were shown in Section 2.5.1.

The key idea of PQS is to decompose the joint evolution into a set of generalised quantum operations, which separately act on each subsystem. I first describe the algorithm assuming discretised time, and then demonstrate how to take the limit of infinitesimal time steps. We aim to simulate the time evolution governed by the Hamiltonian H with time t,

$$\mathcal{U}(T)[\rho] = U(T)\rho U(T)^\dagger, \tag{6.6}$$

where $U(t) = e^{-iHt}$. Considering discrete time δt, we have

$$\mathcal{U}(T) = \prod_{i=1}^{T/\delta t} \mathcal{U}(\delta t) = \prod_{i=1}^{T/\delta t} \left[\mathcal{V}^{\text{int}}(\delta t) \circ \bigotimes_{l=1} \mathcal{U}_l(\delta t) \right] + \mathcal{O}(T\delta t), \tag{6.7}$$

where $\mathcal{U}_l(t) = U_l(t)\rho U_l(t)^\dagger$ with $U_l(t) = e^{-iH_l t}$ and $\mathcal{V}^{\text{int}}(t) = V^{\text{int}}(t)\rho V^{\text{int}}(t)^\dagger$ with $V^{\text{int}}(t) = e^{-iV^{\text{int}}t}$. Here $\mathcal{O}(T\delta t)$ corresponds to the Trotter error, which vanishes when taking the limit of $\delta t \to 0$. Note that the evolution consists of local evolution $\mathcal{U}_l(\delta t)$ on the lth subsystem term and the joint evolution $\mathcal{V}^{\text{int}}(\delta t)$.

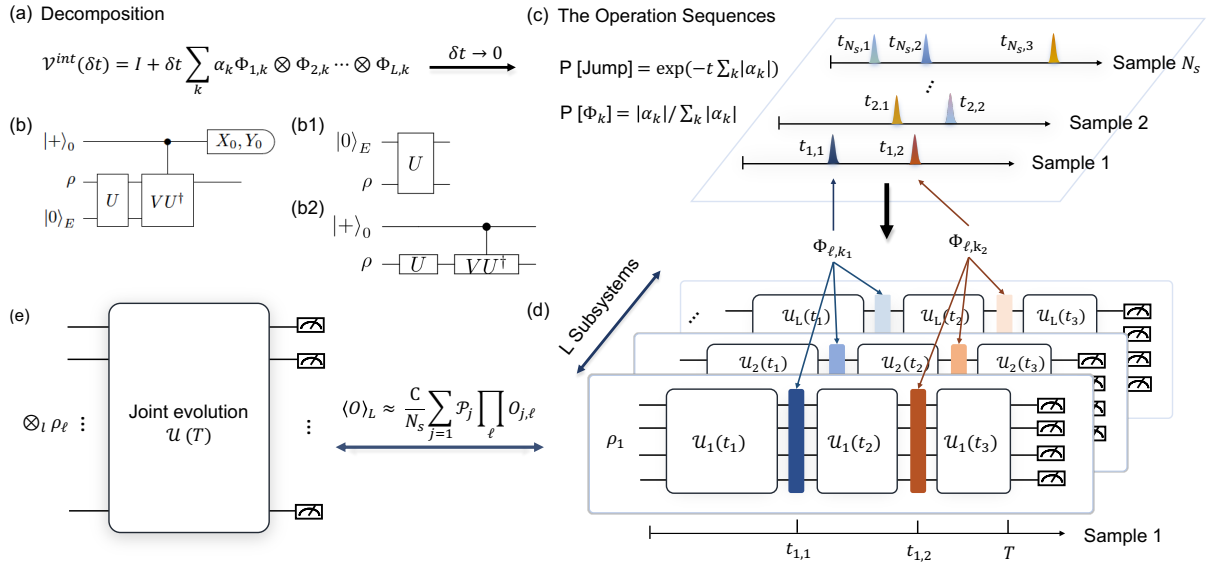


Figure 6.1: Schematic diagram of the perturbative quantum simulation algorithm. (a) The decomposition of interactions into local generalised quantum operations. (b) The implementation of a generalised quantum operation $\Phi(\rho) = \mathrm{Tr}_E[\mathbf{U}(\rho \otimes |0\rangle\langle 0|_E)\mathbf{V}^\dagger]$ using quantum circuits, which reduces to a quantum channel when $\mathbf{U} = \mathbf{V}$ in (b1) and unitary operations $\Phi(\rho) = U\rho V^\dagger$ when there is no ancilla in (b2). (c) We can equivalently realise the discretised scheme with $\delta t \to 0$. The operation sequences for N_s samples are predetermined provided the decomposition. We either continuously apply the local time evolution or randomly apply a generalised quantum operation. The time to apply the operations is determined by the probability P[Jump] and the kth operation is applied with probability $\mathrm{P}[\Phi_k]$. The scheme in (c) does not assume time discretisation and the number of generalised quantum operations that applies during the evolution scales as $\mathcal{O}(\sum_k |\alpha_k| T)$. (d) Schematic of the simulation process for a single sample as an example. For the lth subsystem, we evolve the state under local operations $\mathcal{U}_l$ and apply operations Φ_{l,k_1} and Φ_{l,k_2} at time $t_{1,1}$ and $t_{1,2}$, respectively. We measure the output states with a product observable O and obtain the outcomes $O_{j,l}$ for the jth sample. The process is repeated for N_s times. For any product input state, the expectation value of observable O under the joint evolution $\mathcal{U}$ in (e) can be unbiasedly approximated by $\langle O \rangle_L \approx \frac{C}{N_s} \sum_{j=1}^{N_s} \mathcal{P}_j \prod_l O_{j,l}$ with the overhead C and phase $\mathcal{P}_j$ determined by the decomposition.

The next step is to decompose the joint non-local operation $\mathcal{V}^{\mathrm{int}}(\delta t)$ into local operations that separately act on the subsystems. In particular, we can consider a set of generalised quantum operations as

$$\Phi_k(\rho) = \mathrm{Tr}_E\left[\mathbf{U}_k\left(\rho \otimes |\mathbf{0}\rangle\langle \mathbf{0}|_E\right)\mathbf{V}_k\right] \tag{6.8}$$

where we denote $\mathbf{U}_k = U_{1E_1,k} \otimes U_{2E_2,k} \otimes \cdots \otimes U_{LE_L,k}$, $\mathbf{V}_k = V^\dagger_{1E_1,k} \otimes V^\dagger_{2E_2,k} \cdots \otimes V^\dagger_{LE_L,k}$, and $|\mathbf{0}\rangle\langle \mathbf{0}|_E = |0\rangle\langle 0|_{E_1} \otimes |0\rangle\langle 0|_{E_2} \cdots \otimes |0\rangle\langle 0|_{E_L}$, and each $U_{lE_l,k}$ and $V_{lE_l,k}$ is applied jointly on the lth subsystem and the ancilla E_l. Denoting $\Phi_{l,k}(\rho_l) = \mathrm{Tr}_{E_l}\left[U_{lE_l,k}\left(\rho_l \otimes |\mathbf{0}\rangle\langle \mathbf{0}|_{E_l}\right)V^\dagger_{lE_l,k}\right]$ to be the generalised quantum operation acting on the lth subsystem, we thus have

$$\Phi_k = \Phi_{1,k} \otimes \Phi_{2,k} \otimes \cdots \otimes \Phi_{L,k}, \tag{6.9}$$

which applies separately on each subsystem. When we choose a sufficient number of Φ_k, we can always decompose the instant joint evolution $\mathcal{V}^{\text{int}}(t)$ as a linear combination of local operations,

$$\mathcal{V}^{\text{int}}(\delta t) = \mathcal{I} + \delta t \sum_k \alpha_k \Phi_k = \mathcal{I} + \delta t \sum_k \alpha_k \Phi_{1,k} \otimes \Phi_{2,k} \otimes \cdots \otimes \Phi_{L,k}, \tag{6.10}$$

where $\mathcal{I}$ corresponds to the identity channel $\mathcal{I}(\rho) = \rho$ and α_k are complex coefficients. For example, we can choose $\{\Phi_k\}$ to be a complete basis for all quantum channels. When the set of $\{\Phi_k\}$ is chosen, we can find the coefficients α_k via linear programming.

Now we can express the joint evolution as

$$\mathcal{U}(T) = \prod_{i=1}^{T/\delta t} \left[\left(\mathcal{I} + \delta t \sum_k \alpha_k \Phi_k \right) \circ \bigotimes_{l=1} \mathcal{U}_l(\delta t) \right] + \mathcal{O}(T\delta t). \tag{6.11}$$

Denote $\Phi_0 = \mathcal{I}$, $c(\delta t) = 1 + \sum_k |\alpha_k| \delta t$, $p_0(\delta t) = 1/c(\delta t)$, $p_k(\delta t) = |\alpha_k| \delta t / c(\delta t)$, $\theta_k = -i \ln(\alpha_k / |\alpha_k|)$, we can re-express the above equation as

$$\begin{aligned}
\mathcal{U}(T) &= \prod_{i=1}^{T/\delta t} \left[c(\delta t) \left(\sum_k e^{i\theta_k} p_k(\delta t) \Phi_k \right) \circ \bigotimes_{l=1} \mathcal{U}_l(\delta t) \right] + \mathcal{O}(T\delta t), \\
&= c(\delta t)^{T/\delta t} \sum_{k_1, k_2, \ldots, k_{T/\delta t}} \prod_{i=1}^{T/\delta t} \left[e^{i\theta_{k_i}} p_{k_i}(\delta t) \Phi_{k_i} \circ \bigotimes_{l=1} \mathcal{U}_l(\delta t) \right] + \mathcal{O}(T\delta t), \\
&= c(\delta t)^{T/\delta t} \sum_{\mathbf{k}} e^{i\theta_{\mathbf{k}}} p_{\mathbf{k}} \prod_{i=1}^{T/\delta t} \left[\bigotimes_{l=1} \Phi_{l,k_i} \circ \bigotimes_{l=1} \mathcal{U}_l(\delta t) \right] + \mathcal{O}(T\delta t), \\
&= c(\delta t)^{T/\delta t} \sum_{\mathbf{k}} e^{i\theta_{\mathbf{k}}} p_{\mathbf{k}} \bigotimes_{l=1} \left[\prod_{i=1}^{T/\delta t} \left(\Phi_{l,k_i} \circ \mathcal{U}_l(\delta t) \right) \right] + \mathcal{O}(T\delta t).
\end{aligned} \tag{6.12}$$

Here $\mathbf{k} = (k_1, \ldots, k_{T/\delta t})$, $p_{\mathbf{k}} = p_{k_1} p_{k_1} \ldots p_{k_{T/\delta t}}$, $\theta_{\mathbf{k}} = \theta_{k_1} + \theta_{k_2} \cdots + \theta_{k_{T/\delta t}}$. The phase is denoted as $\mathcal{P}_{\mathbf{k}} = e^{i\theta_{\mathbf{k}}}$.

Under the expansion in Eq. (6.12), the joint evolution $\mathcal{U}(T)$ becomes a series of different trajectories. Here, each trajectory is defined by which operations act at each time, including the local time evolution $\mathcal{U}_l(\delta t)$ of each subsystem and one of the generalised quantum operation $\Phi_k(\rho)$ that on average emulates the effect of $\mathcal{V}^{\text{int}} = e^{-iV^{\text{int}}t}$. The whole evolution $\mathcal{U}(T)$ is now decomposed as a linear combination of operations that act locally on each subsystem. The joint evolution can thus be effectively realised using only local operations, which is the central idea of PQS.

I next discuss the measurement of non-local observable. Suppose the initial state $\rho(0)$ is decomposed as

$$\rho(0) = \sum_{k_0} \alpha_{k_0} \bigotimes_{l=1} \rho_{l,k_0} \tag{6.13}$$

and measure an observable like

$$O = \sum_{k_O} \alpha_{k_O} \prod_{l=1} O_{l,k_O}. \tag{6.14}$$

where we have abused the notation of α. The expectation value of O with respect to the time-evolved state can be expressed as

$$\mathrm{Tr}\Big[\mathcal{U}(T)[\rho(0)]O\Big] = c(\delta t)^{T/\delta t} \sum_{k_O,k_0,\mathbf{k}} \alpha_{k_0}\alpha_{k_O} e^{\theta_{\mathbf{k}}} p_{\mathbf{k}} \prod_{l=1} \mathrm{Tr}\Big[\prod_{i=1}^{T/\delta t}\Big(\Phi_{l,k_i}\circ\mathcal{U}_l(\delta t)\Big)[\rho_{l,k_0}]O_{l,k_O})\Big] + \mathcal{O}(T\delta t). \tag{6.15}$$

Here each term $\mathrm{Tr}\Big[\prod_{i=1}^{T/\delta t}\Big(\Phi_{l,k_i}\circ\mathcal{U}_l(\delta t)\Big)[\rho_{l,k_0}]O_{l,k_O})\Big]$ can be obtained from operations only on the lth subsystem. Consequently, the expectation value of observables with respect to an arbitrary joint state is now a linear combination of products of local measurement results.

6.2.2 Monte Carlo implementation and continuous time

6.2.2.1 Discrete time Monte Carlo method

The above discretised scheme assumes a small discrete time step and requires applying the operations at each time step δt, which is unnecessary since the effect of the weak interacting operation $\mathcal{V}(\delta t)$ in a short time is close to the identity operation. I address this problem by stochastically applying the operations Φ_k depending on the amplitude of its associated coefficient $|\alpha_k|$.

More specifically, the number of expanded terms is proportional to $N_V^{T/\delta t}$, with N_V being the number of terms in the expansion of Eq. (6.10). Although $N_V^{T/\delta t}$ increases exponentially, we do not need to measure all the expanded terms and the Monte Carlo method could more efficiently obtain the measurement outcome.

In particular, the decomposition of Eq. (6.15) can be written in a general form of

$$\langle O\rangle = \sum_k q_k \prod_{l=1} \mathrm{Tr}\left[\Phi_l(\rho_{l,k})O_{l,k}\right] = C\sum_k e^{i\theta_k} p_k \prod_{l=1} \mathrm{Tr}\left[\Phi_l(\rho_{l,k})O_{l,k}\right] + \mathcal{O}(T\delta t), \tag{6.16}$$

with $k = (k_O, k_0, \mathbf{k})$, $q_k = c(\delta t)^{T/\delta t}\alpha_{k_0}\alpha_{k_O}e^{i\theta_{\mathbf{k}}}p_{\mathbf{k}}$, $C = \sum_k q_k = c(\delta t)^{T/\delta t}\sum_{k_0}|\alpha_0|\sum_{k_O}|\alpha_O|$, $\theta_k = -i\ln(q_k/|q_k|)$, $p_k = |q_k|/C$, and $\Phi_l = \prod_{i=1}^{T/\delta t}\Big(\Phi_{l,k_i}\circ\mathcal{U}_l(\delta t)\Big)$. To obtain the measurement $\langle O\rangle$, we can use the following Monte Carlo random sampling method,

1. Generate random numbers k according to the probability $\{p_k\}$;
2. For the lth subsystem, prepare state $\rho_{l,k}$, apply the operation Φ_l, and measure the observable $O_{l,k}$ to get $\langle O_{l,k}\rangle$.
3. Multiply all the outcomes $\langle O_{l,k}\rangle = \mathrm{Tr}[\Phi_l(\rho_{l,k})O_{l,k}]$ of different subsystems, as well as the phase $e^{i\theta_k}$ and C.
4. Repeat steps 1-3 N_s time and output $O_{\mathrm{est}} = \sum_k Ce^{i\theta_k}\prod_l\langle O_{l,k}\rangle$.

Ignoring the effect of Trotter error with a finite time step, the expansion guarantees that the output is an unbiased estimation of the exact measurement outcome. Suppose each $O_{l,k}$ is a Pauli measurement, then with failure probability δ, the estimation error scales as

$$\varepsilon = \mathcal{O}\left(C\sqrt{\frac{\log_2 1/\delta}{N_s}}\right). \tag{6.17}$$

Since the coefficient C boosts the error, it quantifies the cost of the random sampling process. Suppose the input state is a product state, then the additional cost that the perturbative expansion introduces is $C = c(\delta t)^{T/\delta t}$. A detailed analysis of the resource cost is shown in Section 6.2.3.

A major caveat to the above scheme is that it assumes a small discrete time step and requires continuously interchanging the subsystem evolution $\mathcal{U}_l$ and Φ_l with a sufficiently small time step δt. In practice, it could be challenging to 'continuously' interchange the subsystem evolution within a sufficiently small time step δt. I show in the next subsection that an equivalent Monte Carlo method can be applied, which stochastically implements the joint evolution. As such, a general Hamiltonian simulation method other than Trotterisation could be applied to reduce the algorithmic error.

6.2.2.2 Stochastic implementation

First rewrite Eq. (6.11) as follows

$$\mathcal{U}(T) = c(\delta t)^{T/\delta t} \prod_{i=1}^{T/\delta t} \left(p_0 \mathcal{I} + p_{\geq 1}\tilde{\Phi}\right) \circ \bigotimes_{l=1} \mathcal{U}_l(\rho(0)), \tag{6.18}$$

where $p_{\geq 1} = \sum_{k\geq 1} p_k$, $\tilde{\Phi} = \sum_{k\geq 1} \alpha_k \Phi_k / \sum_k |\alpha_k|$. At each time step, we always evolve each subsystem according to $\mathcal{U}_l$, and with a small probability $p_{\geq 1}$, we evolve under $\tilde{\Phi}$. Since the probability $p_{\geq 1} \propto \delta t$ is negligible when taking the limit of $\delta t \to 0$, we can equivalently realise it with a continuous decaying or jump process. Specifically, we can realise the evolution $\mathcal{U}(T)$ with the following stochastic process

1. Generate a uniformly distributed random number $p_{\text{jp}} \in [0, 1]$.
2. Determine t_{jp} by solving $p_{\text{jp}} = Q(t)$ with $Q(t) = e^{-\Gamma(t)}$, $\Gamma(t) = t\sum_{k\geq 1} \tilde{p}_k$, and $\tilde{p}_k = \lim_{\delta t\to 0} p_k/\delta t = |\alpha_k|$.
3. Evolve each subsystem state with $\mathcal{U}_l$ to time t and update $t = t + t_{\text{jp}}$.
4. Generate another random number $q_m \in [0, 1]$ to determine Φ_k and apply $\Phi_{l,k}$ to the lth subsystem.
5. Repeat Step $1 - 4$ until $t = T$.

The average of different trajectories reproduces the evolution under the joint channel $\mathcal{U}(T)$. We can therefore stochastically realise the decomposition without assuming a discrete time.

Note that the number of generalised quantum operations required to realise the joint evolution scales proportionally to the interaction strength as $\mathcal{O}(\sum_k |\alpha_k| T)$, and on average the stochastic implementation is proven to be equivalent to the discretised scheme in Figure 6.1(a). The equivalence is proven in Section D.1. Meanwhile, other advanced Hamiltonian simulation algorithms such as qubitisation introduced in Section F.2 could be used for each time evolution at step 3. I also note that the jump time t_{jp} as well as the evolution could be predetermined, which renders its implementation similar to conventional Hamiltonian simulation.

I illustrate the procedure in Figure 6.1 and summarise the key steps of PQS in Algorithm 3.

Algorithm 3 Perturbative quantum simulation

1: Given a set of generalised quantum operations, find the decomposition Eq. (6.10).
2: Generate a sequence of trajectories where each subsystem evolves and experiences random local generalised quantum operations.
3: Sample from the trajectories. The average behaviour reproduces the joint evolution.

In quantum simulation, we usually prepare a simple initial state, such as a product state, evolve it under a joint Hamiltonian H, and measure an observable that could be decomposed as a linear combination of local ones, i.e., the input state $\rho(0) = \bigotimes_l \rho_l(0)$ and product measurement $O = \bigotimes_l O_l$. When the input state or the measurement is not in a product form, we can similarly decompose them as we discuss above.

By applying our algorithm, the whole simulation process is now decomposed into the average of different ones, each of which only involves operations on the subsystems. If each subsystem consists of n qubits, then we can effectively simulate nL qubits with operations on only $\mathcal{O}(n)$ qubits. Although the aforementioned analysis assumes Trotterisation, the local time evolution $\mathcal{U}_l(t)$ could be implemented with any Hamiltonian simulation methods, such as product formulae discussed in Chapter 5 and Refs. [37, 144], and quantum signal processing [32, 33] discussed in Chapter 2. Our algorithm is compatible with both near-term devices and fault-tolerant quantum computers.

6.2.3 Cost analysis

The above perturbative quantum simulation (PQS) method introduces a sampling overhead which is quantified by

$$C = \lim_{\delta t \to 0} \prod_{i=1}^{T/\delta t} c(\delta t) = \lim_{\delta t \to 0} \prod_{i=1}^{T/\delta t} (1 + \alpha \delta t) = e^{T\alpha}, \tag{6.19}$$

where $\alpha = \sum_k |\alpha_k|$. Since the simulation accuracy is now C times larger, we need to have $C = \mathcal{O}(1)$ and hence $\alpha T = \mathcal{O}(1)$ in order to get an accurate result. This could be satisfied

when T and α are not too large, i.e., when the product of the simulation time and the total interaction strength is constant. While α roughly measures the interaction strength, its analytical relationship to the interaction Hamiltonian V^{int} is not obvious. This is because the value of α depends on the choice of generalised quantum operations and the decomposition. We can thus define the minimal value of α by optimising over all possible decompositions,

$$\alpha_{\min} = \min_{\{\Phi_k\}} \alpha(\{\alpha_k, \Phi_k\}), \tag{6.20}$$

where $\alpha(\{\alpha_k, \Phi_k\})$ is written as a function of the generalised quantum operations and the minimisation is over all possible decomposition strategies.

Below, I provide an analytical lower bound to $\alpha_{\min}$ as a function of the interaction V^{int}. In the next section, I demonstrate an explicit decomposition strategy that achieves this lower bound.

To do so, I consider the Choi state of the instant evolution $\mathcal{V}^{\text{int}}(\delta t)$ by inputting tensor products of the maximally entangled states. Specifically, inputting $|\phi\rangle_{l,l'} = \sum_j |jj\rangle_{l,l'} / \sqrt{d}$ to the lth subsystem with d being the dimension, the output state $\phi^{\text{int}}_{1,1',\ldots,L,L'}$ is the Choi state,

$$\phi^{\text{int}}_{1,1',\ldots,L,L'} = \mathcal{V}^{\text{int}}(\delta t)\left[\bigotimes_l \phi_{l,l'}\right]. \tag{6.21}$$

Suppose that $\mathcal{V}^{\text{int}}(\delta t)$ is decomposed as

$$\mathcal{V}^{\text{int}}(\delta t) = \sum_k \tilde{\alpha}_k \Phi_{1,k} \otimes \Phi_{2,k} \otimes \cdots \otimes \Phi_{L,k}, \tag{6.22}$$

where we have put $\mathcal{I}$ into the summation and denote $\tilde{\alpha}_k$ as the new coefficient incorporating δt. The relation between α and $\tilde{\alpha} = \sum_k |\tilde{\alpha}_k|$ is

$$\alpha = \lim_{\delta t \to 0} \frac{\tilde{\alpha} - 1}{\delta t}. \tag{6.23}$$

Since α depends linearly on $\tilde{\alpha}$, we can equivalently minimise $\tilde{\alpha}$.

Define isomorphisms S and T of a general matrix $M = \sum_{i,j} M_{i,j} |i\rangle \langle j|$ as

$$\begin{aligned} S(M) &= \sum_{i,j} M_{i,j} |i\rangle |j\rangle, \\ T(M) &= \sum_{i,j} M_{i,j} \langle i| \langle j|. \end{aligned} \tag{6.24}$$

Several useful properties of the matrices S and T are

- The definitions of $S(M)$ and $T(M)$ are basis dependent.
- When applying matrices U and V to M, we have

$$\begin{aligned} S(UMV) &= U \otimes V^T \sum_{i,j} M_{i,j} |i\rangle |j\rangle = U \otimes V^T S(M), \\ T(UMV) &= \sum_{i,j} M_{i,j} \langle i| \langle j| U^T \otimes V = T(M) U^T \otimes V. \end{aligned} \tag{6.25}$$

- $S(M)$ and $T(M)$ are related as follows,

$$S(M) = [T(M^*)]^\dagger . \tag{6.26}$$

This is true because $[T(M^*)]^\dagger = \left[\sum_{i,j} M^*_{i,j} \langle i| \langle j|\right]^\dagger = \sum_{i,j} M_{i,j} |i\rangle |j\rangle = S(M)$. Suppose we denote $|M\rangle = S(M)$ then $T(M) = [S(M^*)]^\dagger = \langle M^*|$.

- The norms of S and T are the same,

$$S(M)^\dagger \cdot S(M) = T(M) \cdot T(M)^\dagger = \mathrm{Tr}[M^\dagger M] = \|M\|_2^2, \tag{6.27}$$

which corresponds to Schatten-2 norm of M. This is because

$$S(M)^\dagger \cdot S(M) = \sum_{i',j'} M^*_{i',j'} \langle i'| \langle j'| \sum_{i,j} M_{i,j} |i\rangle |j\rangle = \sum_{i,j} M^*_{i,j} M_{i,j} = \mathrm{Tr}[M^\dagger M].$$

The proof is similar for $T(M)$.

By applying S to the l, l' systems and T to the rest systems, we get a matrix

$$\psi^{\mathrm{int}}_{l,l'} = S_{l,l'} \circ \bigotimes_{j\neq l, j'\neq l'} T_{j,j'}(\phi^{\mathrm{int}}_{1,1',\ldots,L,L'}). \tag{6.28}$$

Using, the above property, we can thus lower bound $\tilde{\alpha}$ as follows.

Theorem 4. *Given a decomposition of Eq. (6.22) with generalised quantum operations $\{\Phi_{l,k}\}$, we have*

$$\tilde{\alpha} \geq \max_l \left\|\psi^{\mathrm{int}}_{l,l'}\right\|_1, \tag{6.29}$$

where $\|A\|_1 = \mathrm{Tr}[\sqrt{AA^\dagger}]$ is the trace norm.

I leave the technical proof of Theorem 4 in Appendix D. Consider the specific form of $\mathcal{V}^{\mathrm{int}}(\delta t)[\rho] = \rho + \delta t(-iV^{\mathrm{int}}\rho + i\rho V^{\mathrm{int}})$ and define the interaction part as

$$\bar{\mathcal{V}}^{\mathrm{int}}[\rho] = -iV^{\mathrm{int}}\rho + i\rho V^{\mathrm{int}}. \tag{6.30}$$

We can then similarly define the Choi state of $\tilde{\mathcal{V}}^{\mathrm{int}}$ as

$$\bar{\phi}^{\mathrm{int}}_{1,1',\ldots,L,L'} = \bar{\mathcal{V}}^{\mathrm{int}}\left[\bigotimes_l \phi_{l,l'}\right], \tag{6.31}$$

and the matrices

$$\bar{\psi}^{\mathrm{int}}_{l,l'} = S_{l,l'} \circ \bigotimes_{j\neq l, j'\neq l'} T_{j,j}(\bar{\phi}^{\mathrm{int}}_{1,1',\ldots,L,L'}). \tag{6.32}$$

Then consider the decomposition Eq. (6.10), we have

Corollary 1. *Given a decomposition of Eq. (6.10) with generalised quantum operations $\{\Phi_{l,k}\}$, we have*

$$\alpha \geq \max_l \left\|\bar{\psi}^{\mathrm{int}}_{l,l'}\right\|_1, \tag{6.33}$$

where $\|A\|_1 = \mathrm{Tr}[\sqrt{AA^\dagger}]$ is the trace norm.

In the following section, a specific decomposition strategy will be introduced and we shall see how to use the analytical lower bound to prove its optimality.

6.2.3.1 Optimising the decomposition

Here, in order to minimise the simulation cost, we could consider an over-complete basis with generalised quantum operations and then find an optimised decomposition through linear programming.

Specifically, consider a set of over-complete basis $\{\Phi_k\}$ which includes the identity channel. Our target is to solve the following problem.

$$\begin{aligned} \min C_1 &= \sum_k \alpha_k^+ - \sum_{k'} \alpha_{k'}^-, \\ \text{such that } \mathcal{V}^{\text{int}}(\delta t) &= \sum_k \alpha_k^+ \Phi_k - \sum_{k'} \alpha_{k'}^- \Phi_{k'}, \\ \alpha_k^+, \alpha_k^- &\geq 0. \end{aligned} \tag{6.34}$$

There are a few problems evident here. First, the optimisation becomes exponentially costly when the channel acts on a large number of qubits. Second, basis operations also contain measurement and state preparation, which might be challenging in an experiment. In the next section, we give another explicit decomposition strategy to resolve these problems. The explicit decomposition could be optimal under mild conditions, and only requires unitary operations without measurements or state preparation.

6.3 An explicit decomposition method

While the decomposition of Eq. (6.10) holds in general for a complete set of $\{\Phi_k\}$, it may involve some operations that render experimental implementation challenging. Here, I address this problem by developing an explicit decomposition with only local unitary operations. Its optimality over other expansions under a mild condition is demonstrated in Section 6.3.2.

6.3.1 Method description

Suppose $V^{\text{int}} = \sum_j \lambda_j V_j^{\text{int}}$ with each V_j^{int} being a tensor product of Pauli operators. Consider an expansion of $\mathcal{V}^{\text{int}}$, which is given by

$$\begin{aligned} \mathcal{V}^{\text{int}}(\delta t)[\rho] &= \mathcal{I}(\rho) - i\delta t(V^{\text{int}}\rho I - \rho V^{\text{int}}) + O(\delta t^2) \\ &= \mathcal{I}(\rho) - i\delta t \sum_j \lambda_j (V_j^{\text{int}}\rho - \rho V_j^{\text{int}}) + O(\delta t^2), \end{aligned} \tag{6.35}$$

where both $V_j^{\text{int}}\rho$ and ρV_j^{int} are generalised quantum operations. Suppose $V_j^{\text{int}} = \bigotimes_l V_{l,j}^{\text{int}}$ and the input state is a product state $\rho = \bigotimes_l \rho_l$ the above decomposition could be expressed generally as

$$\mathcal{V}^{\text{int}}(\delta t)\left[\bigotimes_l \rho_l\right] = c(\delta t) \sum_k e^{i\theta_k} p_k \bigotimes_l \left[\tilde{U}_{l,k}\rho_l \tilde{V}_{l,k}\right] + O(\delta t^2). \tag{6.36}$$

Here each $\tilde{U}_{l,k}$ and $\tilde{V}_{l,k}$ could be I and $V^{\text{int}}_{l,j}$, $c(\delta t) = 1 + 2\delta t \sum_j |\lambda_j|$, p_k and θ_k are defined correspondingly. Denoting the unitary evolution of the lth subsystem as $U_l(\delta t)$ and following the notation of the above discussion, the joint evolution of all the subsystems is

$$\mathcal{U}(T)\left[\bigotimes_l \rho_l\right] = C\sum_{\mathbf{k}} e^{i\theta_{\mathbf{k}}} p_{\mathbf{k}} \bigotimes_l \left[\tilde{U}_{l,k_{T/\delta t}} U_l(\delta t) \ldots \tilde{U}_{l,k_1} U_l(\delta t) \rho_l U_l^\dagger(\delta t) \tilde{V}_{l,k_1} \ldots U_l^\dagger(\delta t) \tilde{V}_{l,k_{T/\delta t}}\right], \tag{6.37}$$

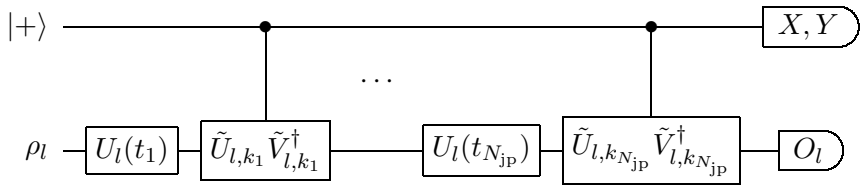
where $C = c(\delta t)^{T/\delta t} = e^{2T\lambda}$ with $\lambda = \sum_j |\lambda_j|$. Now we have decoupled the joint evolution as a linear combination of independent evolution of each subsystem. When we further implement the stochastic Monte Carlo method, the evolution of each subsystem looks like

$$\rho_{l,\mathbf{k}} = \tilde{U}_{l,k_{N_{\text{jp}}}} U_l(t_{N_{\text{jp}}}) \ldots \tilde{U}_{l,k_1} U_l(t_1) \rho_l U_l^\dagger(t_1) \tilde{V}_{l,k_1} \ldots U_l^\dagger(t_{N_{\text{jp}}}) \tilde{V}_{l,k_{N_{\text{jp}}}}, \tag{6.38}$$

where N_{jp} is the number of jumps (decay events), and $t_1 + t_2 + \cdots + t_{N_{\text{jp}}} = T$. Here, each $\tilde{U}_{l,k_i}$ is either I or one of $\{V^{\text{int}}_{l,j}\}$. When we measure O_l, it becomes

$$\text{Tr}[\rho_{l,\mathbf{k}} O_l] = \text{Tr}[\tilde{U}_{l,k_{N_{\text{jp}}}} U_l(t_{N_{\text{jp}}}) \ldots \tilde{U}_{l,k_1} U_l(t_1) \rho_l U_l^\dagger(t_1) \tilde{V}_{l,k_1} \ldots U_l^\dagger(t_{N_{\text{jp}}}) \tilde{V}_{l,k_{N_{\text{jp}}}} O], \tag{6.39}$$

which could be implemented with the following circuit



The measurement result of the whole evolution state is

$$\text{Tr}\left[\mathcal{U}(T)\left[\bigotimes_l \rho_l\right] \cdot \bigotimes_l O_l\right] = C\sum_{\mathbf{k}} e^{i\theta_{\mathbf{k}}} p_{\mathbf{k}} \prod_l \text{Tr}[\rho_{l,\mathbf{k}} O_l]. \tag{6.40}$$

Therefore, after measuring each $\text{Tr}[\rho_{l,\mathbf{k}} O_l]$, we can obtain the exact measurement result.

The perturbative quantum simulation method with explicit decomposition is proven to be equivalent to the infinite-order Dyson series. I would direct the interested reader to Appendix D for a description of the Dyson series, its implementation on a quantum computer, and Theorem 10 for a rigorous proof.

6.3.2 Cost analysis

According to the above discussion, the cost associated with the explicit expansion is

$$C = e^{2T\lambda}, \tag{6.41}$$

with $\lambda = \sum_j |\lambda_j|$. I demonstrate that the expansion is optimal, i.e., with the smallest cost, when V^{int} satisfies the following condition.

Condition 1. *Suppose V^{int} acts nontrivially on the set of subsystems $\mathcal{S}$. Given $V^{\text{int}} = \sum_j \lambda_j V_j^{\text{int}}$ with each $V_j^{\text{int}} = \bigotimes_l V_{l,j}^{\text{int}}$ and $V_{l,j}^{\text{int}}$ being a tensor product of Pauli operators, we have*

$$\begin{aligned} \text{Tr}\left[V_{l,j}^{\text{int}}\right] &= 0,\ \forall j, \forall l \in \mathcal{S}, \\ \text{Tr}\left[V_{l,j}^{\text{int}} V_{l,j'}^{\text{int}}\right] &= 0,\ \forall j \neq j', \forall l \in \mathcal{S}. \end{aligned} \tag{6.42}$$

The first condition requires that each $V_{l,j}^{\text{int}}$ is non-identity and the second condition requires two interaction terms of the same system are orthogonal. When we say V^{int} acts nontrivially on subsystems $\mathcal{S}$, it means that for any $l \in \mathcal{S}$, at least one of V_j^{int} has non-identity Pauli operators on subsystem l. The result is shown in Theorem 5.

Theorem 5. *Suppose the interaction V^{int} satisfies Condition 1. The explicit expansion of Eq. (6.35) has a minimal cost over all possible decomposition strategies.*

Proof. We assume that V^{int} acts nontrivially on all the systems. The proof is similar to the general case. The Choi state of the interaction part $\tilde{\mathcal{V}}^{\text{int}}[\rho] = -iV^{\text{int}}\rho + i\rho V^{\text{int}}$ is

$$\tilde{\phi}^{\text{int}}_{1,1',\ldots,L,L'} = \tilde{\mathcal{V}}^{\text{int}}\left[\bigotimes_l \phi_{l,l'}\right] = -i\sum_j \lambda_j \left[\bigotimes_l V_{l,j}^{\text{int}} \left|\phi\right\rangle_{l,l'} \left\langle\phi\right|_{l,l'} - \bigotimes_l \left|\phi\right\rangle_{l,l'} \left\langle\phi\right|_{l,l'} V_{l,j}^{\text{int}}\right]. \tag{6.43}$$

Taking $\tilde{\psi}^{\text{int}}_{1,1}$ as an example, we have

$$\begin{aligned} \tilde{\psi}^{\text{int}}_{1,1'} = &-i\sum_j \lambda_j \Big[V_{1,j}^{\text{int}} \left|\phi\right\rangle_{1,1'} \otimes \left|\phi\right\rangle_{1,1'} \bigotimes_{l\geq 2} \left\langle\phi\right|_{l,l'} (V_{l,j}^{\text{int}})^T \otimes \left\langle\phi\right|_{l,l'} - \left|\phi\right\rangle_{1,1'} \\ &\bigotimes (V_{1,j}^{\text{int}})^T \left|\phi\right\rangle_{1,1'} \bigotimes_l \left\langle\phi\right|_{l,l'} \otimes \left\langle\phi\right|_{l,l'} V_{l,j}^{\text{int}}\Big]. \end{aligned} \tag{6.44}$$

Denoting

$$\begin{aligned} \left|\psi_{1,j,a}\right\rangle &= V_{1,j}^{\text{int}} \left|\phi\right\rangle_{1,1'} \otimes \left|\phi\right\rangle_{1,1'}, \\ \left|\psi_{1,j,b}\right\rangle &= \left|\phi\right\rangle_{1,1'} \otimes (V_{1,j}^{\text{int}})^T \left|\phi\right\rangle_{1,1'}, \\ \left\langle\psi_{l,j,a}\right| &= \left\langle\phi\right|_{l,l'} (V_{l,j}^{\text{int}})^T \otimes \left\langle\phi\right|_{l,l'}, \\ \left\langle\psi_{l,j,b}\right| &= \left\langle\phi\right|_{l,l'} \otimes \left\langle\phi\right|_{l,l'} V_{l,j}^{\text{int}}, \end{aligned} \tag{6.45}$$

we can express $\tilde{\psi}^{\text{int}}_{1,1}$ as

$$\tilde{\psi}^{\text{int}}_{1,1'} = -i\sum_j \lambda_j c\left[\left|\psi_{1,j,a}\right\rangle \bigotimes_{l\geq 2} \left\langle\psi_{l,j,a}\right| - \left|\psi_{1,j,b}\right\rangle \bigotimes_{l\geq 2} \left\langle\psi_{l,j,b}\right|\right]. \tag{6.46}$$

When $\{V_{l,j}\}$ satisfies Condition 1, elements in $\{\left|\psi_{1,j,a}\right\rangle, \left|\psi_{1,j,b}\right\rangle\}$ are mutually orthogonal, i.e.,

$$\left\langle\psi_{1,j,x}|\psi_{1,j',y}\right\rangle = 0, \forall j \neq j' \text{ or } x \neq y,\ x, y \in \{a, b\}. \tag{6.47}$$

Similarly elements in $\{\left\langle\psi_{l,j,a}\right|, \left\langle\psi_{l,j,b}\right|\}$ are mutually orthogonal. Therefore, Eq. (6.46) is a singular value decomposition of $\tilde{\psi}^{\text{int}}_{1,1}$ and we have

$$\|\tilde{\psi}^{\text{int}}_{1,1'}\|_1 = 2\sum_j |\lambda_j|. \tag{6.48}$$

The above proof holds for all other $\tilde{\psi}^{\text{int}}_{l,l'}$. □

The following provides several examples of V^{int} that satisfy the condition. First, the condition is satisfied when there is only one interaction term V_1^{int}.

Corollary 2. *The decomposition is optimal when V^{int} only has one term (a tensor product of Pauli matrices).*

Condition 1 could also hold when the interaction V^{int} has multiple terms. For example, consider three subsystems and denote (l, m) to be the mth qubit of the lth subsystem. The following examples of different interactions satisfy Condition 1.

$$\begin{aligned}
V^{\text{int}} =& aX_{1,1} \cdot X_{2,1} \cdot X_{3,1} + bX_{1,2} \cdot X_{2,2} \cdot X_{3,2} + cX_{1,3} \cdot X_{2,3} \cdot X_{3,3}, \\
V^{\text{int}} =& aX_{1,1} \cdot X_{2,1} \cdot X_{3,1} + bY_{1,1} \cdot Y_{2,1} \cdot Y_{3,1} + cZ_{1,1} \cdot Z_{2,1} \cdot Z_{3,1}, \\
V^{\text{int}} =& aX_{1,1}Y_{1,2} \cdot X_{2,1}Z_{2,2} \cdot Y_{3,1}Y_{3,2} + bX_{1,1}Z_{1,2} \cdot Z_{2,1}Z_{2,2} \cdot X_{3,1}Y_{3,1} \\
&+ cZ_{1,1}Y_{1,2} \cdot X_{2,1}Y_{2,2} \cdot Z_{3,1}Y_{3,2} + dZ_{1,1}Z_{1,2} \cdot Z_{2,1}Y_{2,2} \cdot X_{3,1}Z_{3,1}.
\end{aligned} \tag{6.49}$$

Since Condition 1 requires that the interaction terms of each subsystem are mutually orthogonal, the number of interaction terms is limited.

Proposition 5. *When $V^{\text{int}} = \sum_{j=1}^{N_{\text{int}}} \bigotimes_l V_{l,j}^{\text{int}}$ satisfies Condition 1 with N_{int} being the number of terms and n being the minimal number of qubits of each subsystem. Suppose the minimal weight of each $V_{l,j}^{\text{int}}$ is k, and we have*

$$N_{\text{int}} \leq 3^k \binom{n}{k}. \tag{6.50}$$

In particular, when $k = 1$, i.e., the interaction on each subsystem only act on one qubit, Condition 1 requires V^{int} to have at most $3n$ terms. We leave a more detailed exploration of conditions for the optimal decomposition to future work on the subject.

6.4 Limitations and applications

In this section, I discuss the potential applications and limitations of the PQS algorithm. Since PQS is a hybrid method that combines quantum computing and classical perturbation theory, it inherits both their advantages and limitations. The major constraint on the use of PQS comes from the limitation of classical perturbation theories, which generally only work for weakly correlated systems. For PQS, since we simulate the coupling of subsystems using the perturbation theory, it is only efficient if the coupling strength is weak. In comparison to a universal quantum simulation algorithm using a fault-tolerant quantum computer, the systems that our method is capable of simulating are limited. Specifically, our method cannot work for large systems with general arbitrary two-body interactions, such as strongly correlated electrons, high-dimensional strongly interacting lattice Hamiltonians, or scenarios when the simulation time is very long. Nevertheless, note that similar limitations prevail in almost all modern classical computing methods apart from perturbation theory, such as density functional theory (DFT), quantum Monte Carlo

(QMC), tensor and neural network methods, etc. As I discuss below, these limitations do not prevent wide applications of PQS to realistic problems.

While our method is restricted in scope compared to a universal quantum simulation algorithm, it is worth pointing out that these universal quantum algorithms generally rely on a fault-tolerant quantum computer, which is still challenging to realise using current technology. This simple fact explains why many contemporary works in quantum computing focus on application in the NISQ era or the early stage of fault-tolerant quantum computing, where both the size (number of qubits) and (circuit) depth of the quantum hardware could be limited. Therefore, as we elaborate below, our method does indeed have broad application from simulating intricate quantum many-body systems and probing interesting physics phenomena, to benchmarking larger quantum processors for NISQ devices and early stage fault-tolerant quantum computers.

Theoretically, our method combines the complementary strengths of quantum computing and perturbation theory, to respectively simulate the subsystem and the inter-subsystem interactions. PQS would be most powerful in investigating large systems with weak inter-subsystem interactions or intermediate systems with general interactions. Since there is no assumption regarding the subsystem interactions, locality of the inter-subsystem interactions, or the initial state of the subsystems, the method is widely applicable.

The most promising and exciting application of perturbative quantum simulation concerns clustered subsystems with weak subsystem-wise interactions. Since all subsystems could have arbitrarily strong interactions, the whole system might not be efficiently solvable using classical approaches, in general. In contrast, our methods are significantly more resource efficient, making the study of system dynamics a viable proposition. One prominent example of a suitable application of our approach is the simulation of 1D systems, where we could easily divide the systems into clustered subsystems and have weak subsystem-wise interactions. It could be argued that under certain assumptions (local and gapped Hamiltonians), the ground state of a 1D system could be efficiently solvable using matrix product states, hence the ability of our approach to simulate 1D systems is unsurprising. By way of response, it is worth pointing out that simulating the dynamics of general 1D systems is actually a very challenging task using classical methods. Indeed, as shown in Ref. [208], one can use a 1D system to simulate a 2D system, which indicates classically simulating an arbitrary 1D systems could be hard. In contrast, our method could be applied to study exotic properties of 1D system dynamics with strong subsystem interactions.

In addition, many real-world physical systems also admit clustered interaction forms. Many types of realistic Hamiltonians, such as molecules (the dimer system or functionalised compound), lattice models for materials, or toy models for high energy physics, admit a similar form as that in our simulation, indicating a natural application of our method. Specific examples include the molecular rings, such as the $(Cr_7Ni)_2$ dimer, in which the subsystem has strong correlations and the two subsystems are weakly interacting via spins on the boundary.

Aside from physical systems featuring these geometrical cluster properties, PQS is also applicable for systems with multiple degrees of freedom. Indeed, quantum many-body systems that consist of both weak and strong correlations in different levels of the system could be suitable for our methods. Taking the Hamiltonian of molecules as an example, we can divide the system into electrons and nuclei. We can then separately simulate the two subsystems of the electrons and process their correlation classically to surpass the Born-Oppenheimer approximation. We may use the PQS method to investigate dynamic correlations beyond the Born-Oppenheimer approximation.

The method is also applicable to electronic structure. When considering a relatively large bond distance, we could divide the spin-orbitals from different atoms or molecules into different subsystems. Similarly, we could divide the spin-orbitals into active and virtual ones and represent them with quantum and classical means, respectively. Our method can consequently be employed to construct an effective Hamiltonian for solving large molecules or molecules with more basis states using a small quantum computer. Other examples in condensed matter physics include dynamics simulation of electron-phonon coupling. Many interesting phenomena emerge due to the dynamic coupling of electrons and phonons. Our method could be directly applied to these problems.

Practically, our method could be applied for probing interesting physics phenomena. As will be demonstrated in Chapter 8, PQS can be applied to study the quantum walk of bosons, dynamical phase transition, the propagation of correlations, and spin-charge separation of bosons, fermions, or spins. Apart from these applications, our method could also be used to examine other more general dynamic behaviours, such as molecular reactions of the dimer and electron-phonon interactions in superconducting models. Following a recent study of cluster simulation schemes [209], our method could also be applicable to variational quantum simulation for molecular Hamiltonians. With proper embedding methods [210–212], such as DFT [40,211,213], dynamical mean-field theory (DMFT) [214,215], or density matrix embedding theory (DMET) [216–218], PQS might also be used to probe larger physics problems. These include, for example, ones in the thermodynamic limit.

Furthermore, PQS would be helpful for studying general strongly interacting problems with short time and benchmarking near-term quantum computation. Using the PQS method, a larger problem with $N \cdot L$ qubits could be processed by a $N + 1$ qubit quantum device. Since a smaller quantum device is generally much more accurate than a larger quantum processor due to crosstalks or other types of errors when controlling large quantum systems, our method could serve as a benchmark of the computing result for larger problems. This advantage was also clearly demonstrated using the IBM quantum devices in Section 8.2. Consequently, when we construct larger quantum hardware or aim to use it to demonstrate quantum advantages for

solving a larger problem, we can first run PQS with a smaller device to test the performance of hardware and the feasibility for faithfully implementing the algorithm.

6.5 Discussion

The PQS algorithm leverages the power of quantum computing to simulate the major component of the Hamiltonian, alleviating the constraints of classical perturbation methods. It uses classical perturbation to approximate the interaction, reducing the requirements for qubit numbers, which is favourable for both NISQ devices and early-stage fault-tolerant quantum computers. While the algorithm cannot efficiently simulate arbitrary systems, especially large 2D or 3D systems, it is applicable to ones with an intermediate size, such as a square lattice with tens to hundreds of qubits, and is particularly useful for simulating large systems with weak inter-subsystem interactions, such as (quasi) one-dimensional systems and clustered subsystems. PQS methods can be applied to study various many-body physical phenomena, and has demonstrated potential in benchmarking large quantum processors with small ones, corresponding to an emerging demand in the NISQ and early FTQC era. Numerical and experimental results will be presented in Chapter 8.

Our work set out the potential of hybrid quantum-classical algorithms which combine classical physics methods with quantum computing, clearing a path for studying large many-body quantum systems with near-future quantum hardware. There are, additionally, other classical perturbation treatments of the interaction, such as one which expands according to the interaction strength; integrating it with quantum computing may provide a more efficient PQS. Furthermore, perturbation theories have also been applied in analogue quantum simulation for synthesising effective Hamiltonians (see [219]), and analogue simulation has already been successfully used to discover novel phenomena such as quantum many-body scars [220]. Whether our PQS method could be generalised to analogue quantum simulation is an interesting future direction of research.

Author contributions. This chapter is relevant to a published work [95]. S.E., X.Y., and J.S. conceived the idea. X.Y. and V.V. supervised the project. J.S. and X.Y. developed the theoretical aspect of the work. J.S. and X.Y. wrote the manuscript with input from S.E., V.V., and H.P..

Part III

Exploring emergent quantum phenomena by quantum simulation and spectroscopy

We are generally interested in the eigenstate properties (such as the magnetic ordering of the ground state), dynamic properties, and spectral features of a quantum many-body system. The central aim of Part III is to explore interesting quantum phenomena using the methods developed in this thesis.

Spectroscopy and quantum computing arguably provide a natural and complementary solution to these problems. As discussed in Section 1.3, spectroscopy techniques have a close relation with quantum simulation, from both a theoretical and operational perspective. Spectroscopy stimulates the development of quantum computing methods in predicting excited state properties of quantum many-body systems with general and tuneable interaction.

In Chapter 7, I introduce an engineered spectroscopy method for probing the spectroscopic features of quantum many-body systems on a quantum simulator. This is a continuation of the discussion in Section 1.3, in which I set out the basics of quantum simulation and spectroscopy, illustrating their relation, and drawing some notation comparisons. In Chapter 8, I show how to study eigenstate and dynamic properties using the quantum simulation techniques developed in Chapter 6 and Chapter 7. In particular, I show how to explore many-body phenomena, and probe excitation spectra of quantum many-body systems, including interacting bosons, fermions, and quantum spins.

Chapter 7

Probing spectroscopic features of quantum many-body systems

Spectroscopy techniques provide a powerful tool for exploring many-body phenomena and have achieved great success in probing the excited states of materials, such as vibrational and magnetic structures. In this chapter, motivated by spectroscopy techniques introduced in Section 1.3.3, I introduce a framework that considers engineered quantum dynamics induced by a local or global perturbation to estimate transition energies between the eigenstates of a quantum system. After applying a local perturbation to an initial state which is prepared by an engineered quantum system, the spectral function of the target quantum system can be obtained by looking at the Fourier transform of time-evolved operator correlation functions. I demonstrate a close relation between our engineered spectroscopy method and the quantum algorithmic cooling method introduced in Chapter 4. I further discuss the algorithmic error, and quantum resource requirements for our method, and show that the quantum circuit complexity has a logarithmic dependence on the desired simulation accuracy.

This chapter is relevant to a manuscript under preparation in collaboration with Lucia Vilchez Estevez, Vlatko Vedral, and Andrew Boothroyd [96].

7.1 Motivation

Spectroscopy has been used extensively to facilitate our understanding of many-body phenomena, such as magnetism and superconductivity. In magnetic neutron scattering, for example, neutrons interact with spins of electrons, and the intensity of the scattered neutrons reflect the magnetic response of electrons in the materials. This in turn carries certain information about the magnetic interaction in the materials being probed. To motivate the discussion on engineered spectroscopy, I will briefly recall the basics of spectroscopy introduced in Section 1.3 with the main focus being a discussion of inelastic neutron scattering.

The observable in inelastic neutron scattering [14, 15] is the dynamical structure factor $S(\mathbf{Q}, \omega)$, also known as the magnetic response function. As discussed in Section 1.3.3, it is

related to a two-point unequal time correlator

$$C(t, t') = \langle \hat{\mathbf{S}}_1(t)\hat{\mathbf{S}}_2(t')\rangle \,, \tag{7.1}$$

by the Fourier transform, where $\hat{\mathbf{S}}$ is a spin operator. This can be regarded as a special case of Eq. (1.8), and thus the discussion in Section 1.3.3 flows naturally from this observation.

The two-point unequal time correlator contains spectral information on the spin dynamics of a many-body system. In a translationally invariant system, the dynamical structure factor $S(\mathbf{Q}, \omega)$ reaches its local maximum when the energy selection rule, as well as the momentum selection rule, are both satisfied, as indicated by Eq. (1.10). In a spectroscopic experiment, consequently, we usually track the peak of intensities in the neutron scattering spectrum, from which we can infer the energy dispersion.

Nevertheless, we should note that there are several constraints to conventional spectroscopy experiments. As discussed in some detail in Chapter 1, the allowed degrees of freedom in conventional spectroscopy experiments are limited, so that we can only access the intrinsic property of materials, instead of an engineered system with general interactions. It is worth reiterating that some degrees of freedom in engineering the system are possible in spectroscopy experiments, such as through the application of an external electric or magnetic field. However, these external fields usually contribute only single-body local terms, instead of many-body interactions, to the Hamiltonian of the system being probed. In addition, the conclusion reached in the discussion in Section 1.3.3 holds under two conditions: (1) the perturbation is weak, and the linear response theory holds, (2) the perturbed system is in an equilibrium state. The second condition indicates that there is no coherence of the initial state, as ρ is diagonal in the eigenbases of the Hamiltonian, $\rho^{nn'} = 0$ for $n \neq n'$, and thus we can only probe the properties in the equilibrium phase. This raises the following questions: (1) can we explore the spectroscopic properties of a general quantum many-body system, and (2) can we explore the properties in the out-of-equilibrium phase. To be more specific in relation to the tasks, spectroscopic properties are referred to as transition energies between eigenstates of the Hamiltonian and the excitation spectrum from the ground state, which are the central objectives in this chapter.

Readers may find that this problem rests within a subset of the problem that was posed in Section 1.2.1. Although it is an out-of-equilibrium problem, the key component of the task considered here is to obtain the eigenstates and the corresponding eigenenergies, which can be achieved through using the quantum algorithmic cooling method developed in Chapter 4 or other eigenstate preparation methods introduced in Chapter 2. The dynamical behaviours can subsequently be obtained by applying the dynamics simulation methods developed in this thesis. A combination of these methods would consequently provide us with a result. Nevertheless, applying the above methods to the study of realistic materials with current quantum hardware could be challenging. Capturing the collective behaviour in a realistic material presents a significant

challenge given the many degrees of freedom involved in the problem description, as detailed in previous sections. More importantly, simply in terms of efficiency, having an experiment-friendly method to access the behaviours of materials, without requiring too many experimental resources, is always desirable. This is even more crucial in the noisy intermediate-scale quantum era, in which available quantum hardware has limited practicality, a point which was underscored in Section 1.4.

Spectroscopic techniques [16,17,19–21,221] provide valuable insights for us to design resource-efficient methods. Motivated by spectroscopy, I demonstrate how to probe spectroscopic features of a quantum many-body system in an efficient way. Our method only requires the realisation of time evolution e^{-iHt} without reliance on any ancillary qubits. This is in contrast to many algorithms, such as algorithmic cooling developed in Chapter 4 and variational dynamics simulation outlined in Section 2.2.3, which usually require controlled-unitary operations. Our method is therefore compatible with an analogue quantum simulator, and has the advantage of potentially being more robust against noise. The quantum circuit complexity of our method for transition energy estimation is shown to be logarithmic in precision, which shows a clear advantage over existing works in this estimation task.

7.2 Methods

7.2.1 Framework

In this section, I discuss how to estimate transition energies between the eigenstates of a quantum many-body system using engineered spectroscopic methods. I first formulate this problem in a similar fashion to that in Chapter 4. To begin with, let us consider a quantum channel

$$\mathcal{G}(\rho,\omega) = \sum_{n,n'=0} |n'\rangle \langle n'| \rho |n\rangle \langle n| p_\tau(E_{n'} - E_n - \omega), \tag{7.2}$$

where $|n\rangle$ is the eigenbasis of the Hamiltonian satisfying $H|n\rangle = E_n|n\rangle$. The function $p_\tau(\cdot)$ that selects the energy difference between $|n'\rangle$ and $|n\rangle$ is introduced. For instance, we may choose the Gaussian function $p_\tau(\omega) = \exp(-\tau^2\omega^2)$ which decays exponentially fast with respect to τ or ω. Here, we arrange that $p_\tau(\omega) = p(\tau\omega)$, such that τ is coupled with ω.

With a properly selected observable $\hat{O}$, we can obtain the measurement outcome $G(\rho,\omega) = \mathrm{Tr}[\mathcal{G}(\rho,\omega)\hat{O}]$. Denote $\Gamma_{n',n} := \rho^{n'n}\langle n|\hat{O}|n'\rangle$ which represents the state-and-observable dependent coherence, yet is time-independent. We have

$$\begin{aligned} G(\rho,\omega) &= \sum_{n,n'=0} \mathrm{Tr}[\hat{O}|n'\rangle\langle n'|\rho|n\rangle\langle n| p(\tau(E_{n'} - E_n - \omega))] \\ &:= \sum_{n,n'=0} \Gamma_{n',n} p(\tau(E_{n'} - E_n - \omega)). \end{aligned} \tag{7.3}$$

We can find that the quantity $G(\rho,\omega)$ contains the information on transition energies. Specifically, given a proper $p(\cdot)$, $G(\rho,\omega)$ takes its local maximum when ω approaches to $E_{n'} - E_n$, from which

we can see that $p(\cdot)$ plays a role as a spectral filter that filters $E_{n'} - E_n$ out from the other transition energies.

Next, I show how to estimate $G(\rho, \omega)$ in Eq. (7.3). The key idea is to consider a Fourier transform of $G(\rho, \omega)$. Let us consider the dual form of p via its Fourier transform

$$\tilde{g}(t) := \int_{\infty}^{\infty} p(\omega) e^{i\tau\omega t} d(\tau\omega), \tag{7.4}$$

and its inverse form

$$p(\tau\omega) = \frac{1}{2\pi} \int_{\infty}^{\infty} \tilde{g}(t) e^{-i\tau\omega t} dt. \tag{7.5}$$

Consider the normalised function $g(t) = |\tilde{g}(t)|/\|\tilde{g}(t)\|$ with $\|\tilde{g}(t)\| := \int_{\infty}^{\infty} |\tilde{g}(t)| dt$, and we have

$$g(t) = \frac{1}{c} \int_{\infty}^{\infty} p(\omega) e^{i\tau\omega t} d(\tau\omega), \tag{7.6}$$

and its inverse form

$$p(\tau\omega) = \frac{c}{2\pi} \int_{\infty}^{\infty} g(t) e^{i\theta_t} e^{-i\tau\omega t} dt. \tag{7.7}$$

with $c := \|\tilde{g}(t)\|$ and the phase $e^{i\theta_t} := \tilde{g}(t)/|\tilde{g}(t)|$.

Plugging the Fourier transform of p into Eq. (7.3), we find that

$$\begin{aligned} G(\rho, \omega) &= \frac{c}{2\pi} \sum_{n,n'=0} \Gamma_{n',n} \int_{-\infty}^{\infty} e^{-i\tau(E_{n'} - E_n - \omega)t} g(t) e^{i\theta_t} dt \\ &= \frac{c}{2\pi} \sum_{n,n'=0} \mathrm{Tr}[\hat{O} \int_{-\infty}^{\infty} e^{-i\tau E_{n'} t} |n\rangle \langle n| \rho |n'\rangle \langle n'| e^{i\tau E_n t}] g(t) e^{i\theta_t} e^{i\tau\omega t} dt \\ &= \frac{c}{2\pi} \int_{-\infty}^{\infty} \mathrm{Tr}[\hat{O}\rho(\tau t)] g(t) e^{i\theta_t} e^{i\tau\omega t} dt = \frac{c}{2\pi} \int_{-\infty}^{\infty} G(\tau t) g(t) e^{i\theta_t} e^{i\tau\omega t} dt \end{aligned} \tag{7.8}$$

where we denote

$$G(t) := \mathrm{Tr}[\hat{O}\rho(t)], \tag{7.9}$$

in the Schrödinger picture. The above equation indicates that we can first obtain $\mathrm{Tr}[\hat{O}(\tau t)\rho]$ by measuring $\hat{O}$ on the time-evolved quantum state at time τt, and then use Eq. (7.8) to obtain $G(\rho, \omega)$. Since $g(t)$ is normalised and hence can be regarded a probability distribution, we can estimate $G(\rho, \omega)$ by

$$\hat{G}(\rho, \omega) = \frac{1}{N_s} \sum_{i=1}^{N_s} \hat{g}_i(\rho, \omega), \tag{7.10}$$

where N_s is the total number of samples, and $\hat{g}_i$ is a single-shot estimator, which takes the form of

$$\hat{g}_i(\rho, \omega) = \frac{c}{2\pi} \hat{o}(\tau t_i) e^{i\theta_{t_i}} e^{i\tau\omega t_i}, \tag{7.11}$$

where t_i is sampled from the probability distribution $g(t)$ and $\hat{o}(\tau t_i)$ is an unbiased estimate of $\mathrm{Tr}[\hat{O}(\tau t)\rho]$. $\hat{G}(\rho, \omega)$ is an unbiased estimator of $G(\rho, \omega)$[1],

$$G(\rho, \omega) = \mathbb{E}\hat{G}(\rho, \omega),$$

[1] $\hat{g}_i(\rho, \omega)$ is also an unbiased estimator of $G(\rho, \omega)$.

where the average is over the probability distribution $g(t)$.

It is worth noting that in Eq. (7.8), we choose to treat $\tau\omega$ as a whole when we perform the Fourier transform. As such, $g(\cdot)$ is a probability distribution that is independent of τ, and the time length[2] for evaluating $G(\rho, \omega)$ is τt, which is extended by a factor τ. Alternatively, we can treat τ as a separate parameter that is independent of ω, and hence will not be Fourier-transformed. One can show that these two ways are equivalent, which will not be elaborated on in this thesis for the sake of conciseness.

Eq. (7.3) indicates that to be able to infer the transition energy $\Delta_{n',n} := E_{n'} - E_n$, there are two necessary requirements: (1) a sufficiently large coherence $\Gamma_{n',n}$, and (2) a proper function $p(\tau\omega)$ (or equivalently its dual form $g(t)$) that ensures that $\Delta_{n',n}$ can be distinguished from other transition energies.

The coherence $\Gamma_{n',n}$ is time-independent yet dependent on the state and the chosen observable. We can consider the following strategy in order to satisfy the first condition. We first prepare the initial state ρ as the ground state of the noninteracting system governed by $\hat{H}_0$. The interaction $\hat{H}_1$ at $t = 0$ is then suddenly turned on. The state will be evolved under the Hamiltonian $\hat{H} = \hat{H}_0 + \hat{H}_1$. In a weakly coupled regime, where the initial state ρ is close to the ground state $|0\rangle$ of $\hat{H}$, the initial state can be expanded using the first-order perturbation as

$$\rho \simeq \rho^{00}|0\rangle\langle 0| + \sum_{n\neq 0} \rho^{0n}|0\rangle\langle n| + \rho^{n0}|n\rangle\langle 0|. \tag{7.12}$$

The state coherence ρ^{n0} is nonzero, which indicates that it allows a transition between the eigenstate $|n\rangle$ and the ground state $|0\rangle$, and consequently we can in principle detect the energy difference $E_n - E_0$.

It can be observed that the task of detecting transition energies is very similar to the problem of searching for eigenenergies which was discussed in Chapter 4. However, there are several differences between the method developed in this chapter and the algorithmic cooling which was introduced in Chapter 4. In terms of implementation, this spectroscopic method does not require controlled unitary operations and is thus free of ancillary qubits. It only requires the realisation of real-time dynamics $e^{-i\hat{H}t}$, a basic and most promising application of quantum computing [13,37]. It is also therefore applicable to most analogue quantum simulators and could be more robust against noise.

It is also worth noting the relation between our method and conventional spectroscopy techniques, since both are capable of obtaining the energy spectrum of a quantum system. The essential component of our method is to obtain the expectation value of a properly chosen $\hat{O}$ on a time-evolved state, and it relies on the realisation of real-time dynamics $e^{-i\hat{H}t_i}$ with different times t_i, which will be implemented on a quantum simulator. As we will find in Section 7.2.2, one advantage of our method is that it only requires evolution with a short time length which

[2]The maximal time length determines the time complexity of the algorithm.

is logarithmic in the target precision ε. In addition, since the Hamiltonian is capable of being engineered, we can engineer quantum simulators to detect the energy excitation spectra of many different types of quantum many-body systems.

7.2.2 Error analysis and resource requirement

In this section, I discuss the computational complexity of our method. I focus on the above weakly coupled case, in which the initial state is populated by a collection of low-lying excited states, and it is assumed that $\Gamma_{n,0}$ is nonvanishing while $\Gamma_{n,n'\geq 1} = 0$. In this context, 'nonvanishing' indicates that the coherence $\Gamma_{n'n}$ of the eigenstates is a constant or can be lower-bounded by a polynomial of the inverse of the system size N as

$$\Gamma_{n'n} \geq \mathcal{O}(\mathrm{Poly}(\frac{1}{N})). \tag{7.13}$$

The objective is to estimate the transition energy $\Delta_{j,0} = E_j - E_0$ within an error ε, i.e., $|\hat{\Delta}_{j,0} - \Delta_{j,0}| \leq \varepsilon$. For the sake of convenience, we shift the full energy spectra by a constant E_0, such that the shifted ground state energy is zero, $E_0 = 0$. As such, finding the transition energy is converted to an eigenenergy problem, and Eq. (7.3) is simplified as

$$G(\rho, \omega) = \sum_n \Gamma_{n,0} p(\tau(E_n - \omega)). \tag{7.14}$$

Generally speaking, the eigenenergy E_j can be determined by searching peaks of $G(\rho, \omega)$ over the frequency domain ω. Under the assumption,

$$\Gamma_{j,0} p(\tau(E_j - \omega)) \gg \Gamma_{i,0} p(\tau(E_i - \omega)), \forall i \neq j \tag{7.15}$$

the eigenenergy could be determined by

$$E_j = \mathrm{argmax}_{\omega \in [E_j^L, E_j^U]} G(\rho, \omega). \tag{7.16}$$

This assumption, which is suggested in relation to the discussion in Chapter 4, holds when the target precision is sufficiently small $\varepsilon \ll \Delta_j$ with $\Delta_j := \min\{E_j - E_{j-1}, E_{j+1} - E_j\}$, and τ is sufficiently large. For example, we can consider the Gaussian function $p(\tau\omega) = e^{-\tau^2\omega^2}$, and set $\tau \geq \varepsilon^{-1}$. Using the results detailed in Chapter 4, the maximal time complexity is $\tilde{\mathcal{O}}(\varepsilon^{-1})$, where polylogarithmic dependence is hidden within the big-$\mathcal{O}$ notation.

Although the total circuit complexity (maximal evolution time $\times$ sampling numbers) reaches the Heisenberg limit for eigenenergy estimation, which is still $\tilde{\mathcal{O}}(\varepsilon^{-1})$ as proven in Chapter 4, it is sub-optimal for the maximal time complexity which is more important for practical implementation with near-term devices due to a relatively short decoherence time. In this chapter, I will demonstrate that in a situation where the required simulation accuracy is sufficiently small, i.e., $\varepsilon \ll \Delta_j$, such a large $\tau \sim \varepsilon^{-1}$ is unnecessary. Indeed, a relatively small τ that is of the order of Δ_j suffices to estimate the eigenenergy.

To see this point, let us rewrite $G(\rho,\omega)$ as

$$G(\rho,\omega) = \Gamma_{j,0}p(\tau(E_j-\omega)) + \sum_{i\neq j}\Gamma_{i,0}p(\tau(E_i-\omega)). \tag{7.17}$$

Given a large τ, the first term will be dominant to $G(\rho,\omega)$. The crucial step is to show that $\tau = \tilde{\mathcal{O}}(\Delta_j^{-1})$ suffices to guarantee that $G(\rho,\omega)$ is close to the peak value $\Gamma_{j,0}$ only if ω is close to E_j. In addition, in the vicinity of E_j, $\partial^2 G(\rho,\omega)/\partial^2\omega < 0$. The eigenenergy can thus be estimated by finding the peak of the estimate $\hat{G}(\rho,\omega)$

$$\hat{E}_j = \mathrm{argmax}_{\omega\in[E_j^L,E_j^U]}\hat{G}(\rho,\omega). \tag{7.18}$$

Intuitively, a time length of $\tau = \mathcal{O}(\Delta_j^{-1})$ is sufficient to suppress contributions from the other eigenstates. As a complete proof would entail the inclusion of many technical details, here I only show an intermediate result. For $\tau = \mathcal{O}(\Delta_j^{-1})$ and $\varepsilon \ll \Delta_j$, one can show that (1) $|D_\tau(\omega) - p_j| < c_1\tau^2\varepsilon^2,\ \forall|\omega - E_j| \leq 0.5\varepsilon$; (2) $|D_\tau(\omega) - p_j| > c_2\tau^2\varepsilon^2$, $\forall|\omega - E_j| \in (\varepsilon, \mathcal{O}(\Delta_j))$. This indicates that the distance $d = |\Gamma_{j,0} - G(\rho,\omega)|$ can be modulated by the estimation error $|\omega - E_j|$, and consequently, the eigenenergy E_j can be distinguished from the other eigenenergies. Here, c_1 and c_2 are some constants that are irrelevant of τ and ε yet are dependent on $\Gamma_{j,0}$. It is assumed that $\Gamma_{j,0}$ is nonvanishing, otherwise we cannot see the peak when ω approaches E_j. In the following discussion, $\Gamma_{j,0}$ is assumed to be a constant and will thus not be explicitly shown in the complexity analysis.

Given the theoretical guarantees, we first get an estimate of $G(\rho,\omega)$ by Eq. (7.10), and E_j is then determined by finding the peak of $\hat{G}(\rho,\omega)$ over the frequency domain ω. Here, it is worth noting that since we only need to measure $\mathrm{Tr}[\hat{O}\rho(\tau t)]$ in Eq. (7.10), calculating $\hat{G}(\rho,\omega)$ as a function of ω is a task involving purely classical computing, and does not cost any quantum resources. Therefore, we can locate the position of eigenenergy E_j, which corresponds to a local maximum of $G(\rho,\omega)$ and is distinguishable from other eigenenergies.

The algorithmic complexity concerning a finite τ, a finite cutoff for the integral, and a finite number of measurements is shown below.

Proposition 6. *Suppose that a system is weakly coupled, in which the initial state is close to the ground state, and that there is a sufficiently large coherence between the excited state $|j\rangle$ and the ground state $|0\rangle$. To guarantee that an estimation $\hat{\Delta}_{j,0}$ of the transition energy $\Delta_{j,0}$ is close to the true value within an error ε, we require that the maximum time is $\mathcal{O}(\Delta^{-1}\mathrm{Polylog}(1/\varepsilon))$, and the total running time is $\mathcal{O}(\Delta^3\varepsilon^{-4}\log(1/\varepsilon))$, where Δ is a chosen lower bound of the gap $\Delta \leq \Delta_j := \min\{E_{j+1} - E_j, E_j - E_{j-1}\}$.*

Proposition 6 indicates that when the required simulation accuracy ε is much less than Δ, the circuit complexity is exponentially improved with respect to ε. Nevertheless, the total running time scales as $\tilde{\mathcal{O}}(\varepsilon^{-4})$, in contrast to that established in Chapter 4 which is $\tilde{\mathcal{O}}(\varepsilon^{-1})$. The cost

of a large total running time can be understood as follows. Recall that the target is to estimate $\Delta_{j,0}$ with accuracy ε, i.e., $|\hat{\Delta}_{j,0} - \Delta_{j,0}| \leq \varepsilon$. By the Taylor expansion, we have

$$|G(\rho, \hat{\Delta}_{j,0}) - G(\rho, \Delta_{j,0})| = \Gamma_{j,0}(\tau^2\varepsilon^2) + \mathcal{O}((\tau\varepsilon)^4). \tag{7.19}$$

Due to a finite sampling number, we can only have an estimate of $G(\rho, \hat{\Delta}_{j,0})$ and $G(\rho, \Delta_{j,0})$. The sampling number required to achieve accuracy $\tilde{\varepsilon}$ scales as $\mathcal{O}(\tilde{\varepsilon}^{-2})$. Here, we choose to set $\tilde{\varepsilon} = (\tau\varepsilon)^2$ to ensure that the energy estimation error is up to ε. Consequently, the sampling number scales as $\mathcal{O}(\Delta^4\varepsilon^{-4})$, and the total running time is $\mathcal{O}(\Delta^3\varepsilon^{-4})$.

A remaining issue is the error from a finite cutoff when evaluating the integral in Eq. (7.8) with the integral range from $(-\infty, +\infty)$ to $[-T, +T]$. One can show that a finite cutoff only contributes to a logarithmic factor to the circuit complexity. This analysis is very similar to that in the algorithmic cooling method in Chapter 4, and I direct the interested reader to Section B.1.2 for details.

Proposition 6 also indicates when we choose to set $\Delta = \varepsilon$, our result is reduced to algorithmic cooling (Theorem 1), of which the total running time reaches $\mathcal{O}(\varepsilon^{-1})$, the so-called the Heisenberg limit. Recently, Wang *it al.* proposed quantum algorithms for ground state energy estimation which achieves a similar result [222]. More specifically, they proposed using a Gaussian derivative function in the form of $p_\sigma(x) = -\frac{1}{\sqrt{2\pi\sigma^3}} x \exp(-\frac{x^2}{2\sigma^2})$ as a filter to estimate the ground state energy, where σ plays a similar role to τ^{-1}. They achieved a maximal time complexity which is logarithmic in ε and a total running rime $\tilde{\mathcal{O}}(\varepsilon^{-2})$. However, this method can only be used to estimate the ground state energy instead of transition energies because the convolution function used in [222], which can be regarded as a modified $G(\rho, \omega)$, will be close to zero instead of reaching its maximum when ω approaches E_j.

7.2.3 Engineered spectroscopy

The above result rests upon the assumption that the coherence between the two eigenstates $|n\rangle$ and $|n'\rangle$ with respect to an observable $\hat{O}$ is large. The choice of the initial state and observable is therefore crucial for observing the transitions. Thus far, several works have discussed how to probe excitation spectra of a quantum many-body system by engineering the system and monitoring the dynamics of observables [16,17,19–21,221]. These works provide insights into the selection of the initial state and observables. In particular, it has been shown that nonequilibrium dynamics after a global quench [16,19] or a local quench [20] is sufficient to unveil the excitation spectrum, which has been termed quench spectroscopy. The basic idea is that quench will drive the initial stationary state out of equilibrium and generate low-lying quasiparticle excitations, the dispersion relation of which can be obtained by measuring a properly chosen observable. For

instance, the basic protocol of local quench spectroscopy[3] for a lattice model with translation invariance is that we first initialise the system in its ground state, then apply a local operation $\hat{L}$ to a single lattice site, and measure the dynamics of a local observable, as discussed in [20].

At the basis of these spectroscopic method is that the initial state is populated by a branch of low-lying excited states, which can be expressed as $|\psi_0\rangle = \sum_j c_j |j\rangle$ where $|j\rangle$ is an eigenstate of $\hat{H}$. The initial state, either generated by quench or local operations, can be formally expressed as

$$|\psi_0\rangle = \hat{L}\,|0\rangle\,, \tag{7.20}$$

where $|0\rangle$ is the ground state of $\hat{H}$. The ground state can be expressed as

$$|0\rangle = \hat{L}^{-1}\,|\psi_0\rangle = \sum_j \hat{L}^{-1} c_j\,|j\rangle\,. \tag{7.21}$$

The observation in relation to the excited state and the ground state can thus be expressed as

$$\langle 0|\hat{O}|n\rangle = \sum_j c_j\,\langle j|(\hat{L}^{-1})^\dagger \hat{O}|n\rangle\,. \tag{7.22}$$

In a simple case where $\hat{L}$ is a unitary, one natural choice is $\hat{L} = \hat{O}$, which provides a guarantee of nonzero observation, since $\langle 0|\hat{O}|n\rangle = c_n$.

In the case of a solid system, translation invariance is usually conserved, where the system Hamiltonian satisfies $[H, \hat{\mathbf{P}}] = 0$, where $\hat{\mathbf{P}}$ is the total momentum operator, and hence each eigenstate $|n\rangle$ has a well-defined momentum of $\mathbf{P}_n$, $\hat{\mathbf{P}}_n\,|n\rangle = \mathbf{p}_n\,|n\rangle$. Suppose we choose the observable at a given position x, which can be expressed as

$$\hat{O}(\mathbf{x}) = e^{-i\hat{\mathbf{P}}\cdot x}\hat{O}e^{+i\hat{\mathbf{P}}\cdot x} \tag{7.23}$$

where we abbreviate $\hat{O} := \hat{O}(\mathbf{0})$ for the sake of simplicity. Assume that translation invariance of the initial state[4] is broken. Taking a space Fourier transform of $G_x(\rho, t)$ in Eq. (7.3) with the observable $\hat{O}(\mathbf{x})$, we have

$$G_k(\rho,\omega) = \int d\mathbf{x} e^{-i\mathbf{k}\mathbf{x}} G_x(\rho,\omega) = 2\pi \sum_{n,n'=0} \Gamma_{n',n} p(\tau(E_{n'} - E_n - \omega))\delta(\mathbf{p}_{n'} - \mathbf{p}_n - \mathbf{k}). \tag{7.24}$$

This indicates translation invariance imposes selection rules of both energy and momentum for transition between eigenstates. This is the key element in spectroscopy experiments, where elementary excitations between eigenstates emerge when the selection rule of energy and momentum is satisfied.

[3]Note that the 'local quench' in the original paper may stretch the conventional meaning of quench. Quench usually refers to a process where parameters in the Hamiltonian are changed in time, and usually the time-scale for the change of parameters is very fast. For example, a system is prepared as an eigenstate of a Hamiltonian $\hat{H}_0$ at $t < t_0$, while at time t_0, the system is evolved dynamically under a different Hamiltonian $\hat{H}_0 + \hat{H}_1$. A more accurate description of 'local quench' in the protocol in [20] could be 'local perturbation'.

[4]For instance, translation invariance of the state after applying a local operation to a single site is broken.

In a weakly coupled system, the particle excitations induced by perturbations are restricted to a manifold of single-particle excitation. Therefore, an excited state can be understood as a single quasiparticle excitation above a vacuum state, $|n\rangle = \hat{\gamma}^\dagger_{\mathbf{q}} |0\rangle$, carrying on momentum $\mathbf{q}$, where $|0\rangle$, is the vacuum state, and $\hat{\gamma}^\dagger_{\mathbf{q}}$ is a creation operator of a quasiparticle with momentum $\mathbf{q}$. The observable can be chosen as

$$\hat{O} = \sum_{\mathbf{p}} A_{\mathbf{p}} \hat{\gamma}^\dagger_{\mathbf{p}} + A^*_{\mathbf{p}} \hat{\gamma}_{\mathbf{p}}, \tag{7.25}$$

and we have

$$\langle 0|\hat{O}|n\rangle = A_{\mathbf{q}*}. \tag{7.26}$$

This indicates that the choice of $\hat{O}$ in Eq. (7.25) enables a nonzero observation, and the excitation spectrum can thus be observed.

We can use a similar fashion to probe the transition energy between the excited states $|n\rangle$ and $|n'\rangle$. The transition can be expressed as $\langle n|\hat{O}|n'\rangle = \langle \mathbf{q}|\hat{O}|\mathbf{q}+\mathbf{k}\rangle$ where the momentum selection rule is imposed. If the excitations are restricted to a single quasiparticle manifold, we can choose an observable that conserves the quasiparticle number

$$\hat{O} = \sum_{\mathbf{p},\mathbf{p}'} A_{\mathbf{p},\mathbf{p}'} \hat{\gamma}^\dagger_{\mathbf{p}} \hat{\gamma}_{\mathbf{p}'}. \tag{7.27}$$

In this case, we can ensure that the observation is nonzero since $\langle n|\hat{O}|n'\rangle = A_{\mathbf{q},\mathbf{q}+\mathbf{k}}$.

It is worth noting that a priori knowledge of the system is still required to prepare the initial state which is composed of the desired excitations. Nonetheless, several spectroscopy protocols have been demonstrated that can successfully create excitations in many quantum systems, such as Bose-Hubbard models [19], spin chains [16–18], and disordered systems [221]. Below, I briefly discuss several representative spectroscopy protocols [16–18].

Let us consider a one-dimensional transverse field Ising model with,

$$H = \sum_{i<j\leq N} J_{ij} \hat{\sigma}^x_i \hat{\sigma}^x_j + B \sum_{j\leq N} \hat{\sigma}^z_j, \tag{7.28}$$

where $\hat{\sigma}^\alpha_i$ ($\alpha = x, y, z$) is a Pauli operator on the ith site, J_{ij} is the strength of spin-spin coupling between the ith and jth site. Ref. [16] considered a strong field case $B \gg \max J_{ij}$, in which the energy spectrum of H is split into $N+1$ decoupled subspaces spanned by different excitation numbers[5]. They proposed observation of quasiparticle spectroscopy by engineering the initial state consisting of the particular quasiparticle excitations. More specifically, by rotating the spins on each site $|\theta_j\rangle = \cos(\theta_j)|0\rangle_j + \sin(\theta_j)|1\rangle$ where $|0\rangle_j$ represents a spin-up state, the initial state $|\psi_0\rangle = \otimes_{j=1}^N |\theta_j\rangle$ could be a good approximation of a superposition of the ground state and

[5]The Hamiltonian H conserves the total excitations numbers $\hat{n} = \sum_j (\hat{\sigma}^z_j + 1)/2$.

the eigenstate of H in the single-excitation subspace. To probe the single quasiparticle excitations E_k, the initial state is prepared as

$$|\psi_0\rangle \approx |0\rangle + \gamma\,|k\rangle\,, \tag{7.29}$$

where γ is a small constant, $|0\rangle$ is the ground state and $|k\rangle$ is the eigenstate with a momentum k[6] It is easy to verify that the state coherence $\langle 0|\psi_0\rangle\,\langle\psi_0|k\rangle = \gamma$. For probing transitions between $|k\rangle$ and $|k'\rangle$, the authors prepared the state as

$$|\psi_0\rangle \approx |0\rangle + \gamma(|k\rangle + |k'\rangle)\,, \tag{7.30}$$

in which the state coherence $\langle k|\psi_0\rangle\,\langle\psi_0|k'\rangle = \gamma^2$. Transitions between quasiparticles in higher-order excitation sectors can be obtained by preparing the state in a similar fashion.

Yoshimura *et al.* considered a time-dependent field $B = B(t)$, which is decreased from a large polarising field to a constant, to create excitations, a method termed diabatic-ramping spectroscopy [18]. The transition energies can be obtained by taking the Fourier transform of the observable dynamics. Senko *et al.* considered a similar time-dependent field $B(t) = B_0 + B_p \sin(2\pi v_p t)$ for probing the energy spectrum of a weakly coupled system. At the basis of this method is the emergence of an energy resonance between $|n\rangle$ and $|n'\rangle$ when the frequency of the external field, v_p, matches the transition energies $|\Delta_{n',n}|$ [17]. The emergence of resonance at $v_p = |\Delta_{n',n}|$ could be understood by time-dependent perturbation theory.

Finally, it is worth noting how our method differs from those of spectroscopy experiments. In spectroscopy experiments, the external injected neutrons act as a weak perturbation and the system remains in equilibrium; it probes the intrinsic properties of materials in the equilibrium state. In global or local quench spectroscopy, the eigenstates are assumed to be nearly unchanged after quench, which is similar to that in spectroscopy experiments, although the state will be out of equilibrium. However, our framework holds for solving a more general quantum many-body problem. In our method, the eigenstates are not assumed to be unchanged after applying the perturbation, as long as the coherence is sufficiently large. Although in general a large coherence cannot be guaranteed, in certain cases discussed above, we could engineer the system to probe the desired transitions.

7.3 Comments on spectroscopic methods and relations between static and dynamic problems

At the heart of engineered spectroscopy is the extraction of eigenenergy information from real-time dynamics e^{-iHt}. This relates to the central topic of this thesis, that is, the relation between

[6] As shown in [16], the eigenstates can be written as $|k\rangle = \sum_{j=1}^{N} \tilde{A}_j^k\,|1\rangle_j \otimes_{i\neq j} |0\rangle_i$ where for nearest-neighbour couplings $\tilde{A}_j^k = \sqrt{2/(N+1)}\sin(kj\pi/(N+1))$. By setting $\theta_j = \tan^{-1}(\gamma\tilde{A}_j^k)$, we have a tensor product state that is a good approximation of Eq. (7.29).

static and dynamic problems, which have been intensively discussed in Chapter 1 and Chapter 2. Based on the observation in Eq. (1.7), all the spectral information is contained in the dynamics of the state, which indicates eigenstate properties and dynamical properties are closely related to each other, and thus one can be inferred from the other. An important question arises as to how to extract the spectral information from $\langle\psi(t)|\hat{O}|\psi(t)\rangle$, which was introduced in Eq. (7.9), and serves as an essential element in our spectroscopic protocol. In this section, I discuss how to infer static properties from $\langle\psi(t)|\hat{O}|\psi(t)\rangle$, and provide some comments on the relation between static and dynamic problems.

A key element in this chapter is the generation of a time-evolved state

$$|\Phi_j\rangle = e^{-iHt_j}\,|\Psi_0\rangle$$

evolving from an initial state $|\Psi_0\rangle = \sum_n c_n\,|n\rangle$ with time length t_j. Consider a matrix element

$$S_{jj'} = \langle\Phi_{j'}|f(H)|\Phi_j\rangle, \tag{7.31}$$

and from the expansion of the initial state, the matrix element can be expressed by

$$S_{j,j'} = \sum_{nn'} c_n c_{n'}^* e^{-i(E_n t_j - E_{n'} t_{j'})}\langle n'|f(H)|n\rangle. \tag{7.32}$$

Here, we have introduced a general matrix function $f(H)$ acting on the Hamiltonian, which is defined as $f(H) := \sum_{i=0} f(E_i)\,|u_i\rangle\,\langle u_i|$. where $f(h) : \mathbb{R} \to \mathbb{C}$ is a generic continuous-variable function determining the transformation of the spectrum of the Hamiltonian. We should note that if Trotterisation is used to implement real-time evolution, it will break the translational symmetry in the time domain, and this becomes an approximation of the element.

$f(H)$ can be regarded as the observable in engineered spectroscopy. A simple choice could be $f(H) = H^k, k \in \mathbb{N}$. Note that the quantity $S_{jj'}$ can be efficiently measured by the Hadamard test on a quantum computer, or using the quantum signal processing methods [33]. Consider the kth moment of H, $f(H) = H^k$, Eq. (7.31) becomes

$$S_{j,j'}^k = \sum_n |c_n|^2 E_n^k e^{-iE_n(t_j - t_{j'})}. \tag{7.33}$$

For a fixed k, the dual form of $S_{j,j'}^k$ via a discrete Fourier transform is given by

$$\tilde{S}^k(\omega) = \sum_{j,j'} \mathcal{F}(\omega) S_{j,j'}^k = \sum_n E_n^k |c_n|^2 \sum_{j,j'} e^{-i(t_j - t_{j'})(E_n - \omega)}, \tag{7.34}$$

with $\mathcal{F}(\omega) = \sum_n \exp(+i\Delta t_{jj'}\omega)$ and $\Delta t_{jj'} = t_j - t_{j'}$. If the time interval is selected as $\Delta t_{jj'} = \tau m$ with $m = -M, -M+1, ..., M$, the matrix can now be represented as

$$\begin{aligned}
\tilde{S}^k(\omega) &= \sum_n E_n^k |c_n|^2 \tau \sum_{j,j'} e^{-i\Delta t_{jj'}(E_n - \omega)} \\
&= \sum_n E_n^k |c_n|^2 \tau \sum_{m=-M}^{M} e^{-i\tau m(E_n - \omega)} \\
&= \sum_n E_n^k |c_n|^2 \tau \left(1 + 2\cos((E_n - \omega)(T + \tau)/2)\frac{\sin((E_n - \omega)T/2)}{\sin(E_n - \omega)\tau/2)}\right)
\end{aligned} \tag{7.35}$$

with $T = M\tau$. At $\tau \to 0$, it becomes

$$\tilde{S}^k(\omega) = \sum_n E_n^k |c_n|^2 \left(\frac{\sin((E_n - \omega)T/2)}{(E_n - \omega)/2)} \right). \tag{7.36}$$

Each term in this expression takes a maximal value at $E_n = \omega$. Several constraints may be imposed as

$$(E_{\max} - E_{\min})\Delta t \sim 2\pi, \quad \Delta E t_{\max} \sim 2\pi. \tag{7.37}$$

The eigenenergies in principle could be extracted from Eq. (7.36) with proper classical post-processing.

In addition to spectroscopic methods, the time-evolved state can be used to construct a reference basis and one can use it to estimate the eigenvalues. The basic idea is to represent the original problem in a new subspace in the span of non-orthogonal states $\{|\Phi_j\rangle\}$, where $|\Phi_j\rangle$ is the basis state generated by time evolution with time length t_j, $|\Phi_j\rangle = e^{-iHt_j} |\Psi_0\rangle$, also known as the subspace expansion method. We can decompose any state in the subspace $\bigoplus_j |\Phi_j\rangle \langle\Phi_j|$ spanned by the basis states. The eigenstate of the Hamiltonian can be decomposed as a linear combination of these basis states as

$$|u_k\rangle \approx \sum_j \alpha_j |\Phi_j\rangle . \tag{7.38}$$

Recall that the original eigenvalue problem $H|u_k\rangle = E_k|u_k\rangle$ has, by exact diagonalisation, an exponential complexity in respect of the system size. Alternatively, we can solve it in a subspace in the span of non-orthogonal states $\{|\Phi_j\rangle\}$. The original eigenvalue problem is then converted to

$$\mathbf{F}(H)\alpha = E_k \mathbf{S}\alpha, \tag{7.39}$$

where α is the column vector of expansion coefficients, and the matrix element of $\mathbf{F}$ and $\mathbf{S}$ are defined as $\mathbf{F}(H)_{jj'} := \langle\Phi_j|H|\Phi_{j'}\rangle$ and $\mathbf{S}_{jj'} := \langle\Phi_j|\Phi_{j'}\rangle$, respectively. We may use different reference states in practice; for example, time-evolved states with different time lengths or states generated by hybrid tensor networks detailed in Chapter 3.

As discussed above, the spectroscopy analysis methods and subspace expansion methods provide insights for quantum computing of spectral properties using real-time dynamics, and both can be efficiently implemented on quantum hardware.

7.4 Discussion

In this chapter, I introduce an engineered spectroscopy protocol to probe spectroscopic features of quantum many-body systems. The key element of the protocol is the realisation of real-time dynamics on a quantum simulator, and the time length is proven to be logarithmic in precision, which is near-optimal for eigenenergy estimation. In terms of practical implementation, since

our protocol is ancilla-free, it could be more experimentally friendly to be implemented using the current generation of quantum simulators. The result in this chapter shows that eigenstate properties can be estimated by dynamic properties of the quantum systems. In addition, it reveals an inherent relation between the two primary classes of quantum many-body problems, i.e., static problems and dynamic problems introduced in this thesis.

To further understand why the engineered spectroscopy methods could be used to select the eigenvalues, one can show that the method established in the chapter is closely related to the cooling or projection operations introduced in Chapter 4, which effectively realises the imaginary time evolution and can thus obtain the eigenstates and the eigenvalues. The derivation and discussion can be found in Section E.2 in Appendix E. I will further show how to use our method for probing the spectroscopic features of spin Hamiltonians (a transverse-field Ising model and a J_1-J_2 model) in Chapter 8, and refer to [96] for other examples, such as the Bose-Hubbard Hamiltonian and fermionic Hamiltonians without particle conservation.

Before ending this section, I propose an interesting future direction for the exploration of unconventional superconducting phases and understanding the pairing mechanism of high-temperature superconductivity. The spin resonance mode is one important feature in unconventional superconductors, such as cuprates and iron-based superconductors. Spin resonance occurs in superconductors with different signs of energy gap, and it can be used to distinguish different pairing mechanism, such as spin-fluctuation-mediated and orbital-fluctuation-mediated pairing. This has been intensively explored both theoretically and through experiments. I refer the interested reader to Refs. [223, 224] for a review on this subject.

A key observation is the dynamical spin susceptibility χ_s, which can be calculated by

$$\chi_s(q, \mathrm{i}\omega) = \frac{1}{3}\int_0^\beta \mathrm{d}\tau \mathrm{e}^{\mathrm{i}\omega\tau} \left\langle \hat{S}(q,\tau)\hat{S}_t(-q,0)\right\rangle \tag{7.40}$$

with τ being the imaginary time and ω a Matsubara frequency. Here, $\hat{S}$ is the spin operator $\hat{S}(q) = \frac{1}{2}\sum_{k,\alpha\beta} c^\dagger_{s\alpha}(k+q)\hat{\sigma}_{\alpha\beta}c_{s\beta}(k)$. The above quantity is obtained from the Matsubara spin-spin correlation function, while it could also be calculated in the real-time domain. The spin susceptibility can be calculated using the methods developed in this thesis, and can thus be used to study the spin resonance mode in unconventional superconductors. However, it is worth noting that this generally involves many bands in practice and 2D simulation is hard, and I leave a detailed discussion and simulation for spin resonance in unconventional superconducting phases to future works.

Author contributions. This chapter is relevant to the theoretical part of a manuscript under preparation [96]. J.S. initiated the project. J.S. developed the theoretical aspects of this project with input from L.E., V.V., and A.B..

Chapter 8

Exploration of interacting physics

In this chapter, I demonstrate how the methods developed in this thesis can be applied to explore interesting quantum many-body phenomena. In Section 8.1, I first show the simulation results for interacting bosons, fermions, and quantum spin systems using the perturbative method developed in Chapter 6. I then demonstrate this method experimentally on the IBM quantum cloud and show the experimental results in Section 8.2. In Section 8.3, I show the simulation of interacting spectroscopy of representative quantum systems using the spectroscopic technique developed in Chapter 7. Section 8.1 and Section 8.2 is relevant to a published work [95] and Section 8.3 is relevant to a manuscript under preparation [96].

8.1 Probing interacting dynamics by perturbative quantum simulation

In this section, I demonstrate the concrete applications of our perturbative approach in simulating quantum dynamics of quantum many-body physics problems with operations on a small quantum simulator. I focus on the algorithm with the explicit decomposition method introduced in Section 6.3. I numerically test the perturbative quantum simulation (PQS) method through simulating several interacting physics with different topologies as examples.

Specifically, I investigate (a) the quantum walk of bosons on a one-dimensional lattice, (b) the separation of charge and spin excitations of fermions with two-dimensional topology, and (c) the correlation propagation of quantum spin systems of two clusters. Appropriate partitioning strategies are designed, in which the whole system consists of two subsystems and each subsystem consists of 8 qubits. Figure 8.1 illustrates four different topological structures and the explicit partitioning strategies considered in this thesis. In each example, the corresponding task-specific partitioning strategy of the quantum systems is presented.

I use the perturbative quantum simulation method, in particular, the explicit decomposition method introduced in Section 6.3, for dynamics simulation[1]. $8 + 1$ qubits are used to simulate

[1]Note that 'explicit decomposition method' refers to a specific perturbative method that is developed in Section 6.3, which considers a specific and explicit decomposition of a joint evolution channel. It is worth clarifying

each subsystem. By post-processing the results of the subsystems, we can obtain the simulation results of the whole quantum system. All unique features are detected just as the whole system is simulated directly; the numerical results using PQS align with those of the exact simulation as shown in the figures below, thus verifying the reliability of the method.

These numerical tests are restricted to 16 qubits at most, since exact simulation of larger quantum systems becomes exponentially costly. To benchmark our method for larger systems, I investigated a 1D 48-site spin chain with nearest-neighbour couplings, using the time-evolving block decimation (TEBD) method with matrix product states as the reference. As shown in Figure 8.6, our simulation results coincide with those of TEBD, again confirming the reliability of PQS for simulating multiple subsystems. Intriguingly, our method only needs to manipulate $8 + 1$ qubits to recover the joint evolution dynamics of the 48 qubits.

In this section, I only consider real-time evolution of small and classically tractable quantum systems for benchmarking our method. However, for all the examples considered here, since the simulation cost is independent of the interaction and initial state of the subsystems, the PQS algorithm would also work when tackling much larger subsystems with more complicated subsystem interactions. In practice, when the subsystem size is increased to around $n = 50$ qubits and general strong interaction is considered, PQS could outstrip the capabilities of classical simulation and reliably probe properties of quantum systems many times the size of the quantum processor.

In the following subsections, I show dynamics simulation of different interacting systems using maximally $8 + 1$ qubits.

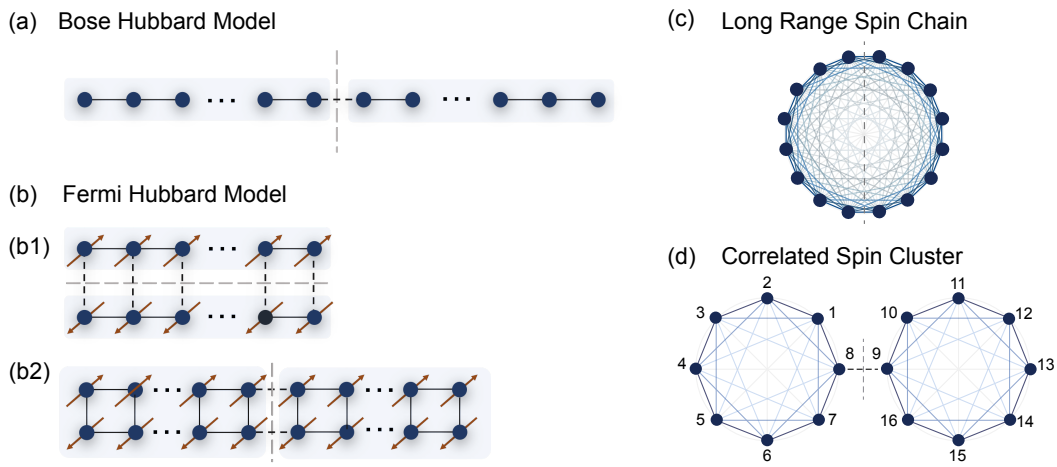


Figure 8.1: Four different topological geometries and the partitioning strategies corresponding to the Bose Hubbard model, the Fermi Hubbard model, the long-range spin chain and the correlated spin cluster considered in this work. The explicit decomposition strategy for the examples (a,b,d) (except (d)) are proven to be optimal over other perturbative expansions according to Theorem 5.

that it is irrelevant to partitioning strategies.

8.1.1 Interacting bosons

Let us first consider the physics of interacting spinless bosons on a lattice [225], which could be described by the extended Bose-Hubbard Hamiltonian

$$H = -\sum_{i,j} t_{ij} \hat{b}_i^\dagger \hat{b}_j + \frac{U}{2} \sum_i \hat{n}_i \left(\hat{n}_i - 1\right) + \sum_i h_i \hat{n}_i, \tag{8.1}$$

where $\hat{b}_i$ and $\hat{b}_i^\dagger$ are the bosonic annihilation and creation operators, $\hat{n}_i = \hat{b}_i^\dagger \hat{b}_i$ gives the number of particles on the site i, t_{ij} describes the hopping strength, U describes the on-site interaction, and h_i is the on-site chemical potential that can be tuned in various quantum systems. The model reduces to the Bose-Hubbard model H_{BHM} when only nearest-neighbour hopping is allowed, i.e., $t_{ij} = \delta_{|i-j|,1} t$. In the large U limit $U/t \to +\infty$, this model reduces to the Tomonaga-Luttinger gas Hamiltonian, which describes the collective behaviour of hard-core bosons [225]. Using the Holstein and Primakoff transformation, the Bose-Hubbard model in the large U limit is mapped onto the XX spin chain model

$$H_{\mathrm{BHM}} = J \sum_j \left(\hat{\sigma}_j^x \hat{\sigma}_{j+1}^x + \hat{\sigma}_j^y \hat{\sigma}_{j+1}^y\right) + \frac{1}{2} \sum_j h_i \left(\hat{\mathcal{I}}_j - \hat{\sigma}_j^z\right) \tag{8.2}$$

with $\hat{\sigma}_j$ representing the Pauli operator on the jth site and the effective interaction $J = -2t$. Hard-core bosons can also be related to one-dimensional free spinless fermions under the Jordan-Wigner transformation.

Quantum walks of the 1D translationally invariant bosons were experimentally demonstrated in Ref. [226]. The device system, a 12-qubit superconducting processor, can be well described by the hard-core boson Hamiltonian in Eq. (8.2). In our numerical simulation, I consider a situation in which translation invariance is broken. Specially, I consider two clusters of the interacting bosons with tuneable hopping strength $t_{ij} = t'$ on the boundary of subsystems and investigate the density distribution and correlations of bosons under time evolution.

The Hamiltonian can be expressed as $H = H_1 + H_2 + V^{\mathrm{int}}$ with the local Hamiltonian and interactions on the boundary as

$$\begin{aligned} H_l^{\mathrm{loc}} &= J \sum_j \left(\hat{\sigma}_j^x \hat{\sigma}_{j+1}^x + \hat{\sigma}_j^y \hat{\sigma}_{j+1}^y\right) + \frac{1}{2} \sum_j h_i \left(\hat{\mathcal{I}}_j - \hat{\sigma}_j^z\right), \\ V^{\mathrm{int}} &= J'(\sigma_{1,N}^x \sigma_{2,1}^x + \sigma_{1,N}^y \sigma_{2,1}^y). \end{aligned} \tag{8.3}$$

Here, $\sigma_{l,i}$ represents Pauli operators acting on the ith site of lth subsystem, and the interactions at the boundary is $J' = -2t'$. This Hamiltonian reduces to the Bose-Hubbard model when $t = t'$.

Next, I divide the whole system into two parts and simulate the dynamics of interacting bosons using our perturbative approach with the explicit decomposition introduced in Section 6.3. Our method enables the simulation of the 16-qubit problem with only $8 + 1$ qubits. It is worth noting that the explicit decomposition is optimal with respect to all possible decomposition

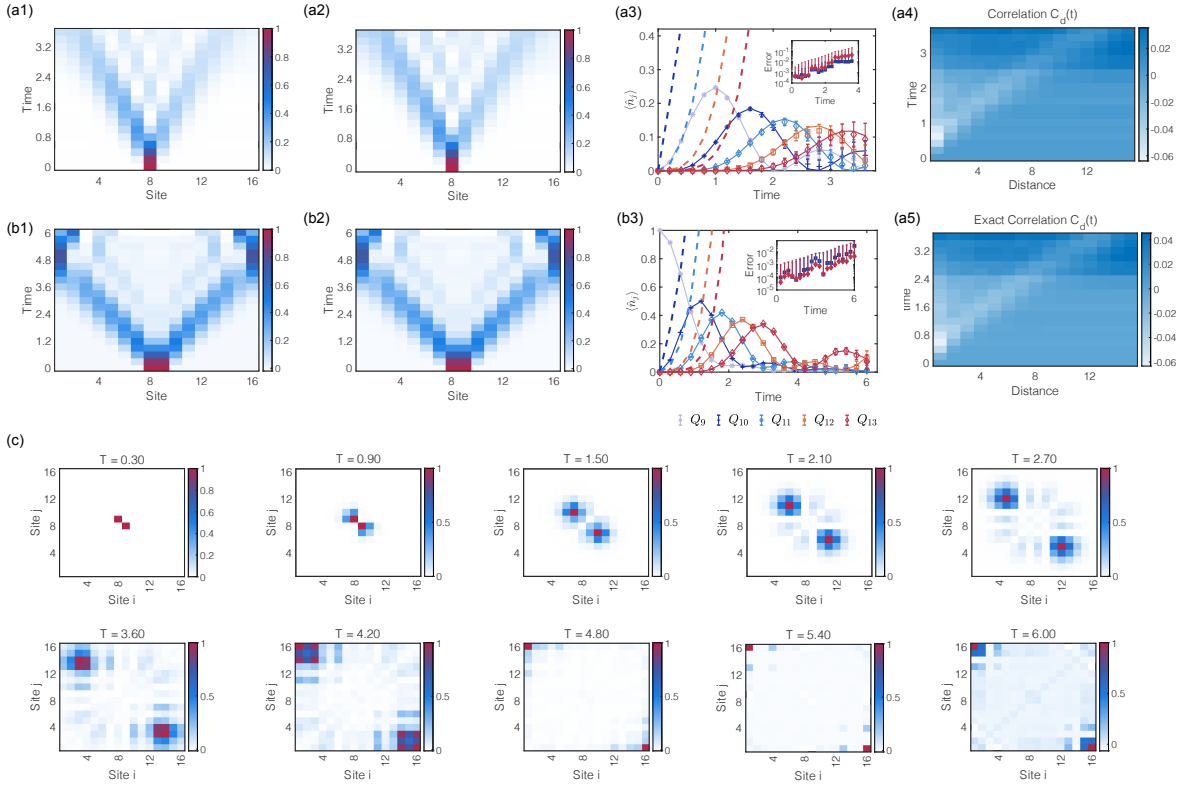


Figure 8.2: Dynamics of 16 interacting bosons on a 1D array in the large on-site repulsion limit $U/t \to \infty$. (a) Quantum walk after single particle excitation at the centre $|\psi_0\rangle = \hat{b}_8^\dagger |\mathbf{0}\rangle$. The interaction strengths are set as $J = 0.5$ and $J' = 0.8J$. (a1) and (a2) show the simulation results and exact results by exact diagonalisation (ED) for time-evolved density evolution $\hat{n}_j = \langle \hat{b}_j^\dagger \hat{b}_j \rangle$, respectively. (a3) The density distribution $\hat{n}_j$ at different sites Q_9 to Q_{13} under time evolution. The nearest-neighbour Lieb-Robinson bounds (dashed line) capture the maximum propagation speed of density spreading. (a4) and (a5) show the evolution of the averaged two-body correlation functions $\bar{C}_d(t) = \frac{1}{N-d}\sum_{j=1}^{N-d} C_{j,j+d}(t)$, which exhibit similar light cone propagation. The inset figure in (a3) shows the errors for the density and the averaged two-body correlation functions. (b) Quantum walk after two particle excitations at the centre $|\psi_0\rangle = \hat{b}_8^\dagger \hat{b}_9^\dagger |\mathbf{0}\rangle$. The interaction strengths are set as $J = 0.5$ and $J' = 0.5J$. (b1) and (b2) show the simulated results and exact results for time-evolved density evolution $\hat{n}_j = \langle \hat{b}_j^\dagger \hat{b}_j \rangle$, respectively. (b3) The density distribution $\hat{n}_j$ at different sites Q_9 to Q_{13} under time evolution. The nearest-neighbour Lieb-Robinson bounds are shown by the dashed line. The inset figure in (b3) shows the errors for the density and the two-body density-density correlation functions $\hat{\rho}_{ij} = \langle \hat{b}_i^\dagger \hat{b}_j^\dagger \hat{b}_i \hat{b}_j \rangle$. (c) Spatial anti-bunching and fermionisation in the quantum walk of two indistinguishable bosons. The two bosons are excited at the centre. The normalised density-density correlation functions $\hat{\rho}_{ij}/\hat{\rho}_{ij}^{\max}$ at several time T. The off-diagonal correlations appear under evolution, which shows the anti-bunching and fermionisation of strongly correlated bosons. This phenomenon is well captured by the non-interacting spinless fermions. In this numerical simulation, the sampling number is set as 5×10^5.

strategies, as proven in Theorem 5. I first demonstrate the dynamics after local perturbation under the interacting Hamiltonian. Previous works have extensively studied the propagation

speed of quantum information in quantum many-body systems with finite range interactions, which is limited by a maximal group speed, known as the Lieb-Robinson velocity v_g [59, 227]. Information that propagates faster than v_g is exponentially suppressed, which exhibits a light-cone-like information propagation analogous to the relativistic theory. One can consider a local perturbation to the initial state $|\psi_0\rangle$ as $|\psi(t=0)\rangle = O_A |\psi_0\rangle$ in the region A. As proven in Ref. [227], the change of the expectation of the observable O_B in the region B under time evolution can be bounded by

$$|\langle\psi(t)|O_B|\psi(t)\rangle - \langle\psi_0|O_B|\psi_0\rangle| = |\langle\psi|O_A^\dagger[O_B(t), O_A]|\psi\rangle| \leq \|[O_B(t), O_A]\|, \tag{8.4}$$

where $O_B(t)$ represents the operator in the Heisenberg picture. This establishes how local operations O_A affect the observables O_B under time evolution. If the interactions decrease exponentially with distance, one can bound the unequal time commutator by

$$\|[O_B(t), O_A]\| \leq C\|O_A\|\|O_B\| \exp\left[-\frac{d - v_g|t|}{\xi}\right], \tag{8.5}$$

where d is the distance in between the region A and B (shortest path connecting A and B), and c, v_g, and ξ are positive constants depending on $g = \max_{i,j} |J_{ij}|$. For the nearest-neighbour interaction, one can have a tighter bound by $|\langle\psi(t)|O_B|\psi(t)\rangle - \langle\psi_0|O_B|\psi_0\rangle| \leq I_d(4Jt)$, where d is the distance in between the site A and B, c and v are the velocity constant, and I_d is the modified Bessel function of the first kind [59]. In our simulation, the particle number is conserved, and I consider the observable as the occupation number operator $O_A = \hat{n}_j$ and local perturbation as $O_B = \prod_{j\in B} \hat{\sigma}_j^x |\psi_0\rangle$.

I now show the propagation of density distribution and non-local two-body correlations after local excitations. One boson is first excited at the centre by $|\psi_0\rangle = \hat{b}_8^\dagger |\mathbf{0}\rangle$, where $|\mathbf{0}\rangle$ is the vacuum. The density spreading of the boson under the interacting Hamiltonian with interaction strength $J = 0.5$ and $J' = 0.8J$ is shown. In Figure 8.2(a1), the evolution of density $\hat{n}_j = \langle\hat{b}_j^\dagger\hat{b}_j\rangle$ indicates a light-cone-like propagation. The propagation is well captured by the nearest-neighbour Lieb-Robinson bound (dashed line), as shown in Figure 8.2(a3). Next, I study the distribution of correlations after the single-particle excitation. I consider the averaged non-local correlations as

$$\bar{C}_d(t) = \frac{1}{N-d}\sum_{j=1}^{N-d} C_{j,j+d}(t) \tag{8.6}$$

with the two-body correlation function $C_{ij}(t) = \langle\sigma_i^z\sigma_j^z\rangle - \langle\sigma_i^z\rangle\langle\sigma_j^z\rangle$. We see the correlation grows nonlocally under evolution, and also exhibits a clear light cone propagation, as shown in Figure 8.2(a4). The exact dynamics are shown in Figure 8.2(a2, a5) for comparison.

Next, I show the strong correlation effects with two bosons excitations. The two adjacent bosons display spatial bunching effects in the non-interacting case while it gradually transforms to spatial anti-bunching in the large U case, which is similar to non-interacting spinless

fermions [228]. The fermionisation phenomenon of the 1D translationally invariant bosons in the large U limit was experimentally demonstrated in Ref. [226]. Here, I consider the correlated Hamiltonian in Eq. (8.3) with reduced interaction strength $J' = 0.5J$ on the boundary. At $t = 0$, two adjacent indistinguishable particles are excited at the centre $|\psi_0\rangle = \hat{b}_8^\dagger \hat{b}_9^\dagger |\mathbf{0}\rangle$. I show the density spreading in Figure 8.2(b1,b3), which exhibits similar propagation as the single particle excitation case. The dynamics of two particle excitation can be sensitive to the particle statistics due to interference. As proposed in Ref. [228], the fermionisation or bosonisation of the particle statistics can be distinguished by measuring the two-body density-density correlators

$$\hat{\rho}_{ij} = \langle \hat{b}_i^\dagger \hat{b}_j^\dagger \hat{b}_i \hat{b}_j \rangle \tag{8.7}$$

In Figure 8.2(c), I show the time evolution of the density operator $\hat{n}_j$ and density-density correlators of two bosons placed at the adjacent centre. The long-range anti-correlations appearing in the off-diagonal pattern reveal the fermionisation of strongly correlated bosons with reduced interaction strength. We can also see the interference pattern in Figure 8.2(c) during the evolution as an indication of interactions between the bosons.

8.1.2 Interacting fermions

In this section, I consider one-dimensional interacting fermions with spin degrees of freedom, which is described by the Fermi-Hubbard Hamiltonian as

$$H = -J \sum_{j,\sigma} \left(\hat{c}_{j,\sigma}^\dagger \hat{c}_{j+1,\sigma} + \text{h.c.} \right) + U \sum_j \hat{n}_{j,\uparrow} \hat{n}_{j,\downarrow} + \sum_{j,\sigma} h_{j,\sigma} \hat{n}_{j,\sigma} \tag{8.8}$$

where $\hat{c}_{j,\sigma}$ ($\hat{c}_{j,\sigma}^\dagger$) is the fermionic annihilation (creation) operators on the jth site with the spin state $\sigma \in \{\uparrow, \downarrow\}$, and $\hat{n}_j = \hat{c}_j^\dagger \hat{c}_j$ is the particle density operator. One-dimensional interacting fermions can be well captured by the Luttinger liquid theory, which indicates that the spin and charge of the electrons disintegrate into two separate collective excitations, spinon (holon) excitations with only spin (charge) degrees of freedom. To understand the separation of spin and charge excitations in a 1D fermionic system, I briefly review the theory of bosonisation [204] for the convenience of readers in the quantum information science community.

The Fermi surface of interacting electrons in 1D only has two points, and therefore it could be reduced to the effective Hamiltonian describing the excitation from one point to the other. The effective Hamiltonian ignoring spins can be expressed as $H = H_0 + V_{ee}$ where $H_0 = \sum_{\zeta=\pm1} \sum_q v_\mathrm{F} q \hat{c}_{\zeta q}^\dagger \hat{c}_{\zeta q}$ and $V_{ee} = \frac{1}{2L} \sum_{kk'q} V_\mathrm{ee}(q) \hat{c}_{k-q}^\dagger \hat{c}_{k'+q}^\dagger \hat{c}_{k'} \hat{c}_k$, which describes the allowed scattering near the Fermi surface. Here, $\zeta = \pm 1$ represents the left or the right side of the Fermi surface, and v_F is the Fermi velocity. For one-dimensional electrons, density modulation is the elementary excitation, and thus it is natural to introduce the bosonic operator,

$$\hat{b}_{\zeta q}^\dagger = \sqrt{\frac{2\pi}{Lq}} \sum_k \hat{c}_{\zeta,k+q}^\dagger \hat{c}_{\zeta,k}, \tag{8.9}$$

in order to map the interacting fermions to the free bosons, where L is a normalisation constant, which is the length of the system. The creation and annihilation operations of bosons satisfy the commutation relation as

$$\left[\hat{b}_{\zeta q}, \hat{b}^{\dagger}_{\zeta' q'}\right] = \delta_{\zeta\zeta'}\delta_{qq'}. \tag{8.10}$$

Therefore, the full interacting Hamiltonian can be mapped to a non-interacting Hamiltonian in terms of the bosonic operators as

$$H = \sum_{q>0,\sigma,\sigma'\zeta=\pm1} \left[v_F q \delta_{\sigma\sigma'} \hat{b}^{\dagger}_{\zeta q\sigma}\hat{b}_{\zeta q\sigma} + \frac{q g_2}{4\pi}\left(\hat{b}^{\dagger}_{\zeta q\sigma}\hat{b}^{\dagger}_{-\zeta q\sigma'} + \hat{b}_{\zeta q\sigma}\hat{b}_{-\zeta q\sigma'}\right) + \frac{q g_4}{2\pi}\hat{b}^{\dagger}_{\zeta q\sigma}\hat{b}_{\zeta q\sigma'}\right], \tag{8.11}$$

where g_2 and g_4 measure the strength of the interaction in the vicinity of the Fermi points as conventionally used in the literature, and σ denotes the spin degrees of freedom. We can write the above Hamiltonian with the bosonic operators of charges and spins $\hat{b}^{\dagger}_{\zeta qc} = \frac{1}{\sqrt{2}}\left(\hat{b}^{\dagger}_{\zeta q\uparrow} + \hat{b}^{\dagger}_{\zeta q\downarrow}\right)$ and $\hat{b}^{\dagger}_{\zeta qs} = \frac{1}{\sqrt{2}}\left(\hat{b}^{\dagger}_{\zeta q\uparrow} - \hat{b}^{\dagger}_{\zeta q\downarrow}\right)$ with c (s) denoting charge (spin). Since the Hamiltonian is quadratic, we can diagonalise the Hamiltonian by a Bogoliubov transformation,

$$H = \sum_{q>0,\zeta=\pm1} \left[v_c\left(\hat{d}^{\dagger}_{\zeta qc}\hat{d}_{\zeta qc} + \frac{1}{2}\right) + v_s \hat{d}^{\dagger}_{\zeta qs}\hat{d}_{\zeta qs}\right]. \tag{8.12}$$

which is represented in a new basis of bosons $\hat{d}$ and $\hat{d}^{\dagger}$ with the velocities $v_c = q\sqrt{\left(v_F + \frac{g_4}{2\pi}\right)^2 - \left(\frac{g_2}{2\pi}\right)^2}$ and $v_s = q v_F$. This clearly shows that the spin and charge densities have a different velocity near the Fermi surface, as predicted by the theory of Luttinger liquids. This observation has been numerically and experimentally investigated [225, 229, 230], and Arute *et al.* reported a simulation of this model using a programmable superconducting quantum processor with high gate accuracy [230].

To simulate the dynamics of interacting fermions carrying spins on a quantum computer, one can use the Jordan-Wigner transformation to map the fermionic operators $\hat{c}_j$ on each site to the qubit Pauli operators as

$$\hat{c}_j \mapsto \frac{1}{2}\left(\hat{\sigma}^x_j + i\hat{\sigma}^y_j\right)\bigotimes_{i=1}^{j-1}\hat{\sigma}^z_i, \tag{8.13}$$

with Pauli operators $\hat{\sigma}^{\alpha}_j$, $\alpha = (x, y, z)$ acting on the jth site, as introduced in Section 2.1. I consider an 8-site interacting 1D Fermi-Hubbard model, which requires $N = 16$ qubits to encode the spin state (spin up and spin down) at each site. The qubit layout is shown in Figure 8.1(b). According to the topology of the interactions, we have two partitioning strategies, by regarding either the nearest hopping or on-site Coulomb interactions as the V^{int}. Therefore, depending on the relative strength of t and U, we can cut the full interacting systems along either transverse or longitudinal directions, which are shown in Figure 8.1(b1) and (b2), respectively. I will then show how to apply our perturbative quantum simulation method to use $8 + 1$ qubits to simulate the dynamics of the 16 qubit system.

The initial state is prepared as the ground state of a non-interacting Hamiltonian. In the non-interacting limit, the Hamiltonian commutes with the total number operators $[H, \sum_j \hat{n}_{j,\sigma}] = 0$. For a one-dimensional chain, one finds that the Hamiltonian in the one-particle sector simply moves the occupied site to the left or right site, and thus can be expanded on the one-particle basis as a tridiagonal matrix. The Hamiltonian has the elements $H_{ij} = \langle i|H|j\rangle$ with $|j\rangle = \hat{c}_j^\dagger |\mathbf{0}\rangle$ and $|\mathbf{0}\rangle$ representing the vacuum. We can use a unitary transformation $U \equiv [u]_{ij}$ to diagonalise the Hamiltonian. This process is also known as a basis rotation. The Hamiltonian after a basis rotation is expressed in a new basis of fermionic operators $\hat{a}_j$ and $\hat{a}_j^\dagger$, which are referred to as rotated bases. The rotated basis is related to the original basis by the unitary transformation,

$$\hat{a}_j^\dagger = \sum_j u_{ij} \hat{c}_j^\dagger. \tag{8.14}$$

In the two-particle sector, there are $\binom{N}{2}$ basis states, and we can similarly diagonalise the matrix of Hamiltonian to obtain the eigenstates and eigenenergies. For the system with a general occupation number N_f (relatively small $N_f \leq N$), we first identify the basis states of N_f particle number, based on which we have the matrix representation of the Hamiltonian. We can then get the transformation from the original basis $\hat{c}_j^\dagger$ to the rotated basis $\hat{a}_j^\dagger$. Refs. [231,232] showed that a linear-depth circuit can be used to prepare the ground state of a non-interacting Hamiltonian. I briefly review the procedure to prepare the initial state using a linear-depth circuit. Denote the operators in the rotated basis that diagonalise the non-interacting Hamiltonian as $\hat{a}$ and $\hat{a}^\dagger$. We can apply a particle-conserving rotation U of the single particle basis to the rotated basis as $U\hat{c}_j^\dagger U^\dagger = \hat{a}_j^\dagger$. Then we obtain the ground state of the non-interacting Hamiltonian from the easy-to-prepare state as

$$|\phi\rangle = U\hat{c}_1^\dagger \cdots \hat{c}_{N_f}^\dagger |\mathbf{0}\rangle\,, \tag{8.15}$$

where $|\mathbf{0}\rangle$ is the vacuum. The two bases are related by a unitary transformation that transforms the original operators $\hat{c}$ ($\hat{c}^\dagger$) of the interacting Hamiltonian to the new operators $\hat{a}$ ($\hat{a}^\dagger$) of non-interacting Hamiltonian, $\hat{a}_i^\dagger = \sum_j u_{ij}\hat{c}_j^\dagger$ where u is a $N \times N$ matrix. The basis-change unitary is given by $U(u) = \exp\left(\sum_{ij}[\log u]_{ij}(\hat{c}_i^\dagger \hat{c}_j - \hat{c}_j^\dagger \hat{c}_i)\right)$ which can be implemented by $\mathcal{O}(N)$ depth circuits using Givens rotations in parallel [232]. In the numerical simulation, the hopping strength is set as $J = 0.5$, and the on-site interaction $U = 0.5J$ or $U = J$. The local potential for spin up is set to be a Gaussian distribution $h_{j,\uparrow} = -\lambda_\uparrow \exp\left(-\frac{(j-(L+1)/2)^2}{2\nu^2}\right)$ with $L = 8$, $\lambda_\uparrow = 4$ and $\nu = 1$ while $h_{j,\downarrow} = 0$ for spin down, which is the same as in Ref. [230] for comparison. The state is initialised with quarter filling $N_\uparrow = N_\uparrow = 2$, in which the charge and spin density are generated in the middle of the chain at $t = 0$ in Figure 8.3.

Next, the two-particle system is evolved under the Fermi-Hubbard Hamiltonian with different strengths of on-site interaction U. The charge and spin densities characterise the collective

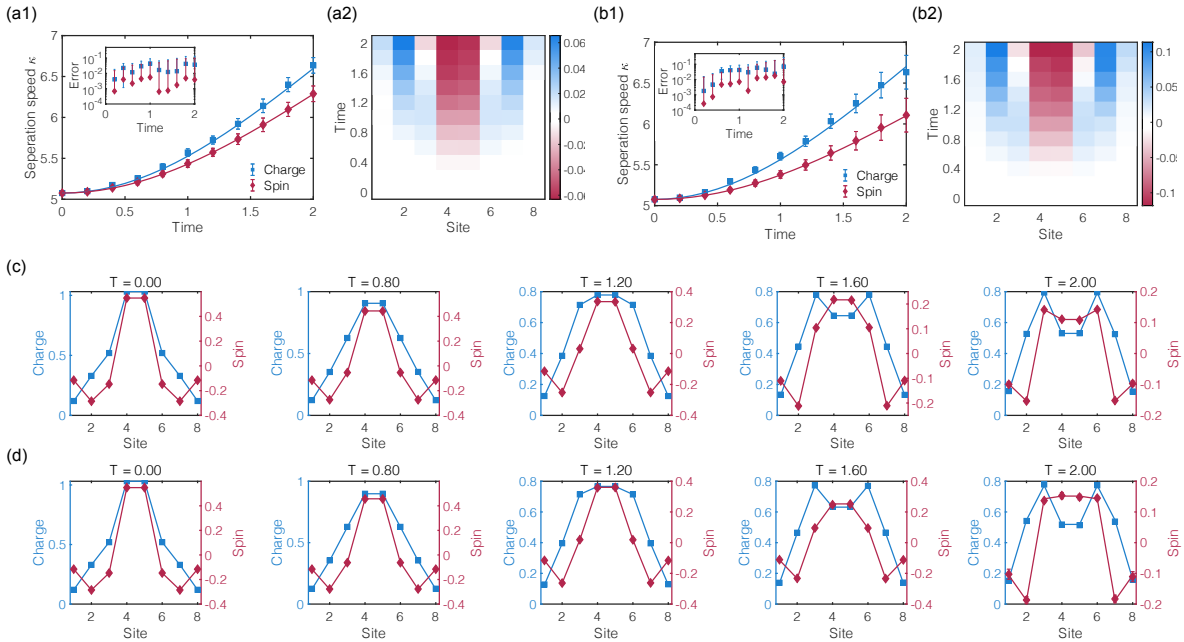


Figure 8.3: Separation of charge and spin density under the one-dimensional Fermi Hubbard model. The quantum state is initialised as the ground state of the non-interacting Hamiltonian with local potential $h_{j,\sigma}$, as specified in Section 8.1.2. I consider the dynamics with two-particle excitations, i.e. $N_\uparrow = N_\uparrow = 2$, which are generated at the middle of the chain at $t = 0$. (a1) and (b1). The time evolution of separation speed κ for charge (square and blue) and spin (diamond and red) with the interaction $U = J/2$ (a1) and $U = J$ (b1), respectively. Solid lines represent the ideal results by exact diagonalisation for comparison. The figure inset shows the errors of κ compared with the exact results over time. (a2) and (b2). The difference of charge and spin densities $\rho_j^c(t) - \rho_j^s(t) -$ const at each site with the interaction $U = J/2$ (a2) and $U = J$ (b2, respectively. The separation are offset by a constant as 0 at $t = 0$, i.e. const $= \rho_j^c(0) - \rho_j^s(0)$. (c) and (d) show the time evolution of density spreading of both charge and spin at different T for $U = J$ to $T = 2.0$. The sampling number is set as 5×10^5.

excitations, which are defined as the sum and difference of the spin-up and -down particle densities over all sites, respectively,

$$\rho_j^\eta = \langle \hat{n}_{j,\uparrow} \rangle \pm \langle \hat{n}_{j,\downarrow} \rangle \tag{8.16}$$

where $\eta = c$ or s represents charge or spin degrees of freedom. I show the density spreading of both charge and spin in Figure 8.3(c) and (d) at different t. The difference of charge and spin density is plotted in Figure 8.3(a2) and (a4) for $U = J/2$ and $U = J$, respectively. Here, the separation of charge density and spin density is offset as 0 at $t = 0$ to make the difference comparable.

The excitations spreading from the middle can be quantitatively distinguished by introducing the separation speed

$$\kappa = \sum_{j=1}^{L} |j - (L+1)/2| \, \rho_j^\eta. \tag{8.17}$$

Under time evolution, we observe a clear separation of spin density and charge density as shown in Figure 8.3(a1) and (b1). As U increases to $U = J$, the separation of spin density and charge density becomes much faster. The error for the separation speed κ_η ($\eta = c/s$) are shown in the figure inset. In the large interaction regime, the initial state is a mixture of excited states, and therefore the effective physics are unable to be well captured by the Luttinger liquid theory [204].

While this effective model for Luttinger liquids is only able to capture the low-energy excitations in a weakly coupled regime, our method can simulate the dynamics in the highly excited regime with a medium or large interaction. The Hamiltonians for these 1D interacting fermions have two parts: (1) kinetic terms due to the nearest-neighbour hopping t, and (2) on-site spin interaction U. According to the topology of interactions, we have two strategies, by regarding either on-site spin interactions (Figure 8.1(b1)) or the nearest hopping (Figure 8.1(b2)) as V^{int}. Therefore, depending on the relative strength of t and U, we can cut the full interacting systems along either transverse (Figure 8.1(b1)) or longitudinal (Figure 8.1(b2)) directions. This enables the simulation in both regimes. To prepare the general entangled state, we can decompose it into a linear combination of local states, which might introduce an additional sampling cost for the state preparation.

It is worth noting that this strategy enables quantum simulation for the two opposite regimes, which aligns with the view from the perturbation theory which applies to the weakly-interacting and strongly-interacting limit. Our method could be used to simulate the dynamics of interacting phenomena with quasi-1D or 2D geometry. In the case of the Fermi-Hubbard model considered above, the explicit decomposition strategy in PQS for both geometric layout of qubits are optimal over other perturbative expansions in terms of the resource cost for the simulation of non-local interactions, as proven by Theorem 5.

8.1.3 Quantum spin systems

8.1.3.1 Dynamical quantum phase transitions

Quantum spin models have been investigated to capture some typical emergent quantum phenomena in condensed matter, such as phase transitions and collective transitions. While many theoretical and numerical methods have been proposed to solve the effective spin models in exact or approximate solutions, a long-range spin chain with general interaction strength could be hard to solve classically. In this section, I consider a long-range spin chain, which is described by

$$H = \sum_{ij} J_{ij} \hat{\sigma}_i^z \hat{\sigma}_j^z + h \sum_j \hat{\sigma}_j^x \tag{8.18}$$

with the interactions obeying the power law decay rule $J_{ij} = J_0 |i - j|^{-\alpha}$. In this section, I study the dynamical quantum phase transitions (DQPT) [52, 233] in a long-range spin chain.

The partitioning strategy of a fully connected spin chain is shown in Figure 8.1(c). Dynamical properties of such a spin chain will be characterised by the local order parameters and the

Loschmidt amplitude. The state is first initialised as the eigenstate $|1\rangle^{\otimes n}$ of the non-interacting Hamiltonian with $h = 0$. The system is quenched by suddenly adding the transverse field h along x direction. In the limit of $\alpha = 0$, the Hamiltonian in Eq. (8.18) reduces to the Lipkin-Meshkov-Glick (LMG) model. The LMG model has an analytical solution, as it can be regarded as a classical model, in which dynamical behaviour can be predicted by a semiclassical limit. In this limit, the Hamiltonian H preserves the magnitude of the total spin and has the spin flip symmetry, i.e., $[H, \mathbf{S}^2] = 0$ and $[H, \prod_i \hat{\sigma}_i^x] = 0$. We can write the Hamiltonian as $H = \frac{J}{N}(\Sigma^z)^2 + h\Sigma^x$ using collective spin operators $\Sigma^\alpha = \sum_i \sigma_i^\alpha$ with $\alpha = x, y, z$. The spin can be represented in a mean-field approach as a classical spin vector $(\Sigma^x, \Sigma^y, \Sigma^z) = N(\cos\theta, \sin\theta\sin\phi, \sin\theta\cos\phi)$ that can be determined by the equation of motion. In Ref. [234], the authors considered a spin Hamiltonian with an external field B_z along the z direction, and showed analytically that the spatially averaged two-point correlation shows a DQPT when B_z/J_0 crosses unity. One can similarly use this analytical method to analyse the dynamical behaviour of Eq. (8.18) with small α near to zero.

Refs. [234, 235] experimentally demonstrated a DQPT and various dynamical results for the long-range spin model with α close to zero on a trapped ion platform [234] and a superconducting processor [235]. Here, I focus on the weakly coupled regime, i.e. large α, as a complementary result for comparison. In the numerical simulation, I set $J_0 = 1$, and the decay rate $\alpha = 3$ in the Hamiltonian Eq. (8.18). The full system is partitioned into 2 or 3 subsystems with each subsystem consisting of at most 8 qubits. The explicit decomposition is used to simulate the large system. Note that the explicit decomposition in this example may not be optimal, as it involves too many Pauli terms at each site. Other decomposition methods within the framework of generalised quantum operations could be numerically searched to obtain a minimal resource cost.

I first show the evolution of order parameters of a 16-site quantum spin chain. Figure 8.4(b1) and (b2) shows the magnetisation $M_z(t)$ (b1) and $M_x(t)$ (b2). We can find that $M_z(t)$ rapidly oscillate across 0 when the external field is large, while the magnetisation oscillates slowly in the low field. The motion of spin can be illustrated in a Bloch sphere in Figure 8.4(a). The order parameters of $M_z(t)$ and $M_x(t)$ provide an evidence for two phases: the ferromagnetic phase and the paramagnetic phase.

The dynamical quantum phase transitions could be observed by the Loschmidt amplitude

$$\mathcal{G}(t) = \left|\langle\psi_0|e^{-iHt}|\psi_0\rangle\right|^2 \tag{8.19}$$

as an indicator to characterise the dynamical echo back to the initial state [52, 233, 235]. A DQPT occurs with the non-analytical behaviour of a rate function

$$\gamma(t) = -N^{-1}\log\left(\mathcal{G}(t)\right) \tag{8.20}$$

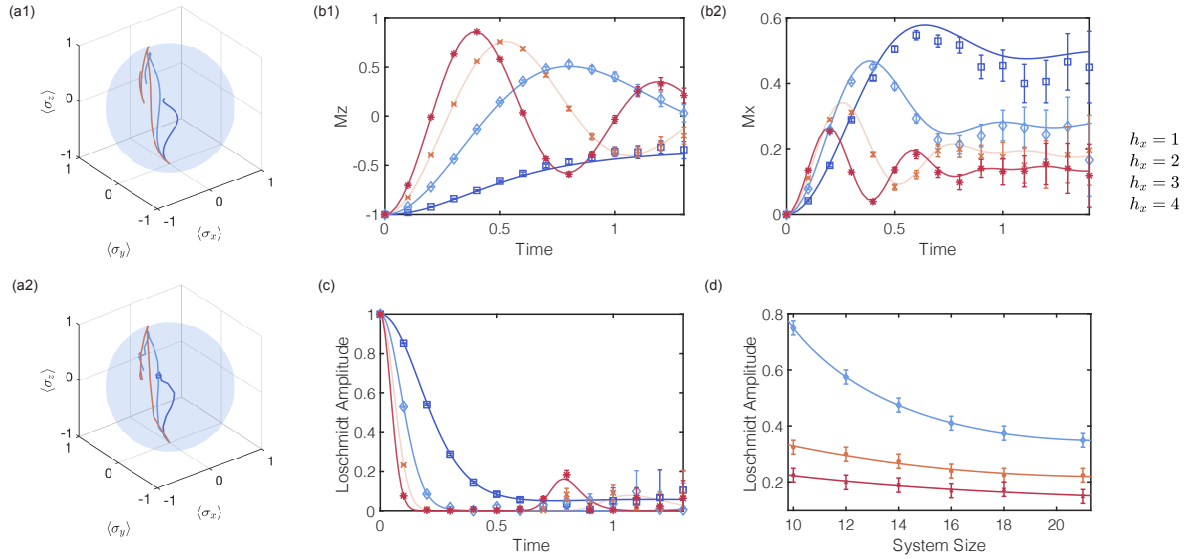


Figure 8.4: Dynamical quantum phase transition of a long-range spin chain with full connectivity and the decay rate $\alpha = 3$. The qubit layout and partitioning strategy are shown in Figure 8.1(c). The system is initialised to the eigenstate of the Hamiltonian with a zero field as $|\psi_0\rangle = |1\rangle^{\otimes N}$, and then the external field along x axis is suddenly turned on at time $t \geq 0$. (a) The numerical (a1) and ideal (a2) time evolution of the average spin magnetisation shown in the Bloch sphere for different strengths of the transverse fields $h_x = 1, 2, 3$. (b) Time evolution of the averaged magnetisation $M_z = \frac{1}{N}\sum_j \langle \sigma_j^z(t)\rangle$ (b1) and $M_x = \frac{1}{N}\sum_j \langle \sigma_j^x(t)\rangle$ (b2) for different strengths of the transverse fields $h_x = 1, 2, 3, 4$. The magnetisation M_z and M_x oscillate rapidly at a large field. The external field drives the system from the dynamical ferromagnetic phase to the dynamical paramagnetic phase. (c) The Loschmidt amplitude $\mathcal{G}(t) = \left|\langle\psi_0|e^{-iHt}|\psi_0\rangle\right|^2$, as an indicator to characterise the dynamical echo back to the initial state for different transverse field strengths h_x. (d) System size dependence of the Loschmidt amplitude. The phase transition appears earlier with larger system size [234, 235]. Solid lines represent the exact results.

in the thermodynamic limit $N \to \infty$, which can be regarded as a dynamic counterpart to a free energy density up to a normalisation N. In the LMG model, the system undergoes DQPT in the thermodynamic limit $N \to \infty$. I consider the weakly-coupled regime and present the dynamical behaviour of Loschmidt amplitude $\mathcal{G}(t)$ for different external field h in Figure 8.4(a). We clearly see that the Loschmidt amplitude rapidly decays to zero when the external field h is above the critical field. The non-analytical behaviour of the rate function $\gamma(t)$ for a large external field h reveals a dynamical phase transition to the paramagnetic phases. The minimal of Loschmidt amplitude is above zero for small h, which indicates the system persists a ferromagnetic phase under evolution. Figure 8.4(d) shows the system size dependence of minimal Loschmidt amplitude for various h_x. We can see that the minimal Loschmidt amplitude appears much earlier with an increasing system size. It is worth noting that the decay rate α of the trapped ions quantum simulator can be tuned in the region of $0 \leq \alpha \leq 3$ due to the physical interaction, while the PQS method could be leveraged to go beyond these limitations.

8.1.3.2 Propagation of correlations

The elementary excitations usually emerge from interactions, and in some cases they could be described in a quasiparticle picture. In quantum systems with finite range interactions, the quantum dynamics exhibits a light-cone-like information propagation, and the speed of information propagation is governed by the interactions of the systems. With nearest-neighbour couplings, like the results presented in Section 8.1.1, the propagation of information has a finite maximal velocity v_g, the so-called Lieb-Robinson velocity. If the interactions exponentially decay with an increasing distance, from Eq. (8.5), we know that the change of the expectation of an observable (O_B) under time evolution is exponentially decreased with the distance d, which indicates that information propagating faster than v_g is exponentially suppressed with the distance d. This light-cone-like information propagation is somewhat analogous to the relativistic theory. The speed of information propagation for power law decay interactions has been experimentally investigated in Ref. [59], which is beyond the light-cone picture. Understanding the effective model to describe quasiparticle excitations and the propagation of information for general interactions is an interesting direction of investigation. In this section, I show the quasiparticle excitations of correlated spin clusters with various interaction strengths using our algorithms.

Now, let us consider the out-of-equilibrium dynamics of a one-dimensional interacting spin system with the Hamiltonian $H = H^{\text{loc}} + V^{\text{int}}$ with the local Hamiltonian and interactions on the boundary as

$$H_l^{\text{loc}} = \sum_{i<j} J_{ij}\hat{\sigma}_{l,i}^x\hat{\sigma}_{l,j}^x + h\sum_j \hat{\sigma}_{l,j}^z, \quad V^{\text{int}} = J_0\hat{\sigma}_{1,N}^x\hat{\sigma}_{2,1}^x. \tag{8.21}$$

Here, $\hat{\sigma}_{l,i}$ represents Pauli operators acting on the ith site of lth subsystem, and the interactions obey the power law decay rule as $J_{ij} = J_0|i-j|^{-\alpha}$. In the regime of a sufficiently large field $h \gg \max(|J_{ij}|)$, the Hamiltonian conserves the total magnetisation along the z direction $M_Z = \sum_i \hat{\sigma}_i^z$. One can show that it could be reduced to the XX model $H = \sum_{ij} J_{i<j}\left(\hat{\sigma}_i^+\hat{\sigma}_j^- + \text{h.c.}\right)$ [59]. This spin model can also be mapped to a model of hard-core bosons, $H = \sum_{i<j} J_{ij}(\hat{a}_i^\dagger\hat{a}_j + \text{h.c.})$ which conserves the total particle numbers as well. For the system with continuous translational symmetry, we can Fourier transform the real-space operator into the operators that are diagonal in momentum space, written as $H = \sum_k \omega_k\hat{a}_k^\dagger\hat{a}_{-k}$ where the modes with energies ω_k have well-defined quasi-momentum k. Here, the operator $\hat{a}_k^\dagger$ creates an excitation with momentum k in the momentum space, and it is related to the original operator by $\hat{a}_k^\dagger = \sum_{i,k}\hat{a}_i^\dagger$. In our simulation, I consider a spin-cluster system and first excite the system by local perturbation, which creates a magnon quasiparticle. For the system with nearest-neighbour interactions, the energy spectrum has a well-known quadratic dispersion $\omega_k \propto k^2$ in the low energy excitation regime. For a spin cluster system, the mode does not have a well-defined momentum, but one can determine the energy dispersion ω_k provided the boundary condition and the interaction J_{ij}.

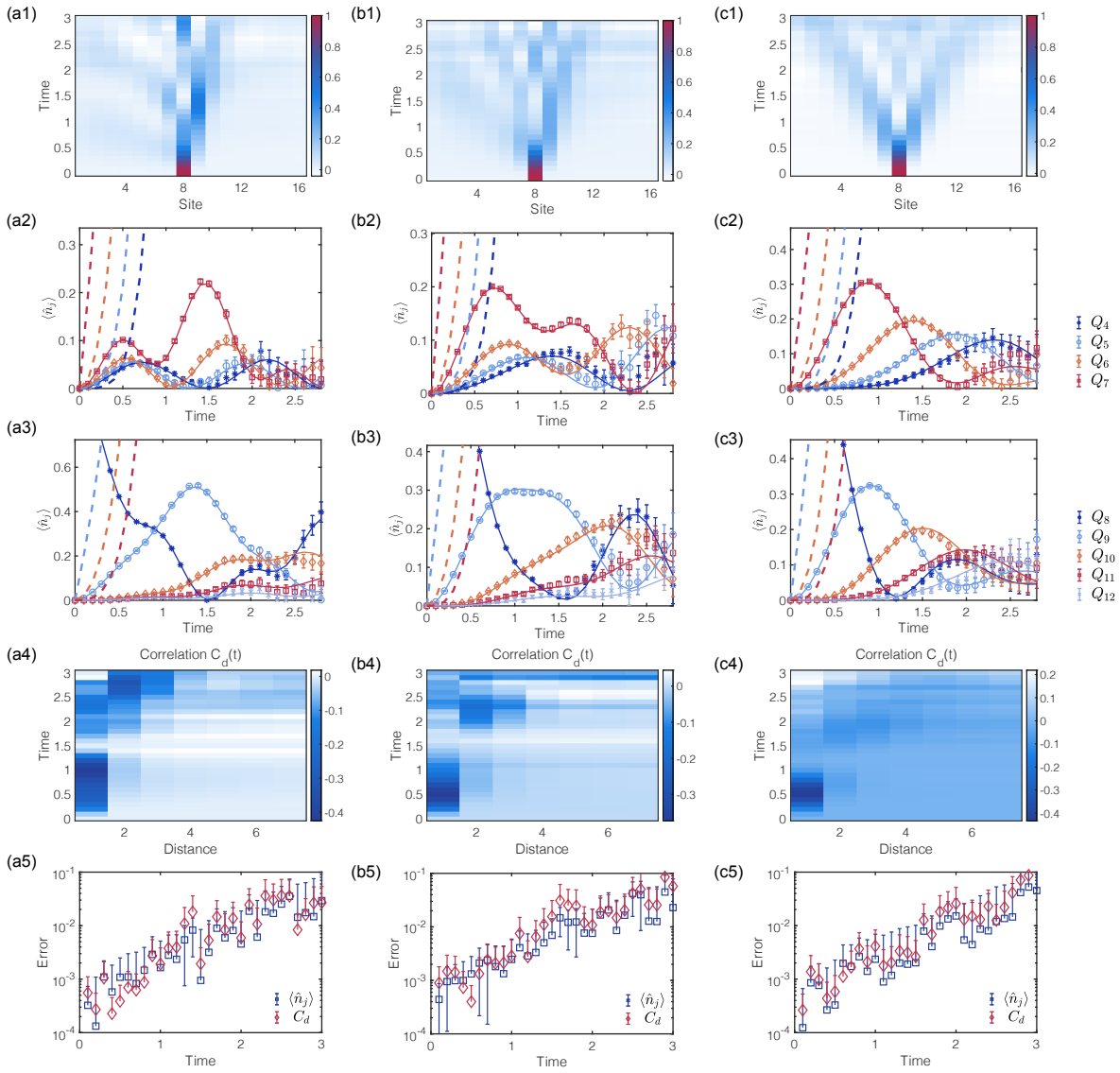


Figure 8.5: Simulation results of a correlated spin cluster (Figure 8.1(d)) with interactions on the boundary and different power law decay interactions α in the subsystems. The systems are initially perturbed at the 8th site at $t = 0$, $|\psi_0\rangle = \hat{\sigma}_8^x |\psi_0\rangle$, and suddenly the interactions between each spin with the interaction strength $J_{ij} = J_0|i-j|^{-\alpha}$ are turned on where $J_0 = 1$. (a), (b) and (c) show $\alpha = 0.5,\ 1,\ 2$, respectively. (a1), (b1) and (c1) show the dynamics of magnon quasiparticle excitations $\langle \hat{n}_j \rangle$, related to the local magnetic moment by $\langle \hat{n}_j \rangle = (1 - \langle \hat{\sigma}_j^z \rangle)/2$. (a2)-(c2) and (a3)-(c3) show the signal of the magnetisation distribution at 4th-7th sites and 8th-12th sites respectively. The nearest-neighbour Lieb-Robinson bounds (dashed lines) do not capture all the signals for this propagation. (a4)-(c4) show the averaged two-body correlation functions C_d from the 8th site. (a5)-(c5) show the errors for magnetisations and the correlation functions.

Below, I will show the simulation dynamics results of a spin cluster system. The geometry of such a spin cluster system and the partitioning strategy are shown in Figure 8.1(d). In our numerical simulation, I consider an intermediate regime where the external field is much larger

than the maximum interaction strength J_0, while being comparable to the total interaction strength $\tilde{J} = \sum_{i<j} J_{ij}$. In this case, the total magnetisation is nearly conserved; however, the field effects cannot be fully ignored. In our numerical simulation, I set $J_0 = 1$, while the external field is set as $h = 2NJ_0$ with $N = 16$ being the total sites in the full system. The system is first initialised as the eigenstate $|\psi_0\rangle = |0\rangle^{\otimes N}$ of a non-interacting H with $J_{ij} = 0$. The system is initially perturbed (at $t = 0$) at the centre (8th site) of the spin chain, with the state $|\psi_0\rangle = \hat{\sigma}_8^x |\psi_0\rangle$. Then, the system is suddenly quenched by turning on the interactions between each spins with the interaction strength J_{ij}. I show the information propagation with different decay rate $\alpha = 0.5$, $\alpha = 1$, and $\alpha = 2$ in Figure 8.5(a), (b) and (c), respectively. In our numerical experiments, I consider the magnetic moment of each spin, and set the sampling number to be 2×10^5.

The magnetic moment distributions over each site are shown in Figure 8.5(a1)-(c1), and the distribution of several neighbour sites Q_4 to Q_7 in (a2)-(c2) with three interaction strengths. We clearly see that the quasiparticle excitations in the first subsystem propagate much faster as the interaction strength increases (α decreases). This is similar to the phenomena of a 1D long range spin chain with the Hamiltonian described by Eq. (8.18), in which the maximum group velocity is predicted to show a divergent behaviour when the decay rate α approaches to zero, as has been reported in [59, 236]. From Figure 8.5, we can compare the maximum group velocity within the two clusters. The quasiparticle excitations for small α (strongly coupled) appear to be much more localised when compared to that of the weakly coupled regime. In addition, the propagation speed violates the Lieb-Robinson bounds, when considering the nearest-neighbour interaction $\max J_{ij}$ or renormalised interaction $\sum_{ij} J_{ij}$, which indicates that long-range physics cannot be well described by light-cone propagation with a finite group velocity. Nevertheless, for the other subsystem, which was unperturbed at the beginning, we observe a different propagation under time evolution, as shown in Figure 8.5(a3), (b3) and (c3). This shows an intermediate behaviour of short- and long-range physics in the spin cluster system, which might be captured by the model of nearest-neighbour interactions $\max J_{ij}$. In addition, we can also study the dynamical phase transition from the quasiparticle distribution[2].

I next present the two-body correlation functions C_d with the spin at the centre, which is expressed as

$$C_d = \langle \hat{\sigma}_j^z \hat{\sigma}_{j+d}^z \rangle - \langle \hat{\sigma}_j^z \rangle \langle \hat{\sigma}_{j+d}^z \rangle \tag{8.22}$$

with $j = 8$ at the centre in Figure 8.5(a4), (b4) and (c4), showing a quasiparticle picture explained above. Refs. [59, 236] discussed the long-range physics and short-range physics in a 1D long range spin chain with the Hamiltonian described by Eq. (8.18). Discussions on the quasiparticle

[2]Specifically, the Loschmidt amplitude can be inferred from the dynamics on the 8th site (lines of Q_8 in Figure 8.5), under the assumption of the conservation of quasiparticle numbers

propagation in a spin cluster could be complementary to that in a 1D long range spin chain (with the Hamiltonian Eq. (8.18)).

Here, we mainly focus on the regime in which the magnetisation is conserved, but we can similarly simulate the highly excited regime using the same method. In the highly excited regime, i.e., $h \sim J_0$, the quasiparticle picture does not hold, and the collective excitations could consequently be different. The investigation of the interacting physics of the spin clusters is an interesting direction.

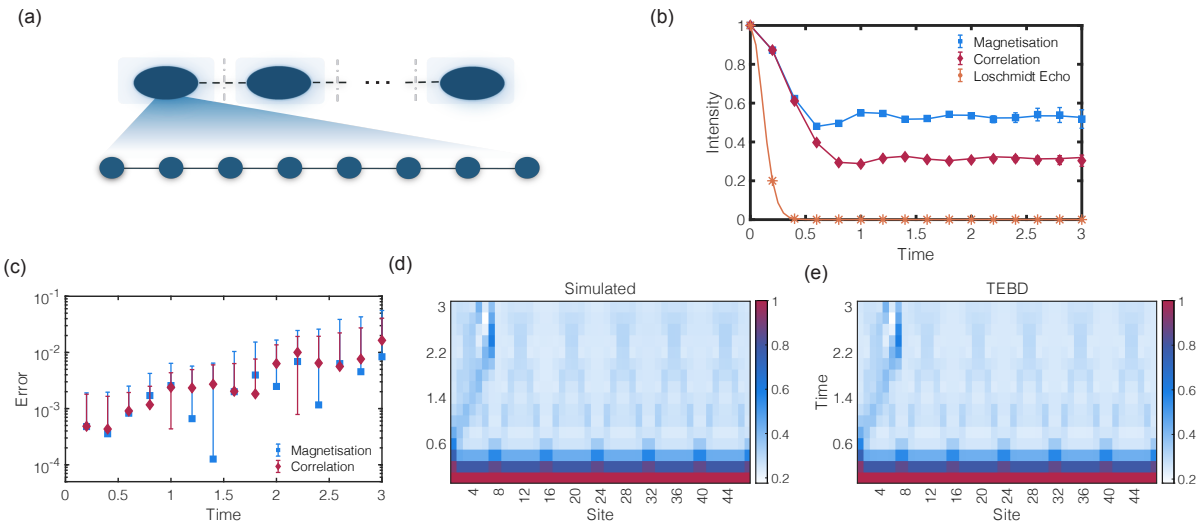


Figure 8.6: Numerical simulation of dynamics of 1D 48-site spin chains with nearest-neighbour couplings. The couplings within each cluster are identical as $J = 1$ while interactions on the boundary are randomly generated from $[0, J/2]$. (a) The partitioning sketch. We group 8 adjacent qubits within each cluster as a subsystem. (b) The averaged magnetisation $\frac{1}{N}\sum_i \hat{\sigma}_i^z$, nearest-neighbour correlation functions $\frac{1}{N-1}\sum_i \langle \hat{\sigma}_i^z \hat{\sigma}_{i+1}^z \rangle$ and Loschmidt echo $\mathcal{G}(t)$, compared with TEBD based on the matrix product state representation as a benchmark. (c) shows the errors for the averaged magnetisation and correlation. (d) and (e) shows the simulated and exact results for the long-range correlation functions $C_i = \langle \hat{\sigma}_1^z \hat{\sigma}_i^z \rangle$ from 2th-48th site, respectively.

8.1.4 Multiple subsystems

Finally, I will demonstrate that our method could be extended to simulate systems consisting of multiple clusters. I consider the out-of-equilibrium dynamics of one-dimensional interacting spin clusters with the Hamiltonian $H = H^{\text{loc}} + V^{\text{int}}$, with the local Hamiltonian and interactions on the boundary as

$$H_l^{\text{loc}} = J_l \sum_i \hat{\sigma}_{l,i}^x \hat{\sigma}_{l,i+1}^x + h \sum_i \hat{\sigma}_{l,i}^z, \quad V_l^{\text{int}} = f_l \hat{\sigma}_{l,N}^x \hat{\sigma}_{l+1,1}^x. \tag{8.23}$$

Here, $\hat{\sigma}_{l,i}$ represents Pauli operators acting on the ith site of lth subsystem. The interactions in each subsystem are identical $J_l = J_0 = 1$, while interactions between subsystems f_l are generated

randomly from $[0, 0.5]$. The external field is set as $h = 1$. The geometry of the spin system and the partitioning strategy are shown in Figure 8.6(a). In the numerical simulation, I consider a spin cluster model consisting of 6 clusters, and there are 8 spins within each cluster. Again, 8 adjacent qubits in each cluster are grouped as a subsystem, and the explicit decomposition strategy is used to simulate up to 48 qubits with operations only on $8 + 1$ qubits. Note that that according to Corollary 2, the explicit decomposition strategy for the example of Eq. (8.23) is optimal.

I show the averaged time-evolved magnetisation $M_z = \sum_i \langle \hat{\sigma}_i^z \rangle$ and the nearest-neighbour correlation function $\bar{C}_1 = \frac{1}{N-1} \sum_i \langle \hat{\sigma}_i^z \hat{\sigma}_{i+1}^z \rangle$ and long-range correlations with the first site $\bar{C}_2 = \frac{1}{N-1} \sum_i \langle \hat{\sigma}_1^z \hat{\sigma}_i^z \rangle$ and the Loschmidt echo $\mathcal{G}(t)$ in Figure 8.6(b). To benchmark our algorithms, I compare our results using the time-evolving block decimation (TEBD) method, which is a numerical method commonly employed to simulate the dynamics of quantum many-body systems based on the matrix product states formalism. Figure 8.6(c) shows that the simulation error can be achieved below 10^{-2} at an intermediate time scale.

8.2 Experimental Implementation on the IBM quantum devices

Our perturbative quantum simulation algorithm is experimentally implemented on the IBM quantum cloud. Remarkably, in contrast to direct simulation, PQS could be more robust to noise attributed to the reduction of quantum sources. In order to verify such an advantage, I used the IBM quantum cloud hardware to experimentally implement PQS, as discussed below.

8.2.1 Experimental results

Let us consider an 8-qubit one-dimensional Ising Hamiltonians

$$H = \sum_{i=1}^{7} \hat{\sigma}_i^z \hat{\sigma}_{i+1}^z + h \sum_{j=1}^{8} \hat{\sigma}_j^x, \tag{8.24}$$

with nearest-neighbour interaction and a transverse magnetic field with different strength h. Starting from an eigenstate of H with $h = 0$, we evolve the state, $|\psi(0)\rangle = |0\rangle^{\otimes 8}$, with $h = 0.5$ from time $T = 0$ to 1 and observe the dynamical quantum phase transition. At time $t \in [0, 1]$, I focus on the expectation value of the spin operator $M_z = \sum_{j=1}^{8} \hat{\sigma}_j^z / 8$ and the Loschmidt amplitude $\mathcal{G}(t) = |\langle \psi(0)| e^{-iHt} |\psi(0)\rangle|^2$, which is equivalent to evaluating the state overlap between $|\psi(0)\rangle$ and $|\psi(t)\rangle = e^{-iHt} |\psi(0)\rangle$.

To get the exact time-evolved state, I consider the Trotterisation with four time-steps. Specifically, we have

$$|\psi(t)\rangle = \left(\prod_{j=1}^{8} e^{-ih\delta t \hat{\sigma}_j^x} \prod_{j=1}^{7} e^{-i\delta t \hat{\sigma}_j^z \hat{\sigma}_{j+1}^z} \right)^{t/\delta t}, \tag{8.25}$$

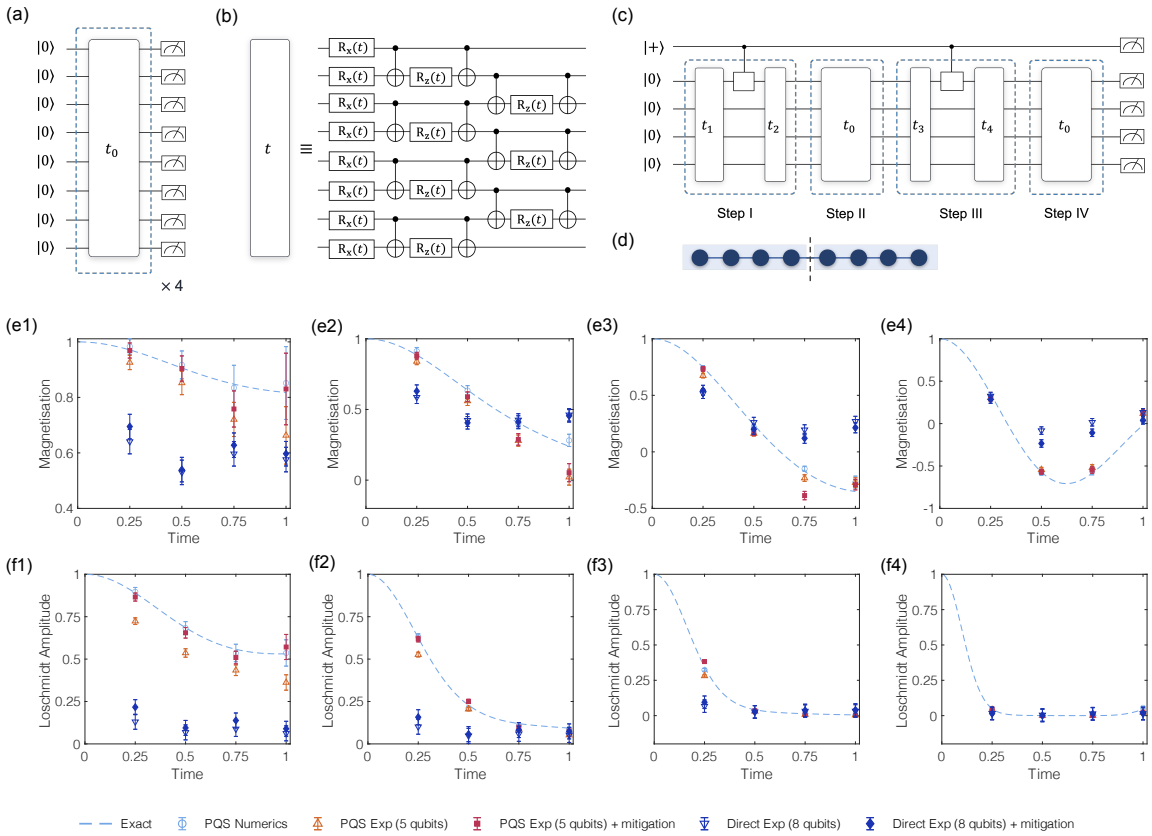


Figure 8.7: Implementing perturbative quantum simulation on the IBM quantum cloud. The DQPT of 8 interacting spins with nearest-neighbour interactions are considered. The initial state $|0\rangle^{\otimes 8}$ is evolved under the Hamiltonian $H = \sum_j \hat{\sigma}_j^z \hat{\sigma}_{j+1}^z + h \sum_j \hat{\sigma}_j^x$. (a) Quantum circuit implementation for 8-qubit simulation based on Trotterisation. (b) An example for the implementation of PQS to simulate 8-qubit system with operations on $4+1$-qubits. (c) The circuit block for single-step evolution. (d) The topological geometry for the spin system and the partitioning strategy. (e1-e4) The magnetisation along the z direction with $h = 0.5, 1, 1.5, 2.5$. (f1-f4) The Loschmidt amplitude with $h = 0.5, 1, 1.5, 2.5$. The comparison of results of exact simulation (dashed line), PQS (numerics, circle), PQS using IBMQ (5 qubits in (c), upper triangle) and the direct simulation using IBMQ (8 qubits in (a), lower triangle). The results using measurement error mitigation are shown, both for PQS (solid square) and direct simulation (solid diamond). 10^3 samples (8192 counts for each sample) are used.

with $t \in \{0.25, 0.5, 0.75, 1\}$ and $\delta t = 0.25$. Each term $e^{-i\delta t \hat{\sigma}_j^z \hat{\sigma}_{j+1}^z} = \mathrm{CNOT}_{j,j+1} \mathrm{R}_z(2\delta t, j+1) \mathrm{CNOT}_{j,j+1}$ could be realised with a single qubit rotation gate $\mathrm{R}_z(2\delta t, j+1) = e^{-i\delta t \hat{\sigma}_{j+1}^z}$ sandwiched by two controlled-X gates $\mathrm{CNOT}_{j,j+1}$ and each $e^{-ih\delta t \hat{\sigma}_j^x} = \mathrm{R}_z(2h\delta t, j)$ is a single qubit gate. As shown in Figure 8.7(c), for each step, all the single qubits gates are implemented in parallel and the two-qubit gates are realised with depth $d = 2$. The Trotter error could be negligible, which is much less than 10^{-2}.

With PQS, we only need to apply operations on $4+1$ qubits with the partitioning strategy shown in Figure 8.7(d). Each subsystem will evolve under the corresponding subsystem Hamil-

tonian, and at some time, a local operation, which is determined by the interaction between subsystems[3], is applied, which we refer to as a decay event. The maximal number of decay events is truncated to four, and the truncation error is found to be small. When a decay event happens at time t, say $t = 0.1$, we further divide the Trotter step from 0 to 0.25 into two steps, i.e., $[0, 0.1]$ and $[0.1, 0.25]$. Then we insert a controlled-Z operation, with the control qubit being the ancilla and the target being the first (last) qubit. As shown in Figure 8.7(b), we design the circuit in a similar way if we have multiple decay events. While the quantum circuit could be further optimised with fewer gates, it is nevertheless sufficient in this form to demonstrate the power of our PQS method.

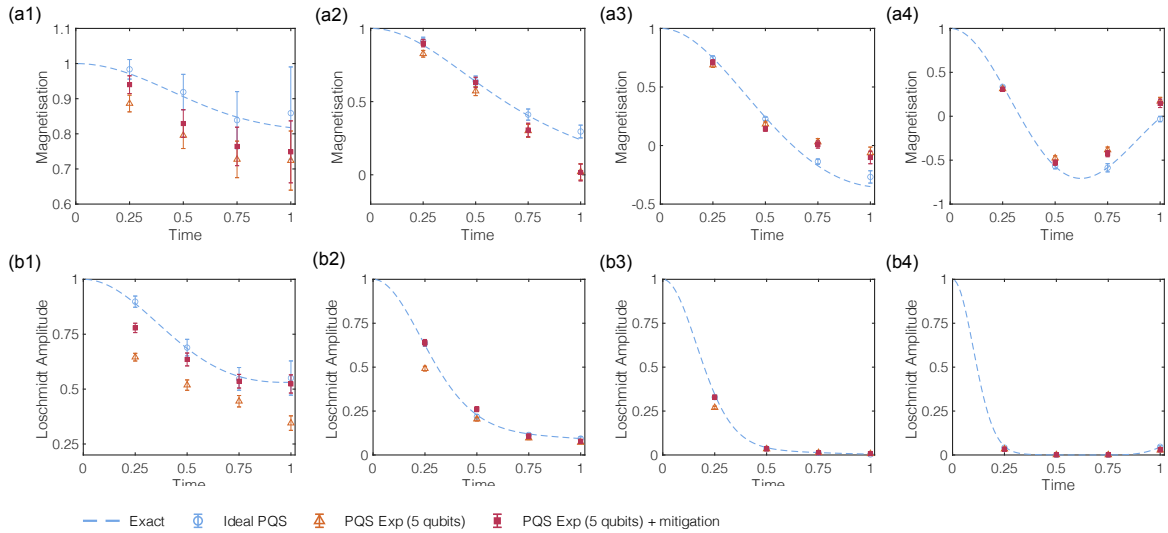


Figure 8.8: Implementing perturbative quantum simulation on the IBM quantum cloud with fewer samples and less optimised quantum circuit. We run 10^3 samples for PQS and collect 128 counts each samples. The quantum circuit for each time evolution block is different from Figure 8.7(c), where we apply the two-qubit gates sequentially with depth 7. (a1-a4) The magnetisation along the z direction with $h = 0.5, 1, 1.5, 2.5$. (b1-b4) The Loschmidt amplitude with $h = 0.5, 1, 1.5, 2.5$. I compare the results of exact simulation (dashed line), PQS (numerics, circle), PQS using IBMQ (5 qubits in (c), upper triangle). We also show the results using error mitigation for measurement both for PQS (solid square).

The direct 8-qubit simulation and our PQS method with 5 qubits were implemented on the IBM quantum cloud. The processor employed to conduct the direct 8-qubit simulation was 'ibmq_16_melbourne', which has 16 qubits with T_2 time ranging from $18 \sim 105\mu s$, CNOT gate error 3.3×10^{-2} and read-out error 4.7×10^{-2}. The processor employed to conduct the 5-qubit PQS method was 'ibmq_santiago', which has 5 qubits with T_2 time ranging from $66.9 \sim 143\mu s$, CNOT gate error 7.1×10^{-3} and read-out error 1.7×10^{-2}. The circuits were implemented through Qiskit [237], a python-based software development kit for working with OpenQASM and the

[3]Specifically, this local operation is Φ_k in Eq. (6.18).

IBMQ processors. The IBM cloud admits multiple job submissions, with each job consisting of a maximal of 72 circuits, where each circuit is fixed and allows 8192 single-shot measurements.

The experimental results are shown in Figure 8.7(e,f). The external field is applied along the x direction with $h = 0.5, 1, 1.5, 2.5$. The results of exact simulation (dashed line), PQS (numerics, circle), PQS using IBMQ (5 qubits in Figure 8.7(b), upper triangle) and the direct simulation using IBMQ (8 qubits in Figure 8.7(a), lower triangle) are compared. For each data point of the direct simulation, 16 circuits are run with 8192×16 single-shot measurements. For the PQS method, 1024 trajectories are considered, with each trajectory corresponding to a circuit measured 8192 times. Even though the number of samples of the PQS method is much larger than the number of samples for the direction simulation method, the shot noise is substantially smaller than the error caused by device imperfections. We could also use a smaller number of samples (128 samples) for each trajectory of the PQS method, and similar results are observed as shown in Figure 8.8. The simulation results are not identical, as we ran a less optimised circuit on the IBM quantum processor.

Measurement error mitigation is applied to increase the simulation accuracy. The measurement error mitigation is implemented by running a set of circuits with different computational-basis input states and computational basis measurements. We can obtain the readout noise matrix from the measurement results. The measurement read errors can be mitigated by applying the inverse of the readout noise matrix to the noisy measurement outcomes. I refer to [238] and [69] for detailed discussion on the theory and implementations, respectively. From our simulation results, we observe that the PQS method outperforms the direct simulation. This is because the five-qubit 'ibmq_santiago' processor is able to carry out more accurate operations than the 'ibmq_16_melbourne' processor. Since our PQS method only requires a small quantum computer with a relatively low circuit depth on which to run, it could be applied to benchmarking large-scale quantum devices, which may have more errors than a small-scale one.

8.2.2 Analysis of noise robustness

An explanation as to why PQS is more robust to noise in the simulation of general systems is pertinent at this point. Suppose we aim to simulate time evolution of an Ising Hamiltonian with Ln qubits. Conventional approaches require $2Ln - 2$ two-qubit gates for each Trotter step, whereas PQS only need $2n-2$ two-qubit gates if L number of n-qubit clusters is being considered. Suppose the fidelity of each two-qubit gate is $1-\varepsilon$, then the infidelities of the conventional method and PQS are $1 - (1 - \varepsilon)^{2Ln-2}$ and $1 - (1 - \varepsilon)^{2n-2}$, respectively. In the regime of small ε and $n\varepsilon$ and relatively large n, the state infidelity using PQS is approximately $n\varepsilon$, which is L times smaller than $Ln\varepsilon$ using conventional quantum simulation methods. For example, when $L = 2$, the infidelity will be half of that obtained from conventional quantum simulation. Therefore,

PQS not only allows for the simulation of larger systems, but also effectively increases simulation accuracy.

8.3 Spectroscopic features

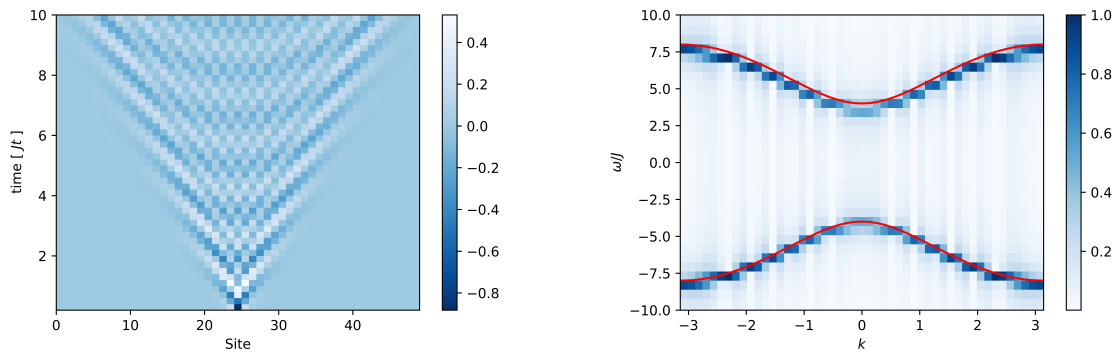


Figure 8.9: Simulation of the Ising model for $h/J = 3$ on a chain with $L = 49$ sites. Left. Evolution of the expectation value of the spin operator $\langle \sigma_i^y(t) \rangle$. Right. Normalised modulus of the quench spectral function. The energy band calculated in theory is ploted by a red line.

In this section, I demonstrate how interesting quantum phenomena can be probed by using the engineered spectroscopy technique developed in Chapter 7. Here, we consider a spectroscopy protocol with a local perturbation applied to the system, which generates a branch of low-lying excitations of the system. The quasiparticle spectroscopic features can consequently be probed by analysing the dynamics of the observables. The protocol is summarised as follows. First, the ground state of a target Hamiltonian H is prepared. A local perturbation is then applied to the initial ground state. Finally, the system is evolved under a unitary operator $U = e^{-iHt_i}$ with different time lengths t_i. The effect of perturbations in different physical systems will be discussed.

The time evolution e^{-iHt} is simulated using classical methods. For instance, tensor networks have been widely used in classically simulating the time evolution of quantum systems, as discussed in Chapter 3. The density matrix renormalisation group (DMRG) is used to initialise the system into the ground state (before applying quench). Then, the time-dependent variational principle (TDVP) algorithm, similarly to the variational principle introduced in Section 2.2.3.2, is applied to compute the time evolution after the local quench. Both of these algorithms are based on the matrix product states (MPS), which can efficiently represent 1D local gapped systems. It is also interesting to examine 2D systems, and we can still employ MPS to simulate two-dimensional arrays at the expense of more computational resources. In general, classical simulation becomes inefficient when the number of lattice sites L increases to a certain large size, or the interaction becomes complex. In these instances, quantum computing technique can be applied to surpass classical limitations and obtain the spectroscopic features of quantum systems.

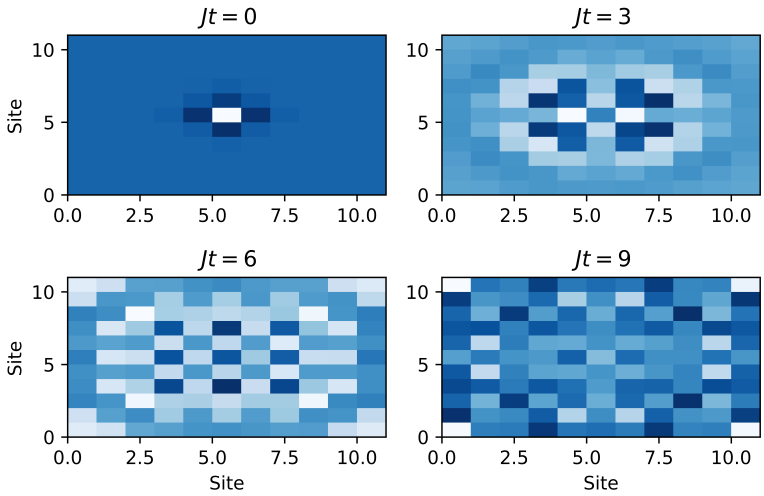


Figure 8.10: Simulation of the Ising model for $h/J = 3$ on a square lattice with $L_x = L_y = 11$. Evolution of the expectation value of the spin operator $\langle \sigma_i^y(t) \rangle$.

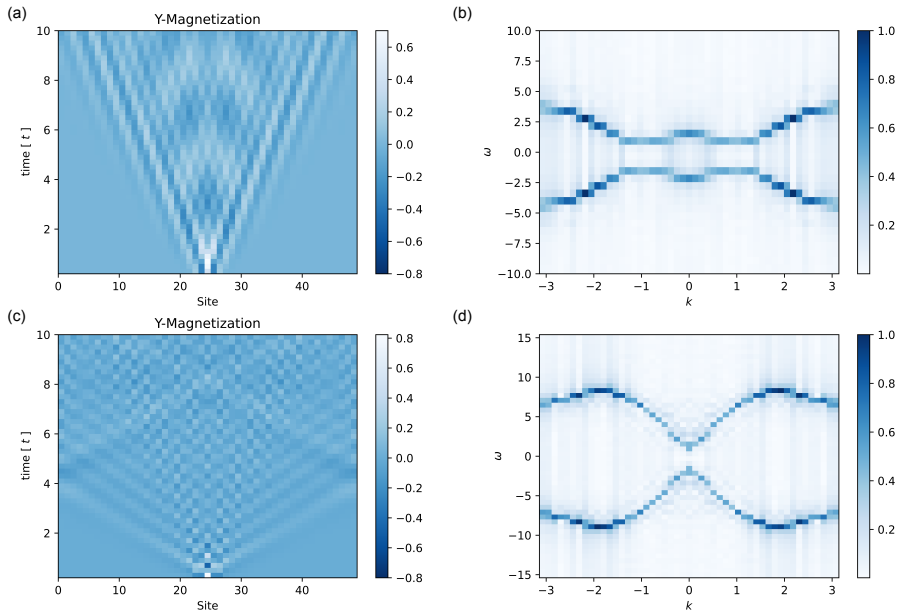


Figure 8.11: Simulation of the $J_1 - J_2$ model on a chain with $L = 49$ sites. Top: Hamiltonian parameters are $J_1 = J_2 = J$ and $h/J = 3$: (a) Evolution of the expectation value of the spin operator $\langle \sigma_i^y(t) \rangle$; (b) Modulus of the spectral function. Bottom: The same plots but for $J_2/J_1 = -0.5$ and $h/J_1 = 1.1$.

First, let us consider a transverse-field Ising (TFI) model with the Hamiltonian

$$H_{\mathrm{TFI}} = -J \sum_i \sigma_i^x \sigma_{i+1}^x - h \sum_i \sigma_i^z. \tag{8.26}$$

The quench spectroscopy[4] by a local perturbation at the centre of the spin chain is shown in

[4]Here, the term 'quench spectroscopy' is adopted from [20]. I refer to Chapter 7 for a more detailed discussion on different spectroscopic methods.

Figure 8.3. The quench spectral function (QSF) is the space-time Fourier transform of $G(t)$ in Eq. (7.9), namely the expectation value of a properly chosen observable under real-time evolution, as discussed in Chapter 7. The quasiparticle excitations agree well with the theoretical analysis which is plotted by the red line, as shown in Figure 8.3(b). For all 1D simulations, we set the bond dimension $\chi = 15$, and the lattice size $L = 49$. We then consider a 2D lattice of dimension $L_x \times L_y$ with $L_x = L_y = 11$ to have sufficient spatial resolution. The simulation results for a 2D lattice are shown in Figure 8.10. One needs to increase the bond dimension of MPS for the simulation of 2D systems, since the 2D state becomes non-local in the MPS representation.

Next, a $J_1 - J_2$ model is considered, which has been widely used in describing magnetism. The Hamiltonian is expressed as

$$H_{J_1 J_2} = -J_1 \sum_i \sigma_i^z \sigma_{i+1}^z - J_2 \sum_i \sigma_i^z \sigma_{i+2}^z - h \sum_i \sigma_i^x, \tag{8.27}$$

with the nearest-neighbour coupling strength J_1, the next-nearest-neighbour coupling strength J_2, and the external field h. Figure 8.11 shows the quasiparticle excitations with different interaction strengths and external fields. Further examples, such as the Bose-Hubbard model and Fermi-Hubbard model, can be found in a manuscript currently under preparation [96].

Author contributions. This chapter is relevant to a publication [95] and a manuscript under preparation [96]. J.S. conceived the applications demonstrated and the algorithm design in this chapter. J.S. carried out the numerical simulation for PQS. J.S. and X.Y. performed experiments on the IBM quantum cloud with the assistance of H.L.. L. E. carried out the simulation of spectroscopy with input from J.S..

Part IV

Towards realistic applications in the near future

In this part, I will discuss the quantum resource requirement and challenges of quantum computing towards realistic applications in the near future.

An interesting question is how far away it is for us to solve a practical problem using the advanced quantum computing methods. To answer this question, it is essential to address the resource requirement for such a problem. In Chapter 9, I will first discuss the requirement to demonstrate quantum advantage in the near future.

In Chapter 10, I will discuss the challenges in the NISQ era. For practical implementation, error is unavoidable; this includes the device error and statistical error. In order to address these challenges appearing in the practical implementation, I propose quantum error mitigation and better quantum state measurement schemes.

Chapter 9

Resource estimation

Estimating eigenstate properties of a quantum many-body system is a fundamental problem in quantum physics. In this chapter, I provide an efficient scheme for estimating eigenenergies and observable expectations on the eigenstates. The method for preparing the eigenstate of a Hamiltonian H is that we apply a filter operator to the initial state, so as to filter out the contributions from the other eigenstates. For instance, we can choose a Gaussian operator $g = e^{-(H-E_j)^2\tau^2}$ to prepare the eigenstate $|E_j\rangle$ with eigenenergy E_j[1]. The filter operator g in general is a nonunitary operator, and consequently cannot be directly implemented on a quantum computer. However, g can be expanded as a linear combination of real-time dynamics e^{-iHt_i} with different time lengths t_i as $g = \sum_i c_i e^{-iHt_i}$, and it can thus be implemented on a quantum computer by using the random sampling method introduced in Chapter 4.

The circuit complexity for eigenstate property estimation is determined by the maximum time $\max_i t_i$, which is proven to be logarithmic in the inverse precision $\mathcal{O}(\log(1/\epsilon))$, and it requires at most one ancillary qubit. A key question is whether we can realise e^{-iHt} in a circuit complexity $\mathcal{O}(\log(1/\epsilon))$ without introducing any additional ancillary qubits so that the total complexity will still be logarithmic in precision. To address this point, I introduce a new Hamiltonian simulation method by employing the Taylor series methods for the realisation of e^{-iHt}. I present the resource cost requirements for typical problems, such as lattice models and chemistry problems, at a quantum gate level (including CNOT gate numbers, T gate numbers, and qubit numbers). Our scheme is compared with existing advanced methods, including phase estimation, qubitisation, and LCU. This chapter provides a basis for the application of quantum computing in the NISQ and FTQC era

This chapter is relevant to a manuscript under preparation in collaboration with Pei Zeng and Vlatko Vedral [92].

[1] Here, the initial state is assumed to have a nonvanishing overlap with the eigenstate $|E_j\rangle$.

9.1 Introduction

Estimating properties of the ground state and excited states of a quantum many-body system is a long-standing problem of fundamental interest in condensed matter physics, quantum chemistry, and material science. It poses a challenging for both classical computing and quantum computing; estimating the ground state energy of a k-local ($k \geq 2$) Hamiltonian is indeed a QMA-complete problem [2]. Nevertheless, quantum computing holds the promise of being able to tackle this challenge on the condition that certain assumptions are met[2]. There has been considerable progress in the development of quantum computing algorithms, such as quantum phase estimation [29], quantum signal processing [31–34], projection by linear combination of unitaries (LCU) [29]. While these algorithms are believed to have the potential to provide a rigorous solution in the long-term, their implementation requires an oracle or a deep circuit, which imposes challenges to current quantum devices. In the noisy intermediate-scale quantum computing (NISQ) era and the early fault-tolerant quantum computing (FTQC) era, it is much more desirable to design resource efficient algorithms. To achieve this goal, the query complexity, as well as the gate complexity after circuit compilation, must be estimated. This in turn involves the systematic analysis of the algorithm complexity with the consideration of logarithmic factors or even constant prefactors.

A few papers have discussed the cost of performing quantum algorithms at gate level for practical problems, such as quantum chemistry [99,101,239,240], condensed-phase electrons [187, 188], etc. In these works, energy estimation or eigenstate preparation are mainly based on phase estimation. A typical strategy is to encode the eigenspectra of the Hamiltonian in a unitary for phase estimation by the evolution e^{-iHt}, which is synthesised by Trotterisation [187, 188], or a qubitised quantum walk [101, 239] with eigenspectrum proportional to $e^{\pm i \arccos(H/\lambda)}$ where λ is a parameter related to the norm of the Hamiltonian. However, due to the cost of phase estimation, the circuit depth will inevitably scale polynomially in precision, which is far from optimal computation.

To circumvent this problem, several works consider using a projection method, for instance, e^{-Ht}, which defines a natural way to prepare the eigenstate. With the development of LCU and qubitisation, these nonunitary projectors can be directly synthesised [33, 37]. For instance, Ge *et al.* proposed filtering the initial state by realising a LCU of $\cos^M(Ht)$, which achieves the ground state preparation to a high degree of precision [29]. Lin and Tong proposed using the block-encoding to realise the projector, and this method achieves a near-optimal asymptotic scaling for eigenstate preparation [31]. However, the block encoding method requires multi-qubit control gates, and thus requires quantum devices with a long coherence time. Subsequently, the quantum eigenvalue transformation of unitaries (QET-U) method was proposed for the task [241],

[2]The common assumptions include a nonvanishing energy gap, and a nonvanishing overlap of the initial state and the target state, outlined in 1 and 2, respectively.

an approach which uses a similar idea of qubitisation, but avoids querying the block-encoding of the Hamiltonian. However, it relies on the controlled unitary oracle, e^{-iHt}, and the use of Trotter formulae eliminates the advantage of the scaling in the observable estimation, or eigenenergy estimation. Indeed, conventional ancilla-free methods such as Trotterisation only achieve a scaling polynomial in the error, and it is still unknown if one can achieve the Hamiltonian simulation with scaling of logarithmic in the error without block-encoding of the Hamiltonian.

To address the aforementioned problems, I provide an efficient end-to-end scheme for the two tasks, with the best asymptotic scaling for the circuit depth with respect to the precision. The method developed in this chapter is much easier to implement, since it does not require multi-control gates. The two problems considered here are eigenenergy estimation and observable estimation on the eigenstate. The latter can be further divided into two cases; one with a priori known eigenenergy and the other with unknown eigenenergy. I proposed using a recently developed method, which realises a nonunitary operator by randomly sampling real-time evolutions with different time length [93] (see Chapter 4). The circuit complexity for each instance is $\mathcal{O}(\log(\varepsilon^{-1}))$. Importantly, I demonstrate that the advantage of logarithmic dependence on the error is preserved by employing a random paired Taylor-series sampling (RTS) method to realise the Hamiltonian simulation. It is also worth remarking that by comparison with Lin and Tong's method for ground state property estimation [48], our circuit complexity, with respect to precision, has better scaling in the polylogarithmic factor, due to the reuse of the quantum measurement results.

Our method is a universal algorithm that can be applied to the study of a broad class of problems of physical interest, such as condensed matter and quantum chemistry. In practical applications, apart from scaling with respect to precision and system-dependent parameters (such as the energy gap and the initial state overlap), another highly relevant question is scaling in respect to the number of terms L in the Hamiltonian. For quantum chemistry problems, the second-quantised Hamiltonian, when represented in the Gaussian basis set, has a number of terms increasing quartically, $\mathcal{O}(N^4)$, with the system size, although it could be reduced by considering other basis sets, symmetry, or factorisation [43, 99, 101, 130]. For the quantum signal processing or LCU method, although the query complexity of quantum signal processing or RTS method only depends on the strength of the Hamiltonian λ, the gate complexity depends on the number of terms in the Hamiltonian when the block-encoding operation is compiled into elementary gates. A major contribution of our work is that we achieve near-optimal system-size dependence for lattice Hamiltonians. This is achieved by exploiting the commutation relation of lattice Hamiltonians. For systems with nearest-neighbour interaction, our algorithm has a scaling of $\mathcal{O}(n^{1+o(1)}\varepsilon^{-(1+o(1))})$[3], which is nearly-optimal with respect to the system size n. This shows a clear advantage over QSP, which has a scaling $\mathcal{O}(n^2 \log(1/\varepsilon))$ [31]. The gate complexity

[3]There is a small overhead in the power of n, which depends on the order of the paired Taylor series.

of QSP depends on the number of terms in the Hamiltonian when the block-encoding operation is compiled into elementary gates. This complexity cannot be decreased by considering the commutation relation of the Hamiltonian, hence non-opitmal for lattice Hamiltonians.

In addition, to improve the dependence on L, we can choose the zeroth-order RTS method for Hamiltonian simulation, and the total complexity for eigenstate preparation is $\mathcal{O}((\lambda t)^2)$ where t is the time length in real time dynamics. However, in a practical scenario, there is a trade-off between the L and t dependence. I will show the comparison in different regimes.

With the establishment of the framework, I analyse the resource in detail concerning its actual overhead and circuit compilation. To meet the requirements of application in the NISQ era and FTQC era, I analyse the gate count for solving practical problems. For near-term application, the major bottleneck is the number of two-qubit gates for a noisy quantum computer, while the major overhead for the error-corrected quantum computer is the T gate count, which requires more gates to perform error correction. We study the resource requirement for ground-state energy estimation and observable estimation for representative physical models in condensed matter and chemistry. I set out a systematic comparison with existing advanced methods for eigenstate preparation, which include the phase estimation method combining the Hamiltonian simulation with Trotter formulae (first-order, higher-order, qDRIFT, and randomised version) or qubitisation method, projection method based on LCU, the QSP method, and the QET-U method. One remarkable contribution of this chapter is that I demonstrate how to analyse the resource cost of different methods with respect to the number of CNOT gates, T gates and qubits, which is achieved by providing a detailed circuit compilation strategy. This chapter provides a useful toolbox for researchers to compare different methods with various initial conditions. More importantly, this chapter answers quantum resources requirements for eigenstate property estimation in realistic applications with both noisy quantum computers and early fault-tolerant quantum computers.

9.2 Methods

In this section, I discuss the framework of eigenenergy estimation and eigenstate property estimation, as introduced in Section 1.2.1.

To access the physical properties of eigenstates, the first step is to prepare the eigenstate by applying a projection operation to the initial state that projects out the contributions from the other eigenstates. While the projector is nonunitary by construction, we can realise it effectively by using a combination of real time dynamics e^{-iHt_i} with different real time t_i. The maximum time complexity, $\max_i t_i$, which determines the circuit complexity required to achieve an additive error ε, is proven to be logarithmic in the inverse error $\mathcal{O}(\log(\varepsilon^{-1}))$. The next step is the realisation of time evolution, while preserving the logarithmic dependence on the error. As discussed in Section 2.4, Trotter methods will make the error dependence polynomial in the

inverse error, and advanced Hamiltonian simulation methods such as qubitisation or LCU will introduce additional cost on the block encoding and ancillary qubits, which is undesirable in this scenario. To circumvent these problems, I develop random paired Taylor-series sampling methods for high-precision Hamiltonian simulation, which is discussed in Section 9.2.1.

9.2.1 Hamiltonian simulation with random Taylor-series sampling

In this section, I consider Hamiltonian simulation by decomposing real-time evolution to a combination of simple realisable unitaries. Let us consider an n-qubit Hamiltonian in the form,

$$H = \sum_{l=1}^{L} H_l = \sum_{l=1}^{L} \alpha_l P_l \tag{9.1}$$

where the coefficients $\{\alpha_l\}_l$ are all positive as the signs have been absorbed into the corresponding Pauli matrices. The following characters of the Hamiltonian is defined,

$$\begin{aligned} n_L &:= \lceil \log_2 L \rceil \\ \lambda &:= \sum_{l=1}^{L} \alpha_l \\ \Lambda &:= \max_l \alpha_l \\ \mathrm{wt}(H) &:= \sum_{l=1}^{L} \mathrm{wt}\left(P_l\right) \\ \mathrm{wt}_m(H) &:= \max_l \mathrm{wt}\left(P_l\right), \end{aligned} \tag{9.2}$$

which are used in this chapter. Here, $\mathrm{wt}(P_l)$ indicates the weight (or called support) of the Pauli operator P_l, i.e., the number of $\{X, Y, Z\}$ terms in P_l.

Next, I introduce a construction of (μ, ε)-random-sampling linear-combination-of-unitary (RLCU) formula in Definition 3.

Definition 3. *A (μ, ε)-random-sampling linear-combination-of-unitary (RLCU) formula of a unitary U is defined to be*

$$\tilde{U} = \mu \sum_i \Pr(i) U_i, \tag{9.3}$$

such that the spectral norm distance $\|U - \tilde{U}\| \leq \varepsilon$. Here, $\varepsilon \leq 1$, $\mu > 0$ is a real number indicating the normalisation cost, $\Pr(i)$ is a probability distribution of i, and $\{U_i\}_i$ is a group of unitaries.

As our method uses random sampling, the state cannot be prepared deterministically. However, in terms of the expectation of observables, our scheme is similar to other schemes that can prepare eigenstates deterministically. In order to compare the performance of our methods and the other deterministic schemes, I introduce Proposition 7, which indicates that once we have a (μ, ε)-RLCU formula, we can estimate the expectation $\mathrm{Tr}(U\rho U^\dagger O)$ with the following performance guarantee. As such, the performance of our method in observable estimation can

be estimated by the error in the construction of the RLCU formula. One can thus compare our method with the other deterministic schemes at the same level.

Proposition 7 (Observable estimation using the RLCU formula)**.** *For a target unitary U and its (μ, ε)-random-sampling formula defined in Definition 3, if we estimate the value $\langle O \rangle_U := \mathrm{Tr}(U\rho U^\dagger O)$ with an initial state ρ and observable O using the variant of Hadamard test circuit for $N = 2\mu^4 \ln(2/\delta)/\varepsilon_n^2$ times, then the distance between the mean estimator value $\hat{O}$ and the true value $\langle O \rangle_U$ is bounded by*

$$|\hat{O} - \langle O \rangle_U| \leq \|O\|(3\varepsilon + \varepsilon_n), \tag{9.4}$$

with success probability $1 - \delta$ and $\|O\|$ being the spectral norm of O.

From Proposition 7 we can see that the normalisation factor μ affects the sample complexity, and the accuracy factor ε introduces an extra bias in the observable estimation. In what follows, I will discuss different RLCU formulae and compare their performance by the normalisation factor μ and accuracy ε.

The most natural and direct Hamiltonian simulation method is to apply the Lie-Trotter-Suzuki (Trotter) formulae, which approximate the real-time evolution operator $U(t)$ by the product of the evolution of the summands e^{itH_l} which are relatively easy to implement. The Trotter formulae have been discussed in Section 2.4.

The major drawback of Trotter methods is their polynomial gate-number dependence on the inverse accuracy $1/\varepsilon$, $\mathcal{O}(\mathrm{Poly}(1/\varepsilon))$. The polynomial gate complexity with respect to ε is due to a large remainder of the Trotter formula. This inspires us that we may reduce the complexity by compensating the Trotter error, as discussed in Section 2.4. Below, I show how to compensate the Trotter error with a linear-combination-of-unitary (LCU) method.

To motivate the discussion, I first expand $U(x)$ by the Taylor series,

$$\begin{aligned} U(x) &= \sum_{s=0}^{\infty} \frac{(-ixH)^s}{s!} \\ &= \sum_{s=0}^{\infty} \frac{x^s}{s!} \sum_{l_1,\ldots,l_s} \alpha_{l_1}\alpha_{l_2}\ldots\alpha_{l_s}(-i)^s P_{l_1}P_{l_2}\ldots P_{l_s}, \\ &= \sum_{j=0}^{\Gamma-1} \beta_j V_j, \end{aligned} \tag{9.5}$$

which is an LCU formula with a finite truncation at the third equation. To realise $U(x)$ coherently on a quantum circuit, one can first prepare ancillary systems with the dimension Γ and amplitude $\{\beta_j\}$, and then perform oblivious amplitude amplification to make $U(x)$ nearly deterministic [38]. However, this requires many ancillary qubits and a large number of multi-controlled Toffoli gates,

which is challenging for near-term devices. Alternatively, we can rewrite Eq. (9.5) as follows,

$$\begin{aligned} U(x) &= \sum_{s=0}^{\infty} \frac{(\lambda x)^s}{s!} \sum_{l_1,\ldots,l_s} p_{l_1} p_{l_2} \ldots p_{l_s} (-i)^s P_{l_1} P_{l_2} \ldots P_{l_s} \\ &= \mu_0(x) \sum_{s=0}^{\infty} \mathrm{Poi}(s;\lambda x) \sum_{l_1,\ldots,l_s} p_{l_1} p_{l_2} \ldots p_{l_s} (-i)^s P_{l_1} P_{l_2} \ldots P_{l_s} \\ &= \mu_0(x) \sum_{s=0}^{\infty} \sum_{l_1,\ldots,l_s} \mathrm{Pr}(s; l_{1:s}) U_0^{(s)}(l_{1:s}), \end{aligned} \tag{9.6}$$

where $\mu_0(x) = e^{\lambda x}$ is the normalisation coefficient, and

$$\mathrm{Poi}(s;x) := e^{-x} \frac{x^s}{s!}, \tag{9.7}$$

is the Poisson distribution with average x, $l_{1:s}$ stands for $\{l_1, l_2, \ldots, l_s\}$, $\mathrm{Pr}(s; l_{1:s}) := \mathrm{Poi}(s;\lambda x) p_{l_1} p_{l_2} \ldots p_{l_s}$, and $U_0^{(s)}(l_{1:s}) := (-i)^s P_{l_1} P_{l_2} \ldots P_{l_s}$. Eq. (9.6) is then a $\mu_0(x)$-RLCU formula satisfying Definition 3.

If we consider the whole time-evolution $U(t)$ with ν time segments, from Eq. (9.6) we can also obtain a random-sampling formula

$$U(t) = U(x)^\nu = \mu_0(x)^\nu \left(\sum_{s=0}^{\infty} \sum_{l_1,\ldots,l_s} \mathrm{Pr}(s; l_{1:s}) U_0^{(s)}(l_{1:s}) \right)^\nu, \tag{9.8}$$

with the normalisation factor $\mu_{0,tot} = \mu_0(x)^\mu = e^{\lambda t}$. Recall that the normalisation factor μ will determine the sample complexity to learn observable information of the target state σ. When $\mu_{0,tot} = e^{\lambda t}$, this indicates that the sample cost scales exponentially with respect to the simulation time t. This is reasonable, otherwise we could efficiently simulate the quantum dynamics on a classical computer, as the circuit using the LCU formula in Eq. (9.6) is composed of Clifford gates, which can be simulated classically.

To understand the reason why the direct Taylor-series expansion leads to a divergent normalisation formula, one can rewrite Eq. (9.6),

$$\begin{aligned} U(x) &= \sum_{s=0}^{\infty} \frac{(-i\lambda x)^s}{s!} \sum_{l_1,\ldots,l_s} p_{l_1} p_{l_2} \ldots p_{l_s} P_{l_1} P_{l_2} \ldots P_{l_s} \\ &= \sum_{s=0}^{\infty} F_0^{(s)}(x), \end{aligned} \tag{9.9}$$

where $F_0^{(s)}(x) = \frac{(-i\lambda x)^s}{s!} \sum_{l_1,\ldots,l_s} p_{l_1} p_{l_2} \ldots p_{l_s} P_{l_1} P_{l_2} \ldots P_{l_s}$ collects all the s-order Pauli terms related to x^s. Since the Pauli terms with different order of x are sampled separately, we can individually estimate the contribution of the terms with different order of x in the overall normalisation factor of $U(x)$,

$$\begin{aligned} \mu_0(x) = \mu(U(x)) &= \sum_{s=0}^{\infty} \mu(F_0^{(s)}(x)) \\ &= \sum_{s=0}^{\infty} \frac{(\lambda x)^s}{s!} = e^{\lambda x}. \end{aligned} \tag{9.10}$$

Here, $\mu(\cdot)$ denotes the normalisation factor of a specific LCU formula satisfying Definition 3.

Eq. (9.9) indicates that the normalisation contributions from Pauli terms with different x orders are distinct. When $\lambda x \ll 1$, the major contribution to the normalisation factor $\mu_0(x)$ comes from terms with small s. Based on this observation, a prescription to suppress the normalisation formula is to eliminate the low-order terms with respect to x in the LCU formula. The overall normalisation factor for ν time segments is then upper bounded by

$$\mu_{k,tot} = \mu_k(x)^\nu \leq (e^{ex^{k+1}})^\nu = e^{e\frac{t^{k+1}}{\nu^k}}. \tag{9.11}$$

In this instance, if the segment number is set as $\nu = \mathcal{O}(t^{\frac{k+1}{k}})$, the normalisation factor will then become a constant.

In what follows, I mainly discuss two ways to eliminate the low-order x-expansion terms, inspired by Ref. [242] and [243], which are illustrated in Figure 9.1.

1. Pairing the terms with different expansion orders. The major observation comes from the following fact: for any Pauli operator P and $x \in \mathbb{R}$,

$$I + ixP = \sqrt{1+x^2}\, e^{i\theta P}, \tag{9.12}$$

   where $\theta = \tan^{-1}(x)$. Then, the normalisation factor is suppressed from $(1+x)$ to $\sqrt{1+x^2} \leq 1 + \frac{1}{2}x^2$.

2. Insert Trotter formulae before the Taylor-series expansion. Instead of expanding $U(x)$ directly, we can implement Trotter formulae first and then expand the remainder of the Trotter formula $V_k(x)$ defined in Eq. (2.50). If we expand $V_k(x)$ by the order of x,

$$V_k(x) = \sum_{s=0}^{\infty} F_k^{(s)}(x), \tag{9.13}$$

   where $F_k^{(s)}(x)$ collects all the Pauli terms related to x^s, then based on the Trotter formula results in Eq. (2.51), we have

$$F_k^{(s)}(x) = 0, \quad s = 1, 2, ..., k. \tag{9.14}$$

   As a result, the low x-order terms are automatically cancelled.

Several different random Taylor-series expansion formulae for real-time evolution $U(t) = e^{-itH}$ and the corresponding sampling algorithms could be considered. The key idea is to first split the controlled-$U(t)$ evolution to ν segments, each of which is a controlled-$U(x)$ evolution with $x = t/\nu$. Each segment is composed of Kth-order Trotter circuits $U_K^L(x)$ and $U_K^R(x)$ and controlled random Taylor-series sampling circuit $\tilde{V}_K^{(P)}(x)$. Here, $K = 0$, 1, or $2k$ ($k \in \mathbb{N}_+$). The superscript P indicates the specific Taylor-series sampling method to be used, which could be D

Table 9.1: Properties of different RLCU formulae for $U(x) = e^{-ixH}$ based on Taylor series expansion, including direct and paired Taylor-series sampling (TS). $c_k < 0.3$ is a small positive number defined as $c_k := \frac{1}{2}\left(\frac{e}{2k+1}\right)^{2k+1}$.

Trotter order	Direct TSS μ	Paired TSS μ	Truncated formula accuracy ε
0th order	$\mu \leq e^{\lambda x}$	$\mu \leq e^{\frac{3}{2}(\lambda x)^2}$	$\varepsilon \leq 2\left(\frac{e\lambda x}{s_c+1}\right)^{s_c+1}$
1st Trotter	$\mu \leq e^{(2\lambda x)^2}$	N/A	$\varepsilon \leq 2\left(\frac{2e\lambda x}{s_c+1}\right)^{s_c+1}$
$2k$th Trotter	$\mu \leq e^{e(2\lambda x)^{2k+1}}$	$\mu \leq e^{(e+c_k)(2\lambda x)^{2(2k+1)}}$	$\varepsilon \leq 2\left(\frac{2e\lambda x}{s_c+1}\right)^{s_c+1}$

(direct sampling) or L (leading-order pairing). The characteristics of these RLCU formulae are summarised in Table 9.1.

The gate complexity of the overall algorithm can then be estimated. To construct controlled-$U(t)$, we split it to ν segments. In each segment, we need to implement Kth-order Trotter circuits and random Taylor-series sampling circuits. The number of gates in the random Taylor-series sampling circuit is $\mathcal{O}(s_c)$. Therefore, the gate complexity of the overall algorithm using Kth Trotter formula ($K = 0, 1, 2k$) is given by

$$\begin{aligned} N_K &= \mathcal{O}(\nu(\kappa_K L + s_c)) \\ &= \mathcal{O}\left(\nu(\kappa_K L + \frac{\log(1/\varepsilon)}{\log\log(1/\varepsilon)})\right), \end{aligned} \tag{9.15}$$

where

$$\kappa_K = \begin{cases} K, & K = 0,\ 1 \\ 2 \cdot 5^{K/2-1}, & K = 2k,\ k \in \mathbb{N}_+. \end{cases} \tag{9.16}$$

Detailed gate complexity of zeroth-order and $2k$th-order leading-order-pairing algorithm are summarised in Proposition 8 and Proposition 9, respectively.

9.2.2 Gate complexity

In this section, I discuss the gate complexity of ground state property estimation using different algorithms. The time complexity to prepare the ground state up to precision ε using a Gaussian cooling function (discussed in Chapter 4) is given by

$$t = \frac{1}{\Delta}\sqrt{\ln\left(\frac{18}{\eta\varepsilon}\right)}\sqrt{\ln\left(\frac{9}{\pi^2\eta\varepsilon}\right)}. \tag{9.17}$$

We can find that the time complexity scales linearly to the energy gap Δ, and logarithmically to the error ε and the lower bound of the initial state overlap η.

The following question is how to realise the Hamiltonian simulation in which the scaling advantage is preserved, and it is ancilla free. If we use the pth order Trotter method to realise e^{-iHt}, the error is

$$\left\| U_{\mathrm{HS}}^{(p)} - e^{-itH} \right\| \leq C_p t^{p+1} \nu^{-p} \tag{9.18}$$

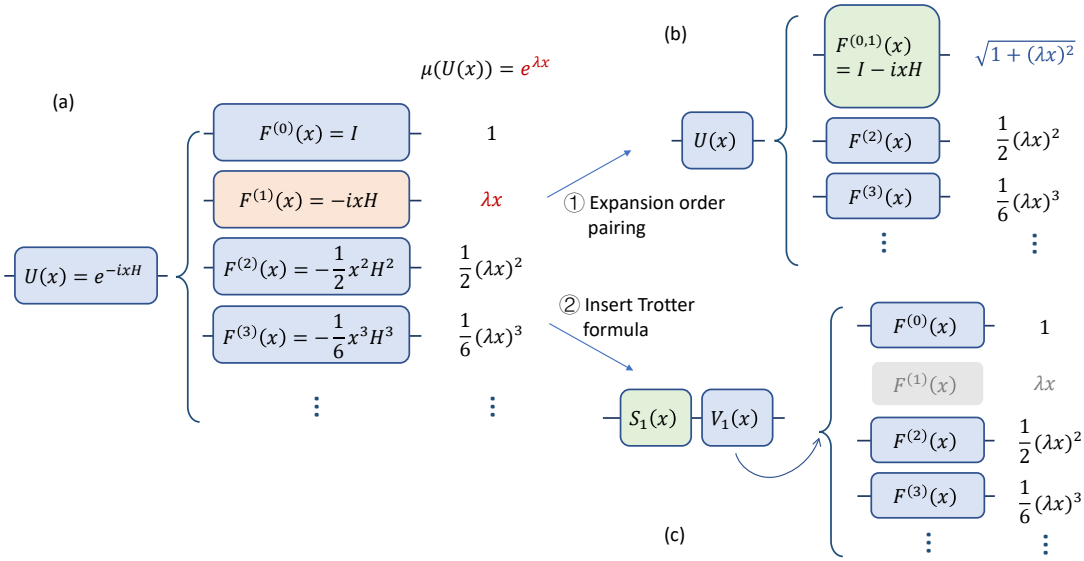


Figure 9.1: Two ways to suppress the normalisation factor of Taylor-series expansion formula, discussed in Ref. [242] and [243]. (a) In the Taylor-series expansion formula, the normalisation factor $\mu = e^{\lambda x}$ diverges. For a small time x, the major contribution to μ comes from the low-order term, especially the first-order term. (b) By pairing the expansion order, I introduce Pauli rotation gates while suppressing the normalisation factor quadratically from $\mathcal{O}(x)$ to $\mathcal{O}(x^2)$. (c) By inserting the Trotter formula before we implement the Taylor-series expansion, we can eliminate the corresponding low-order terms in the expansion.

The requirement for the segment number ν is $\mathcal{O}(C_p^{\frac{1}{p}}(t)^{1+\frac{1}{p}}\varepsilon^{-\frac{1}{p}})$ up to precision ε with an overhead, which can be found in Lemma 1. Therefore, Trotter formulae are undesirable for the high-precision simulation. To address this problem, I employ the random Taylor series sampling method introduced in Section 9.2.1 in this context. The gate complexity of the zeroth-order leading-order-pairing algorithm is summarised in the following proposition.

Proposition 8 (Gate complexity of the zeroth-order leading-order-pairing algorithm)**.** *In the zeroth-order leading-order pairing algorithm in Algorithm 6, if we set the time segment number to be*

$$\nu \geq \frac{3}{2}\frac{(\lambda t)^2}{\ln \mu}, \tag{9.19}$$

and the truncation order to be

$$s_c \geq \max\left\{\left\lceil \frac{\ln\left(\frac{3\mu(\lambda t)^2}{\ln(\mu)\varepsilon}\right)}{W_0\left(\frac{3\lambda t}{2e\ln(\mu)}\ln\left(\frac{3\mu(\lambda t)^2}{\ln(\mu)\varepsilon}\right)\right)} - 1\right\rceil, 1\right\}, \tag{9.20}$$

we can realise a (μ,ε)-RLCU formula where $\mu > 1, 1 > \varepsilon > 0$ are two constants.

We can find that the zeroth-order leading-order-pairing formula has no dependence on L, while the gate complexity scales quadratically to t, which is undesirable for long-time simulation. The RLCU formula for the zeroth-order leading-order-pairing formula is presented in Proposition 21

in Appendix F. To improve the dependence on t, we can use the $2k$th-order leading-order-pairing algorithm which inserts the $2k$th-order Trotter formula to compensate the error. The gate complexity of the $2k$th-order leading-order-pairing algorithm is summarised in the following proposition.

Proposition 9 (Gate complexity of the $2k$th-order leading-order-pairing algorithm)**.** *In the second-order leading-order-pairing algorithm or general $2k$th-order leading-order-pairing algorithm, if the time segment number is set to be*

$$\nu \geq \max\left\{\nu_{2k}^{(L)}(t), 2\lambda t\right\}, \tag{9.21}$$

and the truncation order to be

$$s_c \geq \max\left\{\left\lceil \frac{\ln\left(\frac{2\mu}{\varepsilon}\nu_{2k}^{(L)}(t)\right)}{W_0\left(\frac{1}{2e\lambda t}\nu_{2k}^{(L)}(t)\ln\left(\frac{2\mu}{\varepsilon}\nu_{2k}^{(L)}(t)\right)\right)} - 1\right\rceil, 4k+1\right\}, \tag{9.22}$$

we can realise a (μ,ε)-RLCU formula where $\mu > 1, 1 > \varepsilon > 0$ are two constants. Here, $\nu_{2k}^{(L)}(t) := \left(\frac{2(e+c_k)\lambda t}{\ln\mu}\right)^{\frac{1}{4k+1}} 2\lambda t$, $c_k := \frac{1}{2}\left(\frac{e}{2k+1}\right)^{2k+1}$.

We can find that the time dependence is improved from $\mathcal{O}(\lambda t)^2$ (zeroth-order leading-order-pairing) to $\mathcal{O}(\lambda t)^{1+\frac{1}{4k+1}}$ ($2k$th-order leading-order-pairing). However, compared to the zeroth-order pairing, it requires inserting the Trotter formula in each segment, and thus the gate complexity depends on L. The formal proofs of Proposition 8 and Proposition 9 were mainly carried out by my collaborator Pei Zeng, which involves many technical details. I direct the interested reader to [92] for more detailed discussion and the proof.

The gate complexities of the algorithms with random direct and paired Taylor-series sampling implementation are summarised in Table 9.2. More detailed discussions can be found in Section F.4.

Table 9.2: Summary of the random-sampling Trotter-LCU Hamiltonian simulation protocols and corresponding gate complexities. Here, $\tilde{\log}(x) := \log(x)/\log\log(x)$.

Trotter order	Direct Taylor-series expansion	Paired Taylor-series expansion
0th-order	N/A	$\mathcal{O}((\lambda t)^2\tilde{\log}(1/\varepsilon))$
1st-order	$\mathcal{O}((\lambda t)^2(L+\tilde{\log}(1/\varepsilon)))$	N/A
$2k$th-order	$\mathcal{O}((\lambda t)^{1+\frac{1}{2k}}(L+\tilde{\log}(1/\varepsilon)))$	$\mathcal{O}((\lambda t)^{1+\frac{1}{4k+1}}(L+\tilde{\log}(1/\varepsilon)))$

9.2.3 Main results

Now, I summarise the main theoretical results in this chapter. Our results rely on Assumption 1 and 2.

Theorem 6 (Eigenstate energy estimation)**.** *To estimate the ground-state energy E_0 with a precision κ and a success probability $1-\delta$ using the RLCU scheme, the maximal time in the Hamiltonian simulation e^{-iHt} is $\mathcal{O}(\kappa^{-1}\log(\eta^{-1}))$. The repeated query (sample) complexity is $\mathcal{O}(\eta^{-2}\log(\delta^{-1}))$. The total circuit complexity is $\mathcal{O}(\eta^{-2}\kappa^{-1}\log(\eta^{-1})\log(\delta^{-1}))$. Suppose we divide the time slice into $2k$ segments and use $2k$th-order leading-order-pairing algorithm, the gate complexity in each sample is $\mathcal{O}(L(\lambda\kappa^{-1})^{1+\frac{1}{4k+1}}\log(\eta^{-1}))$. The total gate complexity is $\mathcal{O}(L\eta^{-2}(\lambda\kappa^{-1})^{1+\frac{1}{4k+1}}\log(\eta^{-1})\log(\delta^{-1}))$.*

Theorem 7 (Observable estimation on the eigenstate with known eigenenergy)**.** *To estimate the observable expectation on the eigenstate $\langle u_0|\hat{O}|u_0\rangle$ with a precision ε and a success probability $1-\delta$ using RLCU scheme, the maximal time in the Hamiltonian simulation e^{-iHt} is $\mathcal{O}(\Delta^{-1}\log(\eta^{-1}\varepsilon^{-1}))$. The repeated query (sample) complexity is $\mathcal{O}(\varepsilon^{-2}\eta^{-2}\log(\delta^{-1}))$. The total circuit complexity is $\mathcal{O}(\varepsilon^{-2}\eta^{-2}\Delta^{-1}\log(\eta^{-1}\varepsilon^{-1})\log(\delta^{-1}))$. Suppose we divide the time slice into $2k$ segments and use $2k$th-order leading-order-pairing algorithm, the gate complexity in each sample is $\mathcal{O}(L(\lambda\Delta^{-1})^{1+\frac{1}{4k+1}}\log(\eta^{-1}\varepsilon^{-1}))$. The total gate complexity is*

$$\mathcal{O}(L\eta^{-2}\varepsilon^{-2}(\lambda\Delta^{-1})^{1+\frac{1}{4k+1}}\log(\eta^{-1}\varepsilon^{-1})\log(\delta^{-1})).$$

Compared to other universal quantum algorithms, the total gate complexity has a worse scaling with respect to the initial state overlap, which is due to the random sampling implementation in our scheme. A comparison between the time complexity of our scheme and a variety of typical methods is depicted in Table B.1 and Table B.2.

9.3 Circuit compilation

9.3.1 Stage Setting

In this section, I briefly introduce and estimate the gate cost of the three typical Hamiltonian simulation algorithms, including Trotter formulae, qubitised quantum walk, and quantum signal processing (QSP). After that, I summarise the gate cost of our method which combines RLCU and RTS.

To estimate the gate cost of each algorithm, we synthesise the circuits to CNOT gates, single-qubit Clifford gates and T gates. The CNOT gate number is more important for a near-term application with no or limited fault tolerance; while the T-gate number is more critical for a long-term application on a fully fault-tolerant quantum computer.

In some subroutines of the above algorithms, a direct estimation of the T-gate number is hard to obtain. In these events, we first synthesise the circuits to CNOT gates, single-qubit Clifford gates and single-qubit Z-axis rotation gates $R_z(\theta)$. Then the T-gate number n_T is estimated using the $R_z(\theta)$ gate number n_{Rz}. We could consider the optimal ancilla-free gate synthesis algorithm in Ref. [244], which requires $3\log_2(1/\varepsilon)+\mathcal{O}(\log\log(1/\varepsilon))$ T-gates to approximate the

$R_z(\theta)$ gate to a precision ε. Here, the gate synthesis error of each $R_z(\theta)$ ε_{CS} is set to be a small value compared to the total error. In practice, we should determine the resource overhead c_T based on the number of R_z gates in the quantum algorithm.

If extra ancillary qubits and entangling Clifford gates are allowed, we can further reduce the required T gates to

$$1.15 \log_2(1/\varepsilon) + 9.2, \tag{9.23}$$

using a repeat-until-success strategy [245]. However, this will introduce extra ancillary qubit requirements and more CNOT gate cost.

9.3.2 Basis operations

In this section, I discuss the basis operations involved in the circuit compilation. The implementation is composed of the following operations. In the standard block encoding procedure, the n-qubit Hamiltonian H is encoded in a $(n_L + n)$-qubit unitary, select(H)

$$\text{select}(H) := \sum_{l=1}^{L} |l\rangle\langle l| \otimes H_l. \tag{9.24}$$

If we denote,

$$|G\rangle := \text{PREP}\, |0\rangle^{n_L} = \frac{1}{\sqrt{\lambda}} \sum_{l=1}^{L} \sqrt{\alpha_l} |l\rangle, \tag{9.25}$$

then we have

$$H = \lambda(\langle G| \otimes I)\, \text{select}(H)(|G\rangle \otimes I), \tag{9.26}$$

which indicates that H is block-encoded into select(H). Here, PREP encodes the amplitude into the state on the ancillary space, and it is also referred to as the amplitude-encoding unitary or PREPARE operation in literature.

The reflection unitary R can be expressed as

$$R := (I - 2|0\rangle\langle 0|) \otimes I, \tag{9.27}$$

where the operation $(I - 2|0\rangle\langle 0|)$ is defined on the ancillary space with dimension n_L.

9.3.3 Gate cost in the index enumeration circuit

In this section, I follow the circuit construction in Ref. [239] to build the amplitude encoding operation $B(x)$ and controlled select operation C-select(H). Let us first consider the gate cost of the following operation,

$$\text{C}-\text{select}(X) = \sum_{a=0}^{1} |a\rangle\langle a| \otimes \sum_{l=1}^{L} |l\rangle\langle l| \otimes (X_l)^a\,, \tag{9.28}$$

which is a a major component of both $B(x)$ and C-select(H), where $X_l \in \{I, X\}$ is a single-qubit Pauli operator. The value of X_l depends on the value of l stored in the classical register. We can

regard C-select(X) as a simplified version of C-select(H), where $H = \sum_{l=1}^{L} X_l$ is a single-qubit Hamiltonian where X_l is either I or X, based on the storage in the classical register.

A 'sawtooth' circuit is constructed to realise the C-select(X) gate, which is called indexed operation in Sec. III in Ref. [239]. In the fully simplified circuit of C-select(X) in Fig. 7 in Ref. [239], we need $(L-1)$ computing AND operations, $(L-1)$ uncomputing AND operations, L control-X_l gates, and $(L-1)$ extra CNOT gates. If we decompose the computing and uncomputing AND operations based on Fig. 4 in Ref. [239], and synthesise all the gates to Clifford $+T$ gates, we will have the following observation.

Proposition 10 (Gate cost in the index enumeration circuit)**.** *If we construct the index enumeration circuit* C-*select*(X) *defined in Eq. (9.28) following the sawtooth way in Ref. [239] and synthesise all the gates to CNOT gates, single-qubit Clifford gates and T gates, then we can realise* C-*select* (X) *using* $(6L-5)$ *CNOT gates,* $(4L-4)T$ *gates, and* $(2L-2)$ *Hadamard gates.*

9.3.4 Gate cost in amplitude encoding, select gates, and reflection gates

At this point the gate cost in amplitude encoding, select gates, and reflection gates can be estimated. The amplitude-encoding unitary B realises the following transformation,

$$B|0\rangle = \sum_{l=1}^{L} \sqrt{\omega_l}|l\rangle\,|\mathrm{temp}_l\rangle$$

where

$$\omega_l = \alpha_l/\lambda,$$

is the normalised amplitude of the Hamiltonian. Following Ref. [239], it is assumed that it is allowed to introduce temporary storage $|\mathrm{temp}_l\rangle$ during the amplitude encoding. This will not cause problems so long as we finally disentangle $|\mathrm{temp}_l\rangle$ during the implementation of $B^\dagger$.

The dominant subroutine of the PREPARE circuit is the SUBPREPARE circuit defined in Eq. (48) in Ref. [239], which realises the amplitude encoding to different orbitals, ignoring the spin information first. Let us first ignore the detailed structure of the Hamiltonian H with respect to different spins. Given this, we can now treat the SUBPREPARE circuit as the PREPARE circuit. The essential idea is to first prepare ancillaries with uniformly distributed coefficients over indices l and then use a pre-determined binary representation of a probability keep_l, to perform a controlled-swap on the amplitude register l and another predetermined amplitude location alt_l. With well-designed values of swap probability keep send swap location alt_l, we can use the circuit in Fig. 11 in Ref. [239] to realise the SUBPREPARE circuit.

Suppose we want to realise the amplitude encoding with an accuracy of ε_{AE}, that is, to realise the following transformation,

$$B^{\varepsilon_{AE}}|0\rangle = \sum_{l=1}^{L} \sqrt{\tilde{\omega}_l}|l\rangle\,|\mathrm{temp}_l\rangle \tag{9.29}$$

where $\tilde{\omega}_l$ is an n_{AE} bit approximation to the true value ω_l

$$|\tilde{\omega}_l - \omega_l| \leq \varepsilon_{AE}, \quad l = 1, \ldots, L.$$

and we require

$$n_{AE} := \lceil -\log_2 \varepsilon_{AE} \rceil .$$

Due to the relation of the rescaled spectrum by block encoding, we have the relation

$$\varepsilon_{\text{PREP}} = \frac{\varepsilon}{\lambda}. \tag{9.30}$$

The relation of the amplitude encoding error and PREP error could be derived by considering the norm of the Hamiltonian. A simple relation is given by

$$\varepsilon_{AE} \sim \frac{\varepsilon_{\text{PREP}}}{L}. \tag{9.31}$$

As shown in Fig. 11 in Ref. [239], we need to introduce at least $2n_{AE}+2n_L+1$ extra ancillary qubits, with $n_L := \lceil \log_2 L \rceil$. To simplify the gate cost, we assume L is a power of 2. In this case, the first layer of the circuit in Fig. 11 in Ref. [239] can be realised using Hadamard gates. If L is not a power of 2, additional quantum resources are required.

The second and the third layer of the circuit requires the QROM circuit in Fig. 10 in Ref. [239], which is a modified version of the index enumeration circuit C-select(X) defined in Eq. (9.28).

Based on Proposition 10, the second layer of date loading requires $5(L-1) + L(n_L + n_{AE})$ CNOT gates, $4(L-1)$ T gates. The third layer is a coherent inequality test, which requires $(n_{AE}-1)$ AND and uncomputing-AND operations, and additional $6(n_{AE}-1)$ CNOT gates plus 1 Toffoli gate. Thus, it requires $11n_{AE} - 5$ CNOT gates and $4n_{AE} + 3$ T gates.

The fourth layer is a Fredkin gate, which is a controlled swap gate. This gate can be synthesised into Clifford + T gates using $8n_L$ CNOT gates and $7n_L T$ gates (see Fig. 5 in [246]).

Proposition 11 (Ancillary qubit and gate cost in the second-type amplitude encoding operation)**.** *If the unitary $B^{\varepsilon_{AE}}$ defined in Eq. (9.29) is synthesised to CNOT gates, single-qubit Clifford gates and T gates, the ancillary qubits and gate cost of $B^{\varepsilon_{AE}}$ are listed as follows:*

1. $2n_{AE} + 2n_L + 1$ *extra ancillary qubits;*
2. $n_L(L+8) + n_{AE}(L+11) + 5L - 10$ *CNOT gates;*
3. $4\,(L + n_{AE}) + 7n_L + 3$ *T gates;*

Here, $n_{AE} := \lceil -\log_2 \varepsilon_{AE} \rceil$ *and* $n_L := \lceil \log_2 L \rceil$.

Now, let us analyse the gate cost of the C-select(H) gate. A straightforward implementation of the C-select(H) gate is to replace of X_l gate in C-select(X) defined in Eq. (9.28) to multi-qubit Pauli gates P_l. The gate cost for the lattice Hamiltonian is shown in Corollary 3.

Corollary 3 (Gate cost in the C-select(H) operation of the lattice model)**.** *If we construct the controlled-select circuit* C-select(H) *of the lattice model following the 'sawtooth' way in Ref. [239] and synthesise all the gates to CNOT gates, single-qubit Clifford gates and T gates, then the approximate gate cost of* C-select(H) *is using* $5(L-1)+\mathrm{wt}(H)$ *CNOT gates and* $(4L-4)$ *T gates.*

For instance, consider the transverse-field Ising model

$$H = J\sum_i \sigma_i^z \sigma_{i+1}^z + h\sum_i \sigma_{i+1}^x, \tag{9.32}$$

with a periodic boundary condition. Corollary 3 directly applies here.

Next, I consider the gate cost for fermionic Hamiltonians, which admit a general form in second quantisation,

$$H = \sum_{pq} T_{pq} a_p^\dagger a_q + \sum_{pqrs} V_{pqrs} a_p^\dagger a_q^\dagger a_r a_s \tag{9.33}$$

with the creation and annihilation operator $a^\dagger$ and a. The fermionic Hamiltonian can be mapped to a qubit form by the Jordan-Wignar transformation. To further improve the gate cost in a fermionic Hamiltonian, Ref. [239] introduces an accumulator during the Pauli gate query process (Sec. IIIB and Fig. 8 in Ref. [239]). The accumulator will 'accumulate' the effect of the Pauli operators accessed in the previous data queries and consequently save the CNOT gate cost. Using this improved select operation, we are able to reduce the CNOT cost for each Pauli operator P_l to a constant independent of the weight of P_l. I will use an optimistic estimate of the CNOT gate cost for controlled-P_l operations, which is $5(L-1)+3L = 8L-5$. Finally, gate cost in the reflection operation is presented in Proposition 12.

Proposition 12 (Ancillary qubit and gate costs in the reflection operation [247])**.** *If we construct reflection operation $I-2|0\rangle\langle 0|$ on n qubits following methods in Proposition 4 in [247] and synthesis all the gates to CNOT gates, single-qubit Clifford gates and T gates, then the approximate ancillary qubit and gate costs are listed as follows:*

1. $\lceil \frac{n-3}{2} \rceil$ *ancillary qubits,*
2. $(6n-12)$ *CNOT gates,*
3. $(8n-17)T$ *gates.*

9.3.5 Gate cost for Trotter formulae

The basic setup for the first-order and higher-order Trotter formula has been presented in Section 2.4. In the following, the gate cost for Trotter formulae is provided.

9.3.5.1 Simple Trotter error bound

Lemma 1 (Simple Trotter error bound for the $2k$th-order Trotter formula [144]). *Let $H = \sum_{l=1}^{L} H_l$ be a Hamiltonian consisting of L summands and $t \geq 0$. We denote*

$$\begin{aligned} a_{2k}(\nu) &:= 2\frac{\left(2\cdot 5^{k-1}L\Lambda t\right)^{2k+1}}{(2k+1)!\nu^{2k+1}}e^{2\cdot 5^{k-1}L\Lambda t/\nu} \\ b_{2k}(\nu) &:= \frac{L^{2k}\left(2\cdot 5^{k-1}\Lambda t\right)^{2k+1}}{(2k-1)!\nu^{2k+1}}e^{2\cdot 5^{k-1}L\Lambda t/\nu}, \end{aligned} \tag{9.34}$$

where $k \geq 1, \nu$ is the time segment number. If we set the segment number ν to be

$$\nu_{2k}^{\text{det}} = \min\left\{\nu \in \mathbb{N} : \frac{\nu}{2}a_{2k}(\nu) \leq \varepsilon\right\} \tag{9.35}$$

for the deterministic first-order Trotter formula, or set ν to be

$$\nu_{2k}^{random} = \min\left\{\nu \in \mathbb{N} : \frac{\nu}{2}\left(a_{2k}(\nu)^2 + 2b_{2k}(\nu)\right) \leq \varepsilon\right\}, \tag{9.36}$$

for the randomised first-order Trotter formula, then the spectral norm distance of the resulting simulation channel to the unitary channel of e^{-iHt} is at most ε.

From Eq. (9.35), the time segment can be roughly approximated by

$$\nu_{2k} \leq \frac{\left(2\cdot 5^{k-1}L\Lambda t\right)^{1+\frac{1}{2k}}}{((2k+1)!)^{\frac{1}{2k}}\varepsilon^{\frac{1}{2k}}} \tag{9.37}$$

and a more tight segment number can be calculated by Eq. (9.35).

9.3.6 Overview of ground state property estimation by quantum signal processing

In this section, I introduce the key ingredient of the near-optimal ground state preparation algorithm proposed by [31]. Their method relies on the block encoding of a nonunitary matrix in the quantum circuit. For ease of explanation, I follow the notation and convention used in [31].

A matrix $A \in \mathbb{C}^{N\times N}$ where $N = 2^n$ can be encoded in the upper-left corner of an $(n_L + n)$-qubit unitary matrix if

$$\|A - \alpha(\langle 0^{n_L}| \otimes I)U(|0^{n_L}\rangle \otimes I)\|_2 \leq \varepsilon. \tag{9.38}$$

and U is referred to as an $(\alpha, n_L, \varepsilon)$-block-encoding of A. In this chapter, I consider the Hamiltonian written in a linear combination of unitaries as

$$H = \sum_{l=1}^{L} \alpha_l H_l$$

In the standard block encoding procedure, the n-qubit Hamiltonian H can be explicitly block-encoded into $U_H := \text{PREP}^\dagger \cdot \text{select}(H) \cdot \text{PREP}$,

$$H = \lambda(\langle G| \otimes I)\,\text{select}(H)(|G\rangle \otimes I). \tag{9.39}$$

To simplify the results, let us denote the number of CNOT gates and T gates required for C-select operation as S_{CNOT} and S_{T}, and the number of CNOT gates and T gates required for PREP operation as P_{CNOT} and P_{T}, respectively.

I summarise the state preparation based on QSP below.

1. Obtain the $(\lambda, n_L, 0)$-block-encoding of a Hermitian matrix $H = \sum_k E_k |\psi_k\rangle \langle\psi_k| \in \mathbb{C}^{N\times N}$, $N = 2^n$, $E_k \leq E_{k+1}$. This block encoding is constructed by U_H.

2. Construct a $(\lambda + |\mu|, n_L + 1, 0)$-block-encoding of matrix $H - \mu I$ using of [34, Lemma 29] for any $\mu \in \mathbb{R}$.

3. Construct an $(1, n_L + 2, \varepsilon)$-block-encoding of

$$R_{<\mu} = \sum_{k:\mathbb{E}_k<\mu} |\psi_k\rangle \langle\psi_k| - \sum_{k:E_k>\mu} |\psi_k\rangle \langle\psi_k| .$$

   This is realised by constructing a block encoding of the sign function $-S(\frac{H-\mu I}{\alpha+|\mu|}; \delta, \varepsilon)$ for any δ and ε where $S(\cdot, \delta, \varepsilon)$ is the sign function of degree $d = \frac{e}{2\delta} \ln(32\pi^{-1/2}\varepsilon^{-1})$. Note that if we further assume that $\Delta/2 \leq \min_k |\mu - \mathbb{E}_k|$, then let $\delta = \frac{\Delta}{4\lambda}$, all the eigenvalues of $-S(\frac{H-\mu I}{\lambda+|\mu|}; \delta, \varepsilon)$ are ε-close to either -1 or 1, and thus $-S(\frac{H-\mu I}{\lambda+|\mu|}; \delta, \varepsilon)$ is ε-close, in operator norm, to the reflector about the direct sum of eigen-subspaces corresponding to eigenvalues smaller than μ:

4. Using the block encoding of $R_{<\mu}$, we can construct an $(1, n_L + 3, \varepsilon/2)$ block encoding of the projection operator $P_{<\mu} := \frac{1}{2}(R_{<\mu} + I)$.

5. Obtain the ground state with success probability close to 1 by amplitude amplification.

6. Observable estimation.

Here, I apply the amplitude amplification by using the Reflector R instead of the Projector P, which introduces the cost from the reflection operation by Proposition 12.

The gate cost using quantum signal processing is shown in Section F.2 in Appendix F.

9.4 Resource estimations

9.4.1 Compilation to standard gates

In this section, I elaborate on compilation of the quantum circuit to standard quantum gates that are easy to implement. A controlled-Pauli-rotation gate can be realised by $2(wt(H) - 1)$ CNOT gates, a controlled single-qubit Pauli rotation gate, and some single-qubit Clifford gates. Furthermore, we can decompose the controlled single-qubit Pauli rotation gate to two single-qubit Z-axis rotation gates and two CNOT gates. Therefore, we need $(s + 2)$wt(H) CNOT gates and 2 single-qubit Pauli rotation gates for each time segment.

Note that the RLCU algorithm is not a deterministic state preparation that cannot prepare the target state σ. However, when we focus on the property estimation of the target state, the estimator has a similar performance in relation to the normal Hamiltonian simulation methods: the sample complexities of the former algorithms and RCLU algorithm to learn the observable properties of the state are similar, as long as the norm of the estimator is a constant. This is guaranteed by Proposition 7. To estimate the gate costs of RLCU algorithm, we set the normalisation factor $\mu = 2$ to ensure that the sample complexity of the RCLU algorithm is similar to other quantum simulation algorithms.

When considering a real physical model, such as a chemistry problem, the coefficient of the Hamiltonian is constructed by calculating an integral and represents a feature of the quantum system. Due to finite precision, there will be an amplitude encoding error when we perform the PREP operation. To ensure that the amplitude encoding error is less than a threshold, more qubits are required to encode the coefficients.

Here, the amplitude encoding error is included in the analysis when aiming at realistic application. That is the reason why we require more qubits for the algorithms that involve amplitude amplification. It is worth noting that in several cases, the absolute value of the coefficients may not be highly important for capturing the actual physics, and hence the amplitude encoding error does not have to be taken into account.

9.4.2 Numerical results

As a numeric test, I choose the transverse-field Ising model and the one-dimensional Fermi-Hubbard model. The quantum circuit is synthesised to CNOT gates, single-qubit Clifford gates and non-Clifford gate (including single-qubit Z-axis rotation R_z gates and T gates). The requirements for the gate number in the representative algorithms are estimated. Since the central objective is for application in the early FTQC or NISQ era, I mainly focus on the circuit depth in a single-run experiment. The amplitude amplification is therefore not considered in the algorithms, which can deterministically prepare the state closer to true ground state yet at the cost of a deeper circuit.

The algorithms compared in this section can be divided into three classes: phase estimation, quantum signal processing, and our method based on random linear combination of unitaries (RLCU) and random paired Taylor series sampling (RTS). The phase estimation method relies on Hamiltonian simulation, which can be realised by Trotter formulae methods (first-order, second-order, fourth-order, qDRIFT and the randomised version), the Taylor-series method, and the qubitisation method. For our method, I consider using the zeroth-order pairing, second-order pairing, and fourth-order pairing for Hamiltonian simulation. In the following, I first consider the comparison with Trotter-based methods, which are oracle-free and ancilla-free. We shall find that the fourth-order random Trotter has the optimal performance out of all the other Trotter

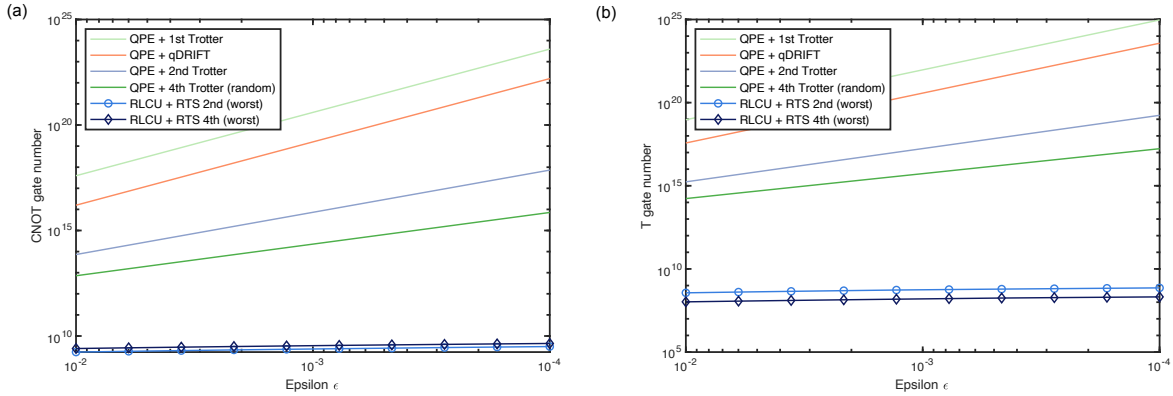


Figure 9.2: Resource estimation for the ground state property estimation for the 50-site Ising Hamiltonian, compared to the phase estimation combining Trotterisation. The gate count is estimated for realising a target precision ε. The energy gap is set as $\Delta = 0.01$ and the initial state overlap is set as $p_0 = 0.5$.

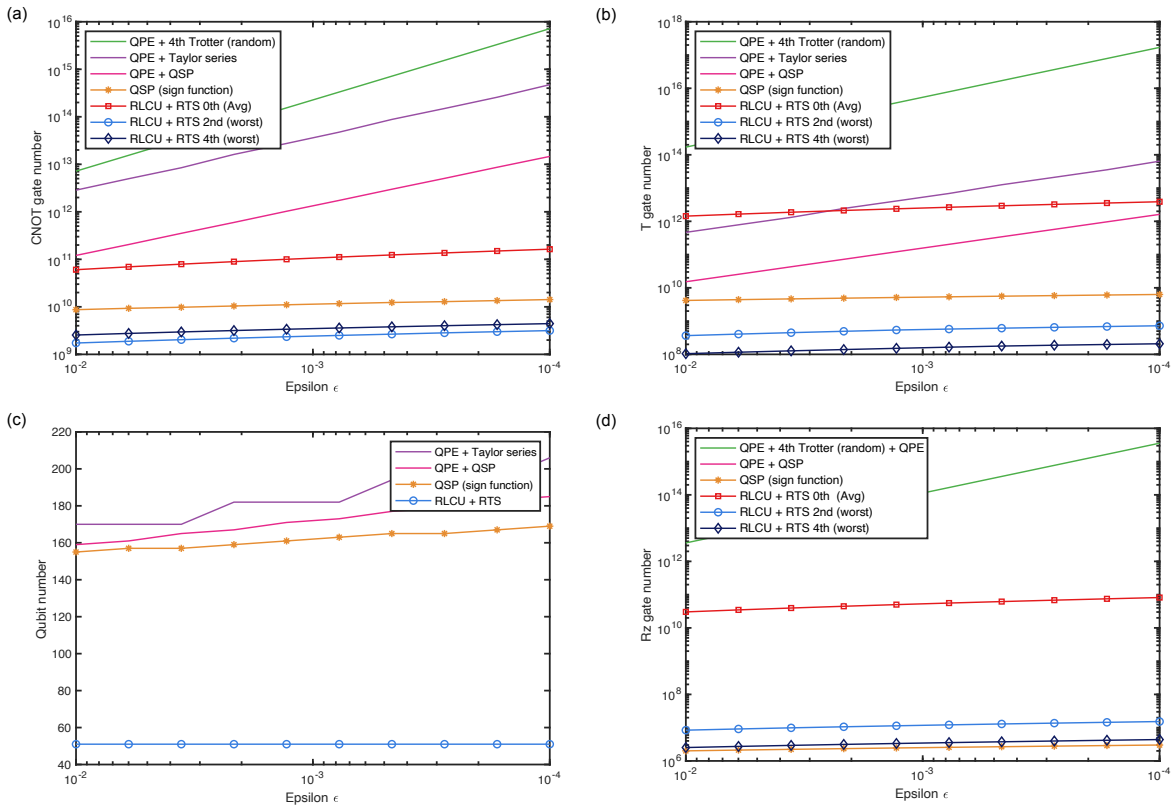


Figure 9.3: Resource estimation for the ground state property estimation for the 50-site Ising Hamiltonian. (a-d) show the ancillary qubit number, the number of CNOT gates, non-Clifford (single-qubit R_z rotation) gates, T gates, respectively. The parameter setup is the same as that in Figure 9.3.

methods in this scenario, which is consistent with the results in [37]. Then, I show a comparison between our method and more advanced methods, with a particular focus on the qubitisation

method, which achieves near optimal scaling in the task of ground state property estimation.

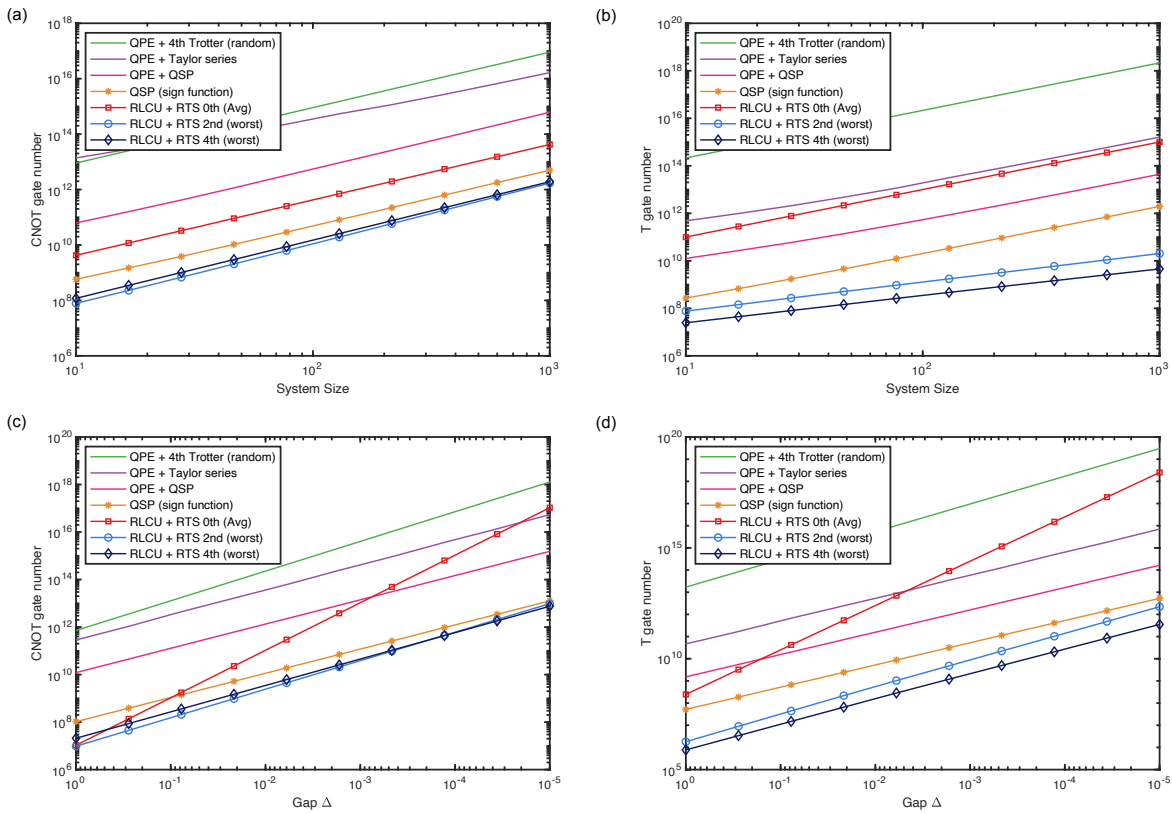


Figure 9.4: Resource estimation (the number of CNOT gates and T gates) for the ground state property estimation for the 1D Ising Hamiltonian. The gate count is estimated for realising a fixed precision $\varepsilon = 10^{-3}$. (a-b) shows the number of CNOT gates and T gates with an increasing system size with a fixed energy gap $\Delta = 0.01$. (a-b) shows the number of CNOT gates and T gates with a decreasing energy gap on 50 sites.

I first consider the task of ground state property estimation of a 50-site transverse-field Ising model. The comparison to Trotter-based methods are presented in Figure 9.2. We can find that our method shows a clear advantage over the phase estimation combining Trotterisation. The comparison for resource estimation with more advanced methods is shown in Figure 9.3. Figure 9.3(a,b,c,d) show the CNOT gate count, T gate count, total number of qubits, and the non-Clifford Z-axis rotation R_z gate count, respectively. Here, I consider synthesising R_z gates into T gates using a repeat-until-success strategy in Eq. (9.23). The second-order RTS shows advantages with respect to the CNOT gate count, while the fourth-order RTS shows advantages with respect to the T gate count. This aligns with our expectation, since the T gate mainly comes from the segment number. In the fourth-order pairing, the segment number is less than the zeroth-order or second-order pairing, while the usage of higher-order Trotter formula results in more CNOT gates. In the numerical simulation, I set the energy gap $\Delta = 0.01$ and the lower bound of the initial state overlap $\eta = 0.5$.

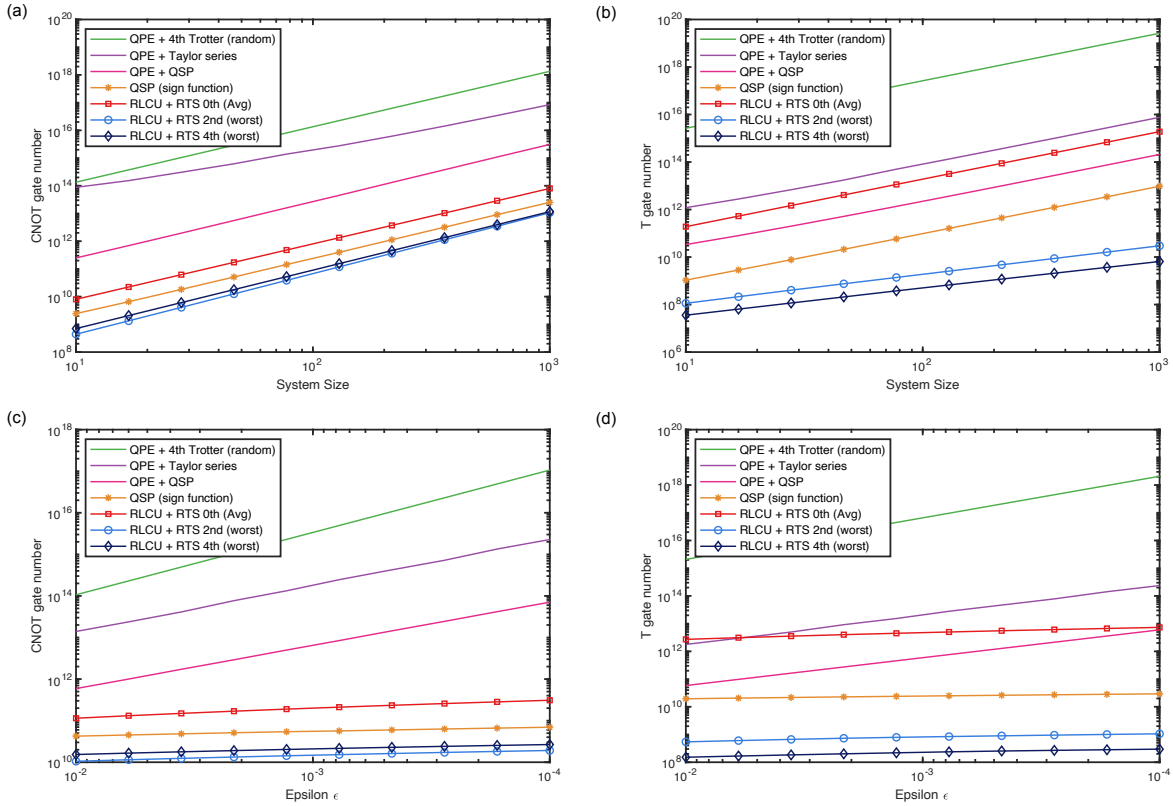


Figure 9.5: Resource estimation (the number of CNOT gates and T gates) for the ground state property estimation for the 1D Fermi-Hubbard model. (a-b) shows the number of CNOT gates and T gates with an increasing system size with a fixed energy gap $\Delta = 0.01$. (a-b) shows the number of CNOT gates and T gates with a decreasing energy gap with 50 sites.

Next, I study the resource scaling with respect to the system size and energy gap. Figure 9.4(a,b) show the CNOT gate count and T gate count with an increasing system size. Figure 9.4(c,d) show the CNOT gate count and T gate count with a decreasing energy gap, respectively.

The above results clearly show the advantages of our method over the existing methods in a wide range of problems of physical interest. Results for the one-dimensional Fermi-Hubbard Hubbard model are shown in Figure 9.5.

9.5 Outlook and discussion

In this chapter, I focus on the tasks of eigenenergy estimation and observable estimation on the eigenstate using a combination of RLCU and RTS. The paired Taylor-series with leading-order-rotation method for Hamiltonian simulation is employed such that it preserves the advantage of logarithmic dependence on the error, without requiring additional ancillary qubits. I further show a concrete gate count analysis on CNOT gate and T gate concerning the circuit synthesis.

For systems with a certain symmetry, we can directly measure the state overlap without a

Hadamard test circuit, and thus our method could be ancilla free. This is particularly useful for implementation on hardware since, in most algorithms, due to the restriction of connectivity, there could be a large overhead when compiling it into a nearest-neighbour operation.

One can also think of the resource requirement when considering certain topological layout of the qubits, such as the 2D lattice. The gate complexity could be reduced by fermionic swap networks or fast fermionic Fourier transform (FFFT) [43,188]. The optimal variant for quantum chemistry simulation is by using a split-operator formula with a fast fermionic Fourier transform [187]. We may embed these methods in the simulation of correction operations to avoid the use of long-range gate.

A more interesting direction is to apply the technique developed in this chapter to realistic molecules or materials. For instance, we may choose different bases, such as plane wave basis, to encode the Hamiltonian, which gives the number of terms in the Hamiltonian quadratically increasing with the system size N, $L = \mathcal{O}(N^2)$ [43]. This will essentially answer the question of quantum advantages in realistic applications.

The dependence on sparsity of the Hamiltonian, which appears due to the leading-order rotation, may be removed by considering a composite simulation strategy. In addition, we may consider measuring commuting terms simultaneously by basis rotation for observable estimation, such that the complexity of observable estimation could be further reduced.

Author contributions. This chapter is relevant to a manuscript under preparation [92]. In this work, P.Z. and J.S. initiated the project. J.S. and P.Z. developed the theoretical framework of eigenstate preparation. P.Z. developed Hamiltonian simulation by Trotter-LCU with input from J.S.. J.S. developed the framework of resource estimation and carried out numerical simulation. J.S. and P.Z. wrote the manuscript with input from V.V. and X.Y..

Chapter 10

Mitigating quantum process error and statistical error

In preceding chapters, I mainly discuss the algorithmic error, while the error from device imperfections and shot noise from finite measurement samples are neglected. Those errors appearing in the practical implementation could accumulate and affect the simulation accuracy. As pointed out in Section 1.4, in this NISQ era, we face severe challenges due to the limitation of quantum hardware technology. In this chapter, I discuss how to address these challenges in dealing with the device error due to unwanted interactions with environment and the statistical error when it comes to the practical implementation with imperfect noisy quantum devices in the NISQ era.

The error mitigation part is relevant to a published work [97], while the quantum state measurement part is relevant to the theoretical part in published works [87, 98].

10.1 Quantum process error mitigation

10.1.1 Background

With the experimental demonstration of quantum supremacy [11], whether current or near-future noisy intermediate-scale quantum (NISQ) devices are sufficient for realising quantum advantages in practical problems becomes one of the most exciting challenges in quantum computing [71]. Since NISQ devices have insufficient qubits to implement fault-tolerance, effective quantum error mitigation (QEM) schemes are crucial for suppressing errors to guarantee the calculation accuracy to surpass the classical limit. Among different QEM schemes via different post-processing mechanisms [81, 82, 116, 117, 149, 150, 238, 248–267], the probabilistic QEM method is one of the most effective techniques [149,150], which fully inverts noise effect by requiring a full tomography of the noise process and assuming noise independently appears either before or after each gate in a digital gate-based quantum computer. While these assumptions are adopted for many QEM schemes, realistic noise is more complicated. Specifically, since every gate is experimentally realised via the time evolution of quantum controls [8, 11, 192, 193, 268–273], noise happens along with the evolution, whose effect inevitably mixes with the gate or process and even scrambles nonlocally.

For example, as one of the major noises in superconducting qubits, crosstalk of multi-qubit gates originates from the imperfect time evolution with unwanted interactions [8, 192, 268, 273–275]. Therefore, such inherent dynamics-based and nonlocal noise effects make conventional QEM schemes less effective for practical NISQ devices. Meanwhile, a more natural and noise-robust computation model is via analogue quantum simulators [50–57, 57, 58, 234, 276–283], which directly emulate the target system without even implementing gates. It also remains an important open challenge to suppress errors for reliable medium- or large-scale analogue quantum simulators [7, 284–286].

In this chapter, I present QEM schemes without assumptions of gate-based circuits or simplified local noise models of each gate. Specifically, I introduce *stochastic error mitigation* for a continuous evolution with noise described by imperfections of the engineered Hamiltonian or super-operators induced from the interaction with the environment [7, 280, 286, 287]. Compared to existing methods, such as dynamical decoupling, which are generally limited to low frequency noise and small simulations [288–291], our work introduces a universal way to mitigate realistic noise under experiment-friendly assumptions. Our work considers continuous evolution of the system and assumes accurate single-qubit operations, which is applicable to all digital quantum simulators and various analogue simulators. Our method is compatible with existing QEMs, and its combination with Richardson extrapolation can be further leveraged to suppress errors in inaccurate model estimations and recovery operations. I numerically test our scheme for various Hamiltonians with energy relaxation and dephasing noise and a quantum circuit with two-qubit crosstalk noise. A resource estimation for near-term devices involving up to 100 qubits is conducted, which shows the feasibility of our QEM scheme in the NISQ regime.

10.1.2 Framework

In this section, I introduce the background of analogue quantum simulation (AQS) and digital quantum simulation (DQS) with noisy operations. In a digital gate-based quantum computer, the effect of noise is usually simplified as a quantum channel appearing either before or after each gate, whereas realistic noise occurring in the experimental apparatus is more complicated. Specifically, every gate in digital circuits or every process in analogue simulation is physically realised via a continuous real time evolution of a Hamiltonian and therefore errors can either inherently mix with the evolution — making it strongly gate or process dependent, or act on a multiple number of qubits — leading to highly nonlocal correlated effects (crosstalks). For instance, dominant errors in superconducting qubits are inherent system dephasing or relaxation, and coherent errors (or crosstalk) when applying entangling gates. While AQS are believed to be less prone to noise, this holds true mostly in comparison to DQS, and when considering an intermediate simulation scale [284], outcomes of AQS could be sensitive to noise (for example,

see theoretical studies on the sensitivity to errors [285, 286] and noisy simulation result [234] of AQS).

Since conventional quantum error mitigation methods are restricted to gate-based digital quantum computers and over-simplified noise models, they fail to work for realistic errors and general continuous quantum processes. For instance, owing to the restricted set of allowed operations in analogue quantum simulators, it is challenging to suppress or correct errors of a continuous process in this context [285] and it remains an open challenge to suppress errors for reliable medium- or large-scale quantum simulators [284]. Our work addresses this problem by considering a more general scenario of a continuous process with realistic noise models occurring in quantum simulators.

In particular, I introduce the model that describes either gate syntheses or continuous processes in digital or analogue simulation. We consider the ideal evolution of state $\rho_I(t)$ with a target Hamiltonian H_{sys} as

$$\frac{d\rho_I(t)}{dt} = -i[H_{\text{sys}}(t), \rho_I(t)]. \tag{10.1}$$

In practice, we map H_{sys} to a noisy controllable quantum hardware H_{sim}, whose time evolution is described by the Lindblad master equation of the noisy state $\rho_N(t)$ as

$$\frac{d\rho_N(t)}{dt} = -i[H_{\text{sim}}(t), \rho_N(t)] + \lambda\mathcal{L}_{\text{exp}}\big[\rho_N(t)\big]. \tag{10.2}$$

Here, $H_{\text{sim}} = H_{\text{sys}} + \delta H$ corresponds to coherent errors (such as crosstalk or imperfections of Hamiltonian) and $\mathcal{L}_{\text{exp}}[\rho] = \frac{1}{2}\sum_k(2L_k\rho L_k^\dagger - L_k^\dagger L_k\rho - \rho L_k^\dagger L_k)$ is the noise superoperator with error strength λ that describes inherent coupling with the environment (such as dephasing and damping) [7, 280]. Instead of assuming a local single-qubit noise channel of each gate in conventional QEM, we consider a local noise model by assuming local Lindblad terms. Note that local noise operators at instant time t can easily propagate to become global noise after integrating time, which may cause nonlocal noise effects in reality.

Suppose we are interested in measuring the state at time T with an observable O. The task of QEM is to recover the noiseless measurement outcome $\langle O\rangle_I = \text{Tr}[O\rho_I(t)]$ via noisy process. In general, it would be difficult to efficiently mitigate arbitrary noise with any noise strength. Here, it is assumed that the noise operators act weakly, locally and time-independently on small subsystems. Note that even though the coherent error δH and the Lindblad operators L_k act locally on the quantum system, the effect of errors propagates to the entire system after the evolution. Therefore, such global effects of noise cannot be effectively mitigated using the conventional quasi-probability method, which assumes a simple gate-independent error model described by single- or two-qubit error channels before or after each gate. We also assume that accurate individual single-qubit controls are allowed, which holds for digital NISQ devices where

single-qubit operations can achieve averaged fidelity of 99.9999% [292] whereas the record for two-qubit fidelity is three orders lower [11, 75, 293–295]. While not all analogue quantum simulators support individual single-qubit controls, they can indeed be achieved in various platforms with superconducting qubits [8, 296–299], ion trap systems [9, 234, 300], and Rydberg atoms [301]. Therefore, our framework is compatible with various practical NISQ devices. In the following, I focus on qubit systems and assume time-independent noise. The discussion can be naturally generalised to multi-level systems, as well as general time-dependent noise (See Section G.1.1).

10.2 Continuous quantum error mitigation

Quantum gate in digital circuits or joint evolution process in analogue simulation is physically realised via a continuous real time evolution of a Hamiltonian. This section extends the QEM method to a more practical scenario, and shows how to mitigate errors for these inherent dynamics-based and nonlocal noise in practical noisy quantum devices. To motivate the discussion, I first introduce 'continuous' QEM as a preliminary scheme, as shown in Figure 10.1(a). Consider a small time-step δt, the discretised evolution of Eq. (10.1) and Eq. (10.2) can be represented as

$$\rho_\alpha(t+\delta t) = \mathcal{E}_\alpha(t)\rho_\alpha(t). \tag{10.3}$$

Here $\alpha = I, N$ and $\mathcal{E}_\alpha(t)$ denotes the ideal ($\alpha = I$) or noisy ($\alpha = N$) channel that evolves the state from t to $t+\delta t$ within a short time period δt. We can find a recovery operation $\mathcal{E}_Q$ that approximately maps the noisy evolution back to the noiseless one as

$$\mathcal{E}_I(t) = \mathcal{E}_Q\mathcal{E}_N(t) + \mathcal{O}(\delta t^2). \tag{10.4}$$

The operation $\mathcal{E}_Q$ is in general not completely positive, hence cannot be physically realised by a quantum channel. Nevertheless, similar to probabilistic QEM for discrete gates [149, 150], we can efficiently decompose $\mathcal{E}_Q$ as a linear sum of a polynomial number of physical operators $\{\mathcal{B}_j\}$ that are tensor products of qubit operators,

$$\mathcal{E}_Q = c\sum_j \alpha_j p_j \mathcal{B}_j, \tag{10.5}$$

with coefficients $c = 1 + \mathcal{O}(\delta t)$, $\alpha_j = \pm 1$, and a normalised probability distribution p_j. The decomposition and its optimisation via linear programming is shown in Section G.1.4 and Section 2.5.3. Under this decomposition, the whole ideal evolution from 0 to T can be mathematically decomposed as

$$\prod_{k=0}^{n-1}\mathcal{E}_I(k\delta t) = C\sum_{\vec{j}}\alpha_{\vec{j}}p_{\vec{j}}\prod_{k=0}^{n-1}\mathcal{B}_{j_k}\mathcal{E}_N(k\delta t) + \mathcal{O}(T\delta t), \tag{10.6}$$

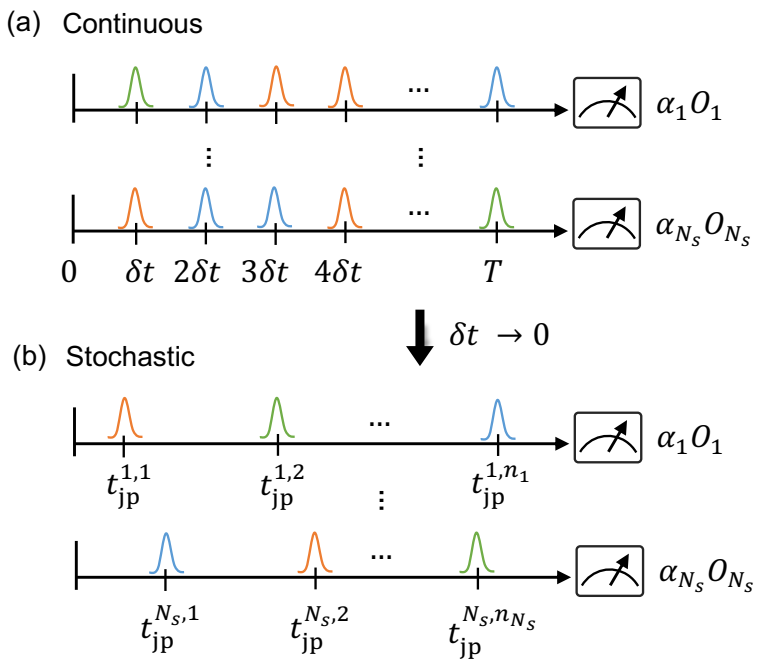


Figure 10.1: (a) Continuous QEM. With discretised time step δt, each recovery operation is weakly and 'continuously' acted after each noisy evolution of time δt. Here different colours represent different recovery operations. The output state is measured and repeated to obtain N_s outcomes $\{O_m\}$, and their average corresponds to the error mitigated outcome. (b) Stochastic QEM. We can equivalently realise (a) by $\delta t \to 0^+$ and randomly applying a small number n_m of strong recovery operations as in Algorithm 7. The time $\{t_{\text{jp}}^{m,k}\}_m$ to apply recovery operations of the mth run are predetermined, which can be further pre-engineered into the original evolution via a noisy time evolution of a modified Hamiltonian.

where $n = T/\delta t$, $C = c^n$, $\alpha_{\vec{j}} = \prod_{k=0}^{n-1} \alpha_{j_k}$, $p_{\vec{j}} = \prod_{k=0}^{n-1} p_{j_k}$, and $\vec{j} = (j_1, \ldots j_{n-1})$. Denote the ideally evolved state as $\rho_I(T) = \prod_{k=0}^{n-1} \mathcal{E}_I(k\delta t)\rho(0)$ and the noisily evolved and corrected state as $\rho_{Q,\vec{j}}(T) = \prod_{k=0}^{n-1} \mathcal{B}_{j_k}\mathcal{E}_N(k\delta t)\rho(0)$, we can approximate the ideal state $\rho_I(T)$ as a linear sum of noisy states as

$$\rho_I(T) = C \sum_{\vec{j}} \alpha_{\vec{j}} p_{\vec{j}} \rho_{Q,\vec{j}}(T) + \mathcal{O}(T\delta t). \tag{10.7}$$

When measuring an observable O of the ideal state, the ideal measurement outcome $\langle O \rangle_I = \mathrm{Tr}[\rho_I(T)O]$ is also approximated as a linear sum of the noisy measurement outcomes $\langle O \rangle_{Q,\vec{j}} = \mathrm{Tr}[\rho_{Q,\vec{j}}(T)O]$ as

$$\langle O \rangle_I = C \sum_{\vec{j}} \alpha_{\vec{j}} p_{\vec{j}} \langle O \rangle_{Q,\vec{j}} + \mathcal{O}(T\delta t). \tag{10.8}$$

In practice, we can randomly prepare $\rho_{Q,\vec{j}}(T)$ with probability $p_{\vec{j}}$, measure the observable O, and multiply the outcome with the coefficient $C\alpha_{\vec{j}}$. Then the average measurement outcome $\langle O \rangle_{Q,\vec{j}}$ of the noisy and corrected states $\rho_{Q,\vec{j}}$ approximates the noiseless measurement outcome.

To measure the average outcome to an additive error ε with failure probability δ, we need $N_s \propto C^2 \log(\delta^{-1})/\varepsilon^2$ samples according to the Hoeffding inequality. Since the number of samples needed given access to $\rho_I(T)$ is $N_0 \propto \log(\delta^{-1})/\varepsilon^2$, the error mitigation scheme introduces a sampling overhead C^2, which can be regarded as a resource cost for the stochastic QEM scheme. The overhead scales as $C^2(T) = \exp(\mathcal{O}(\lambda T))$ given noisy strength λ and evolution time T. Here, a normalisation λ is chosen so that the contribution from $\mathcal{L}_{\text{exp}}$ is bounded by a constant. Therefore, the condition that the scheme works efficiently with a constant resource cost is $\lambda T = \mathcal{O}(1)$. By regarding λ as the error rate, the condition can be intuitively interpreted as that the total noise rate is a constant, aligning with the result for conventional QEM.

This error mitigation scheme works for general errors, such as correlated stochastic noise and unwanted interactions between (a small number of) multiple qubits. In Section G.1.1, we provide more details of the continuous error mitigation, including the decomposition of recovery operations and the resource cost for this method. In addition, this scheme can be naturally applied to multi-level systems when we can prepare the basis operations $\{\mathcal{B}_j\}$ for them.

10.3 Stochastic quantum error mitigation

In practice, it could be challenging to 'continuously' interchange the noisy evolution and the recovery operation within a sufficiently small time-step δt. Since $\mathcal{E}_I(t) \approx \mathcal{E}_N(t)$ and the recovery operation at each time is almost an identity operation

$$\begin{aligned} \mathcal{E}_Q &= (1 + q_0 \delta t)\mathcal{I} + \sum_{j\geq 1} q_j \delta t \mathcal{B}_j \\ &= c\left(p_0 \mathcal{I} + \sum_{j\geq 1} \alpha_j \tilde{p}_j \delta t \mathcal{B}_j\right), \end{aligned} \tag{10.9}$$

with $\mathcal{B}_0$ being the identity channel $\mathcal{I}$. The probability to generate the identity operation $\mathcal{I}$ and $\mathcal{B}_i$ $(i \geq 1)$ is $p_0 = 1 - \sum_{j\geq 1} \tilde{p}_j \delta t = 1 - \mathcal{O}(\delta t)$ and $\tilde{p}_j = p_j/\delta t = \mathcal{O}(1)$, $c = 1 + (q_0 + \sum_{j\geq 1} |q_j|)\delta t$. In addition, the parity α_0 for $\mathcal{B}_0 = \mathcal{I}$ is always unity, and the parity α_i corresponding to $\mathcal{B}_j$ $(i \geq 1)$ is equal to $\text{sign}(q_j)$.

We can further apply the Monte Carlo method to stochastically realise the continuous recovery operations, as shown in Figure 10.1(b). Specifically, we initialise $\alpha = 1$ and randomly generate $q \in [0, 1]$ at time $t = 0$. Then evolve the state according to the noisy evolution $\mathcal{E}_N$ until time t_{jp} by solving $p(t_{\text{jp}}) = q$ with $p(t) = \exp(-\Gamma(t))$ and $\Gamma(t) = t \sum_{j\geq 1} \tilde{p}_j$. At time t_{jp}, we generate another uniformly distributed random number $q' \in [0, 1]$, apply the recovery operation $\mathcal{B}_j$ if $q' \in [s_{j-1}, s_j]$, and update the coefficient as $\alpha = \alpha_j \alpha$. Here $s_j(t) = (\sum_{i=1}^{j} \tilde{p}_i)/(\sum_{i=1}^{N_{\text{op}}} \tilde{p}_i)$, N_{op} is the number of basis operations, and the sum omits the identity channel. Then, we randomly initialise q, and repeat this procedure until time reaches T. On average, we can prove that the *stochastic* QEM scheme is equivalent to the 'continuous' one (see the author's paper [97]

for the proof). In contrast, the stochastic QEM does not assume time discretisation, and only requires to randomly apply a few recovery operations, scaling linearly to the total noise strength as $O(\lambda T)$ (See Section G.1.1). We can insert the recovery operations by 'pausing' the original noisy evolution. Alternatively, since we can determine the time t_{jp} and the recovery operations before the experiment, they can be pre-engineered into the original evolution. Therefore, we can effectively implement stochastic QEM via the noisy time evolution of Eq. (10.2) with an adjusted Hamiltonian.

The stochastic error mitigation scheme is summarised in Algorithm 7 in Appendix G.

At the jump time t_{jp}, we apply the basis operation other than the identity operation. We can determine the basis operation by generating another uniformly distributed random number $q' \in [0, 1]$. If $q' \in [s_{k-1}, s_k]$, we set the basis operation to $\mathcal{B}_k$, where $s_k(t) = (\sum_{j=1}^{k} \tilde{p}_j)/(\sum_{j=1}^{N_{\text{op}}} \tilde{p}_j)$ and N_{op} is the number of the basis operations.

For time-independent noise, the jump time can be simply determined as $t_{\text{jp}} = -\log(q)/\sum_{i\geq 1} \tilde{p}_i$ with q randomly generated from $[0, 1]$, and thus we can generate the sequence of jump time a priori. Given evolution time T, the average number of recovery operations is proportional to $\mathcal{O}(\lambda T)$. In the numerics, the average number of recovery operations is about 0.3 times per evolution on average given a realistic noise model and simulation task.

I will discuss its implementation for both analogue quantum simulation and digital gate-based quantum simulation in Section 10.4.

10.4 Implementation of the scheme with analogue and digital quantum simulators

In this section, I discuss the implementation of our scheme with analogue and digital quantum simulators. To implement stochastic error mitigation with an analogue quantum simulator, we insert the single-qubit recovery operations at each jump time, which is pre-determined by Algorithm 7. As the evolution of most quantum simulators is based on external pulses, such as trapped ions and superconducting qubits, it would be practically feasible to interrupt the continuous evolution by simply turning off the external pulses and then turning on the single-qubit recovery pulses. The joint evolution and the single-qubit dynamics can be pre-engineered as a modified evolution of AQS, as shown in Figure 10.2. In practice, when turning on/off the joint evolution cannot be realised in a short time, we can alternatively apply single-qubit recovery pulses with a short duration and a sufficiently strong intensity compared to the parameters of the AQS Hamiltonian, as shown in Figure 10.2(b) II. This is similar to the banged analogue-digital quantum computing protocol introduced in Refs. [302, 303], which implements single-qubit gates without turning off the background Hamiltonian. In this case, when single-qubit rotations are performed in a time δt that is much smaller than the timescale of the joint evolution, the additional error per single-qubit rotation introduced by the background evolution of the

Hamiltonian is on the order of $\mathcal{O}(\delta t^3)$. Therefore, errors induced from the error mitigation procedure could be very small, and they could be further mitigated via the hybrid approach in the Section 10.5.

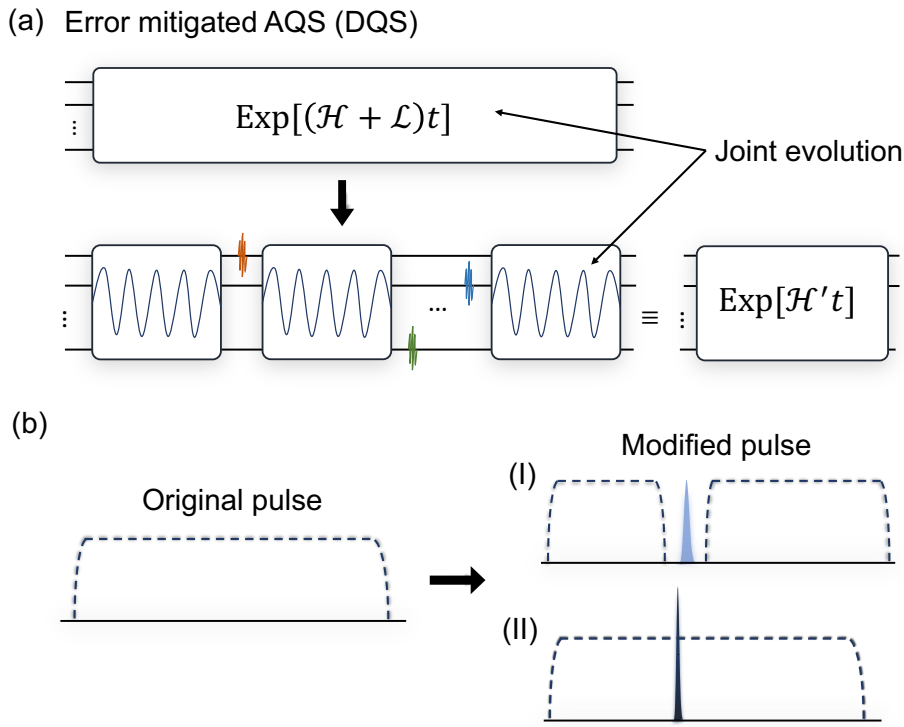


Figure 10.2: (a) Schematic diagram of the error mitigated AQS or DQS with controllable single-qubit operations. The AQS or DQS is realised by a continuous process under the ideal driving Hamiltonian. We denote the noisy unitary and stochastic processes by $\mathcal{H}$ and $\mathcal{L}$ with the super-operator formalism. Our work considers joint dynamics of all qubits sandwiched with a small number ($\mathcal{O}(1)$) of pre-engineered single-qubit dynamics to mitigate the errors accumulated in the evolution, which can also be regarded as a modified evolution $\mathcal{H}'$. (b) Two schemes of the error-mitigated process or gate with modified pulse sequences. Dashed lines represent the original pulse that constructs the target process or multi-qubit gates. Provided controllable drive that could be freely turned on/off, we can synthesise the error mitigated process/gate by modifying the original pulse sequence as shown in scheme I, which corresponds to (a). In the case of restricted driving operations, we can alternatively apply a strong and fast single-qubit pulse to the original pulse to mitigate either process errors or gate errors as shown in scheme II. Note that scheme II could similarly be applied in (a), which introduces a negligible error of $\mathcal{O}(\delta t^3)$ when each single-qubit gate is implemented in $\delta t \ll T$.

On the other hand, the stochastic error mitigation scheme could be naturally implemented on a digital gate-based quantum computer. Digital gates are experimentally realised via continuous pulse sequences [8, 11, 192, 193, 268–273], thus we can construct the error mitigated gates by modifying the original pulse sequence with the pre-determined pulses (recovery operations). A similar process has been experimentally demonstrated in Ref. [271], where the effect of the new pulse sequence is to effectively mitigate the unwanted terms in the driving Hamiltonian. Our

QEM method can be used to eliminate the general coherent and incoherent errors of the gates to achieve high gate fidelity. Therefore, provided control of pulse sequence, we can engineer the pulse sequence as shown in Figure 10.2(b), and prepare the error mitigated gates to perform quantum computing or quantum simulation tasks. If the driving operations are restricted, we can similarly apply the fast single-qubit pulse to the original pulse to mitigate the errors, where the additional error induced in this process is on the order of $\mathcal{O}(\delta t^3)$, as shown in the bottom of Figure 10.2(b).

As shown in Figure 10.2, the exact quantum hardware that we need to implement our error mitigation scheme actually lies in between fully analogue simulation (we can control only all qubits with a predetermined Hamiltonian) and a fully digital quantum computer (we can apply any operation on a small number of qubits). The scenario can be regarded as a modified analogue-digital quantum simulator, where we need only to apply strong local gates along the background dynamics (no matter whether it is digital or analogue). In contrast to a fully AQS, single-qubit operations are inserted into real-time dynamics. The joint evolution and the single-qubit dynamics can be pre-engineered by using Algorithm 7. The readers could regard it as analogue-digital quantum simulation or a time-dependent Hamiltonian dynamics (although we only require controllable single-qubit operations instead of arbitrary Hamiltonian simulation). Compared to a fully DQS, our scheme does not need to apply any two-qubit gate and hence it significantly avoids crosstalks. It is worth noting that given evolution time T in AQS or pulse sequence of the target gate in DQS, the average number of recovery operations is linear in λT. In practice, the average number of recovery operations could be a small number (for instance, less than 1 in our numerical simulation), and therefore easy to implement in reality. To summarise, our scheme can be implemented on all digital and most analogue quantum simulators, and analogue-digital quantum simulators described above, as long as accurate and fast single-qubit operations are allowed.

10.5 Reduction of model estimation error

For systems with finite-range interactions, local Markovian dynamical process can be reconstructed by using only local measurements, which can be found in Refs. [304, 305] for details. Given a prior knowledge of the noise model, the above stochastic QEM schemes can eliminate the physical errors by applying basis operations at jump time. Nonetheless, the realistic noise $\mathcal{L}_{\text{exp}}$ and the estimated noise $\mathcal{L}_{\text{est}}$ may differ due to imprecise estimation of the noise model. Here we combine the extrapolation QEM method [116, 150] to mitigate model estimation error and the errors associated with imperfect recovery operations.

I first show how to boost model estimation error, which will be used for its mitigation. Assume

that the evolution of the quantum system is described by the open-system master equation

$$\frac{d}{dt}\rho_\lambda = -i[H(t), \rho_\lambda] + \lambda\mathcal{L}_{\text{exp}}\left[\rho_\lambda\right]. \tag{10.10}$$

I show in the Section G.2.1 that by evolving the state ρ'_λ under the re-scaled Hamiltonian $\frac{1}{r}H\left(\frac{t}{r}\right)$ for time rt, we can effectively boost physical errors of quantum systems, which can be expressed as $\rho'_\lambda(rt) = \rho_{r\lambda}(t)$. Here, it is assumed that the noise superoperator $\mathcal{L}$ is invariant under rescaling, and the initial conditions holds $\rho'_\lambda(0) = \rho_{r\lambda}(0)$.

The effective evolution after applying the error mitigation method with $\mathcal{L}_{\text{est}}$ is

$$\frac{d}{dt}\rho_\lambda^{(Q)}(t) = -i[H(t), \rho_\lambda^{(Q)}(t)] + \lambda\Delta\mathcal{L}\left[\rho_\lambda^{(Q)}(t)\right], \tag{10.11}$$

where $\rho_\lambda^{(Q)}(t)$ is the effective density matrix and $\Delta\mathcal{L} = \mathcal{L}_{\text{exp}} - \mathcal{L}_{\text{est}}$. By re-scaling $H(t) \to \frac{1}{r}H(\frac{t}{r})$, the evolution for rescaled time rt is

$$\frac{d}{dt}\rho_{r\lambda}^{(Q)}(t) = -i[H(t), \rho_{r\lambda}^{(Q)}(t)] + r\lambda\Delta\mathcal{L}\left[\rho_{r\lambda}^{(Q)}(t)\right], \tag{10.12}$$

which can be implemented by rerunning the error-mitigated experiment for a re-scaled time rt under the re-scaled Hamiltonian.

As the value of $r \geq 1$ can be tuned, we choose several values of r and suppress the model estimation error via Richardson extrapolation. Specifically, with more than two values of r denoted as $\{r_j\}$ and constants $\beta_j = \prod_{l\neq j} r_l(r_l - r_j)^{-1}$, we have

$$\langle O\rangle_I = \sum_{j=0}^{n} \beta_j \langle O\rangle_{r_j\lambda} + \mathcal{O}\left(\frac{\gamma_n \left(r_{\max}\lambda T \left\|\Delta\mathcal{L}\right\|_1\right)^{n+1}}{(n+1)!}\right), \tag{10.13}$$

where $\langle O\rangle_{r\lambda}$ is the measurement outcome after stochastic error mitigation, corresponding to $\rho_{r\lambda}^{(Q)}(T)$, $\gamma_n = \sum_j |\beta_j|$, $r_{\max} = \max_j r_j$, and $\|\Delta\mathcal{L}\|_1 = \max_\rho \text{Tr}|\mathcal{L}(\rho)|$. Therefore, in addition to $\lambda T = \mathcal{O}(1)$, the scheme is efficient provided

$$r_{\max}\|\Delta\mathcal{L}\|_1 = \mathcal{O}(1). \tag{10.14}$$

The derivation of Eq. (10.13) is shown in Section G.2.2.

From Eq. (10.13), the deviation between the ideal measurement outcome and the error-mitigated one is bounded independently with the to-be-simulated Hamiltonian. The bound only relies on the noise model, the evolution time, the number of samples, and the parameters used in extrapolation. Moreover, since imperfections of the basis operations $\mathcal{B}_i$ lead to deviation of $\mathcal{L}_{\text{est}}$, which can be regarded as another type of model estimation error, they can be corrected via the extrapolation procedure. Note that noise could have fluctuations or drift in the experimental apparatus, which in practice could be challenging to obtain the precise noise model. Our hybrid QEM incorporating extrapolation is therefore practically useful, as this method alleviates the requirement of precise estimation of the noise model and can be robust to the error of recovery operations. Analysis of error mitigation for the model estimation is shown in Section G.2.

10.6 Numerical simulation

As have been shown in the previous section, the variation of the performance of our error mitigation methods in terms of different Hamiltonians and noise models is theoretically well bounded, which indicates that the theory does apply for general Hamiltonian simulation with NISQ devices. In this section, I test the QEM schemes for both analogue quantum simulators and gate-based digital quantum circuits.

10.6.1 Analogue quantum simulation

I first consider a 2D anisotropic Heisenberg model

$$H = J\sum_{\langle ij\rangle}\left[(1+\gamma)\sigma_x^{(i)}\sigma_x^{(j)} + (1-\gamma)\sigma_y^{(i)}\sigma_y^{(j)} + \sigma_z^{(i)}\sigma_z^{(j)}\right] - \gamma h\sum_{i=1}^{4}\sigma_y^{(i)} \tag{10.15}$$

on a 2×2 square lattice, where $\langle ij\rangle$ represents nearest neighbour pairs. This model has been extensively used to investigate the quantum magnetism and criticality [306–310]. I then consider analogue simulation via a noisy superconducting quantum simulator with energy relaxation $\mathcal{L}_1$ and dephasing $\mathcal{L}_2$ noise [191, 280, 311, 312]. Here the Lindblad operator takes the form of

$$\mathcal{L}_\beta[\rho] = \sum_j \lambda_\beta\big(L_\beta^{(j)}\rho L_\beta^{(j)\dagger} - \frac{1}{2}\{L_\beta^{(j)\dagger}L_\beta^{(j)}, \rho\}\big) \tag{10.16}$$

for $\beta = 1, 2$, $L_1^{(j)} = \sigma_-^{(j)} = |0\rangle\langle 1|$, and $L_2^{(j)} = \sigma_z^{(j)}$. Such a noise model is also relevant for other quantum simulators such as trapped ions [9,50–52], NMR [55–57], ultracold atoms [53,54], optical lattices apparatus [58], etc. The noise can be characterised by measuring energy relaxation time T_1 and dephasing time T_2 without full process tomography [191,287,312,313], and more generally via local measurements [304, 305]. I also consider physical errors for the single-qubit recovery operations as single-qubit inhomogeneous Pauli error, $\mathcal{E}_{\text{inh}} = (1-p_x-p_y-p_z)\mathcal{I}+p_x\mathcal{X}+p_y\mathcal{Y}+p_z\mathcal{Z}$ with $\mathcal{I}, \mathcal{X}, \mathcal{Y}, \mathcal{Z}$ being the Pauli channel and p_α being the error probability. In our simulation, we set $J = h = 2\pi\times 4$ MHz, $\gamma = 0.25$, and the noise strength $\lambda_1 = \lambda_2 = 0.04$ MHz [191, 192, 287]. We set the estimated physical errors as $p_x = p_y = 0.25\%$ and $p_z = 0.5\%$, which can be achieved with current superconducting simulators [8, 49], and consider the real noise strength to be 10% greater than the estimated one, i.e., $\lambda_{\text{exp}} = 1.1\lambda_{\text{est}}$. The initial state is set as $|+\rangle^{\otimes 4}$ with $|+\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$. We evolve it to time $T = 16\pi/J$, and measure the expectation value of the normalised nearest-neighbour correlation function $\sum_{\langle ij\rangle}\sigma_x^{(i)}\sigma_x^{(j)}/4$ with 10^6 samples.

The numerical result without model estimation error is shown in Figure 10.3 (a)(b)(c). Specifically, we compare the time evolution of the expectation value of the correlation operator in Figure 10.3 (a)(b) and the fidelity $F(\rho_I, \rho_{\text{eff}}) = \text{Tr}\sqrt{\rho_{\text{eff}}^{1/2}\rho_I\rho_{\text{eff}}^{1/2}}$ of the effective density matrix ρ_{eff} and the ideal one ρ_I in Figure 10.3 (c). We can see that Richardson extrapolation and stochastic

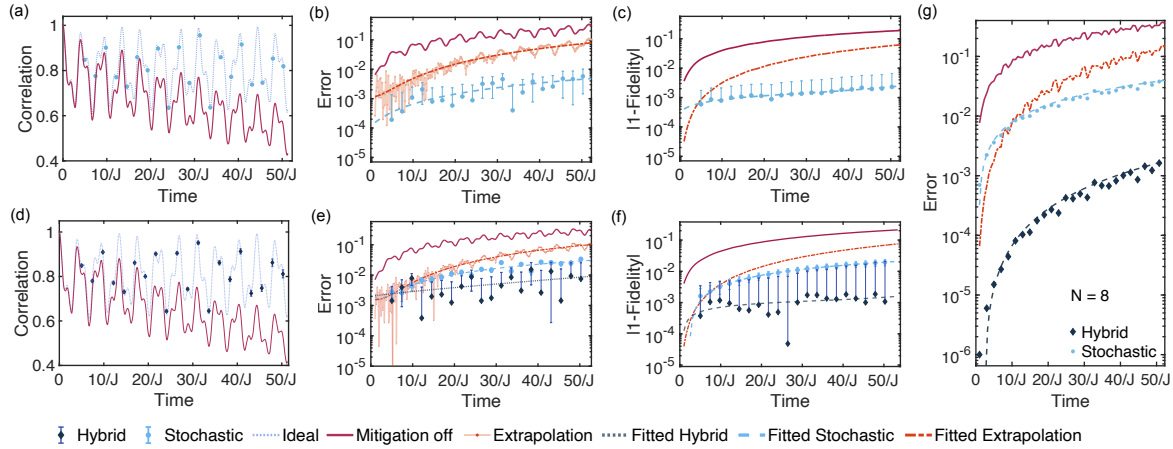


Figure 10.3: Numerical test of the QEM schemes without ((a)(b)(c)) and with 10% model estimation error $\lambda_{\text{exp}} = 1.1\lambda_{\text{est}}$ ((d)(e)(f)(g)). Consider the dynamics of 2D anisotropic Heisenberg Hamiltonian with energy relaxation and dephasing noise. (a)−(f) consider a four-qubit Hamiltonian with finite (10^6) number of samples. (a) and (d) compare the time evolved nearest-neighbour correlation function $\sum_{\langle ij\rangle} \sigma_x^{(i)}\sigma_x^{(j)}/4$. (b) and (e) show the error between the exact value and the error-mitigated value. (c) and (f) show the fidelity of the effective density matrix $\rho_{\text{eff}}^{\alpha}$ and the ideal one ρ_I under different error mitigation scheme α. (g) considers an eight-qubit Hamiltonian with an infinite number of samples. The hybrid error mitigation scheme suppresses the error up to about four orders of magnitude even with 10% model estimation error.

QEM improve the accuracy by one and two orders, respectively. The result with model estimation error is shown in Figure 10.3 (d)(e)(f). Here, we also consider the hybrid method with both stochastic QEM and linear extrapolation, with optimised $r_0 = 1$ and $r_1 = 1.8$. We can see that stochastic QEM still outperforms Richardson extrapolation with large evolution time, and the hybrid method can be further used to improve the simulation accuracy. The simulation result thus indicates that the hybrid QEM scheme can be robust to the drift of noise [314–316]. The performance of the QEM schemes can be made clearer without considering sampling errors. Figure 10.3(g) shows simulations of the eight-qubit anisotropic Heisenberg model on 2×4 lattice under different QEM schemes with infinite samples. The result indicates that both stochastic and hybrid QEM can effectively eliminate the accumulation of errors during the evolution.

10.6.2 Digital quantum simulation

Finally, I consider an eight-qubit, d-depth parameterised quantum circuit (Figure 10.4(c)) and show how stochastic QEM can suppress coherent errors in multi-qubit operations. Here, the controlled-NOT (CNOT) gate in the quantum circuits is generated by cross-resonance (CR) gates, which are experimentally realised by using microwaves to drive the control qubit (c) at the frequency of the target qubit (t), resulting in a driving Hamiltonian $H \approx \Omega(-\sigma_z^{(c)}\sigma_x^{(t)} + \gamma\mathbb{I}^{(c)}\sigma_x^{(t)})$ [192, 193, 271–273]. Here, Ω is the effective qubit-qubit coupling and γ represents the effect of crosstalk between qubits [271]. We consider inherent environmental noise and recovery

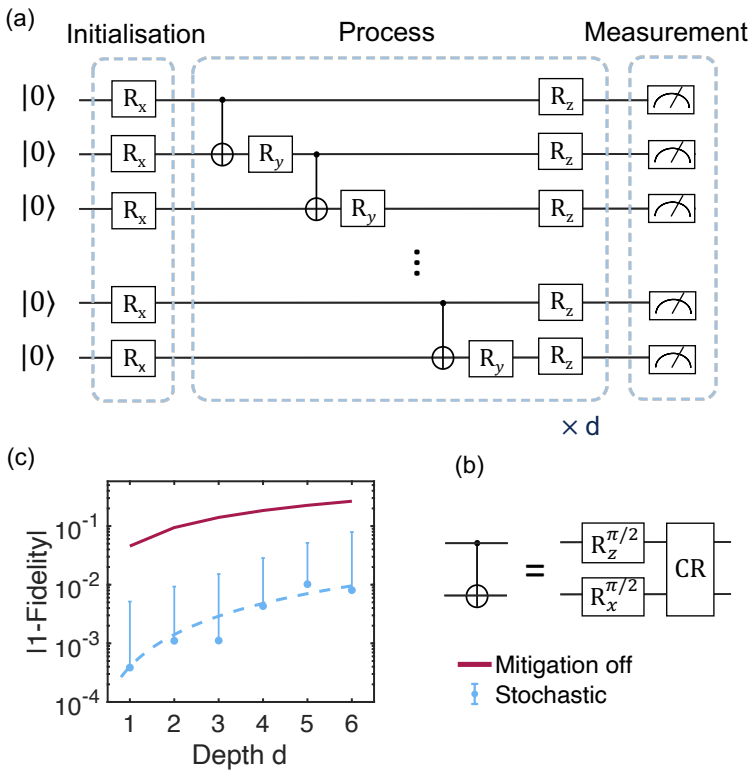


Figure 10.4: Stochastic QEM for eight-qubit superconducting quantum circuits with environmental and crosstalk noise. (a) considers a d-depth parameterised quantum circuits with single-qubit rotations R_α ($\alpha \in \{X, Y, Z\}$ and CNOT gates. The rotation angles are randomly generated from $[0, 2\pi]$. (b) shows the realisation of the CNOT gate via the CR gate $U_{CR} = \exp(i\pi\sigma_z^{(c)}\sigma_x^{(t)}/4)$ and single-qubit gates $R_z^{\pi/2}$ and $R_x^{\pi/2}$ up to a global phase $e^{i\pi/4}$. (c) shows the fidelity dependence of circuit depth d with/without QEM.

operation error as in the above analogue simulator, and additional coherent crosstalk errors $\gamma = 1\%$. We set $\Omega = 2\pi$ MHz, the evolution time $T = \pi/4\Omega$, and run 10^5 samples. We mitigate the noisy two-qubit pulse sequence by inserting basis operations, and shows the fidelity dependence of circuit depth d with/without QEM in Figure 10.4(c). The result clearly shows that stochastic QEM improves the computing accuracy by two orders.

10.7 Resource cost estimation

In this section, I estimate the resource cost for stochastic error mitigation with NISQ devices. Given a precise noise model, the stochastic error mitigation method in principle enables exact recovery of the ideal evolution. However, to achieve the same accuracy of the measurement on the ideal evolution, we need C^2 times more samples or experiment runs with the error-mitigated noisy evolution. The overhead C^2 is likely to be prohibitively large due to large noise on a NISQ device. Nevertheless, I show that the overhead can be reasonably small (less than 100) when

the total error (defined below) is less than 1. In particular, we consider a noisy superconducting simulator with up to $N = 100$ qubits, which suffers from single-qubit relaxation and dephasing noise with equal noise strength $\lambda_1 = \lambda_2$. While the noise strength is defined as the noise rate at instant time, we define the total noise strength

$$\Lambda = NT(\lambda_1 + \lambda_2) \tag{10.17}$$

as the noise strength of the whole N-qubit system within time T. The dependence of the overhead C^2 on the total noise strength Λ and number of qubits is shown in Figure 10.5(a). For a practical case with $T = 1$ μs, $\lambda_1 + \lambda_2 = 0.01$ MHz, $N = 100$, the cost $C^2(\Lambda = 1) = 30$ and I further show the number of measurements needed to achieve a given simulation accuracy in Figure 10.5(b). Note that the overhead C^2 is independent of the Hamiltonian H_{sim}, so the results apply for general NISQ devices (see Section G.1.1).

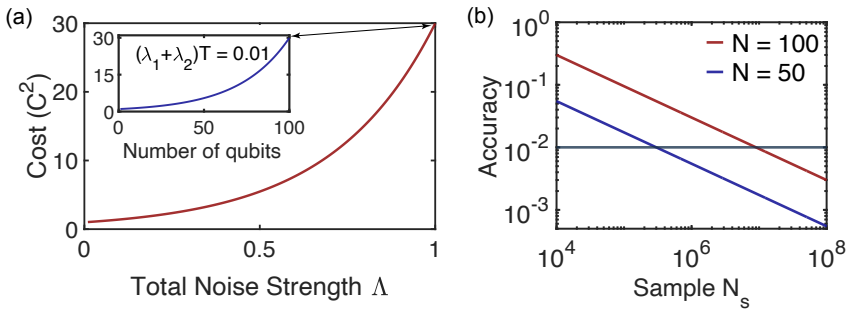


Figure 10.5: (a) Cost versus total noise strength $\Lambda = NT(\lambda_1 + \lambda_2)$. Consider a general N-qubit Hamiltonian H_{sim} with single-qubit energy relaxation (λ_1) and dephasing noise ($\lambda_2 = \lambda_1$), and evolution time T. The inset shows the cost versus a different number of qubits with $(\lambda_1 + \lambda_2)T = 0.01$. (b) Simulation accuracy $\varepsilon \propto C/\sqrt{N_s}$ with different number of samples N_s with $T = 1$ μs, $\lambda_1 + \lambda_2 = 0.01$ MHz, $N = 100$ (red) and $N = 50$ (blue). We consider a pessimistic estimation of the overhead, $C/\sqrt{N_s}$, and the error ε can be much smaller in practice.

10.8 Statistical fluctuation in measurements

In the above discussion, I discuss how to mitigate the errors in the quantum simulation process due to undesired interactions with the environment. Another primary error comes from the statistical fluctuation due to finite samples. This kind of shot noise could be reduced by exploiting more advanced measurement schemes. The basic idea is to either exploit observable compatibility, importance sampling, or additional quantum circuit to more efficiently measure observables. In this section, I propose a unified framework for quantum state measurement with no increase in the circuit depth, and show explicit measurement schemes to estimate observables with fewer measurements.

10.8.1 A unified framework for quantum state measurement

Suppose that the target observable can be decomposed into the Pauli basis $\mathbf{O} = \sum_l \alpha_l \mathbf{O}_l$ with $\mathbf{O}_l \in \{I, X, Y, Z\}^{\otimes n}$. Here, the bold format is used to represent the n-qubit Pauli operators $\mathbf{O}_l$ and the subscript l of $\mathbf{O}_l$ to represent the lth n-qubit Pauli operators in the decomposition. Without loss of generality, denote an n-qubit Pauli operator as $\mathbf{Q} = \otimes_{i=1}^n Q_i$ with $Q_i \in \{I, X, Y, Z\}$ being the single-qubit Pauli operator that acts on the ith qubit. The expectation value of $\mathbf{Q}$ can be obtained by measurements in any Pauli basis $\mathbf{P} = \otimes_{i=1}^n P_i$ whenever $Q_i = P_i$ or $Q_i = I$ for any i, which we refer to as $\mathbf{P}$ hits $\mathbf{Q}$ and denote by $\mathbf{Q} \triangleright \mathbf{P}$. In this case, $\mathbf{Q}$ element-wisely commutes with $\mathbf{P}$.

Provided the target observable and the measurement scheme, we first determine a measurement basis set $\{\mathbf{P}\}$ and the corresponding probability distribution $\{\mathcal{K}\}$, and then generate an estimation of $\mathrm{Tr}(\rho\mathbf{O})$ by measuring ρ with $\mathbf{P}$ selected from the basis set over the distribution $\mathcal{K}(\mathbf{P})$. The estimator for the observable $\mathbf{O}$ with measurement $\mathbf{P}$ is given by

$$\hat{\mathbf{o}}(\mathbf{P}) = \sum_l \alpha_l f(\mathbf{P}, \mathbf{O}_l, \mathcal{K}) \mu(\mathbf{P}, \mathrm{supp}(\mathbf{O}_l)), \tag{10.18}$$

where $\mu(\mathbf{P}, \mathrm{supp}(\mathbf{O}_l)) := \prod_{i \in \mathrm{supp}} \mu(P_i)$ with $\mu(P_i)$ being the single-shot outcome by measuring the ith qubit of state ρ with the Pauli basis P_i, and the support of $\mathbf{Q}$ defined by $\mathrm{supp}(\mathbf{Q}) := \{i | Q_i \neq I\}$. Below, I will show the explicit forms of the probability distribution $\mathcal{K}(\mathbf{P})$ and function f for importance sampling, grouping and classical shadow algorithms, which give an unbiased estimation

$$\mathbb{E}[\hat{\mathbf{o}}] = \mathrm{Tr}(\mathbf{O}\rho). \tag{10.19}$$

The variance of the estimator in Eq. (10.18) for a single sample could be calculated by the definition as

$$\mathrm{Var}[\hat{\mathbf{o}}] = \mathbb{E}_{\mathbf{P}} \sum_{l,l'} \alpha_l \alpha_{l'} \mathrm{Tr}(\rho \mathbf{O}_l \mathbf{O}_{l'}) f(\mathbf{P}, \mathbf{O}_l, \mathcal{K}) f(\mathbf{P}, \mathbf{O}_{l'}, \mathcal{K}) - \mathrm{Tr}(\rho\mathbf{O})^2, \tag{10.20}$$

where the following equality is used,

$$\mathbb{E}_{\mu(\mathbf{P})} \mu(\mathbf{P}, \mathrm{supp}(\mathbf{O}_l)) \mu(\mathbf{P}, \mathrm{supp}(\mathbf{O}_{l'})) = \mathbb{E}_{\mu(\mathbf{P})} \mu(\mathbf{P}, \mathrm{supp}\,.(\mathbf{O}_l \mathbf{O}_{l'})) = \mathrm{Tr}(\rho \mathbf{O}_l \mathbf{O}_{l'})$$

The detailed proof can be found in Refs. [87, 200]. In the following, I will discuss the relations of these representative measurement algorithms within this framework.

10.8.2 Importance sampling

Importance sampling is also referred to as l_1 sampling. The measurement $\{\mathbf{P}_l\}$ is chosen to be the observable $\{\mathbf{O}_l\}$, and the corresponding probability is determined by the weight of the observable

as $\mathcal{K}(\mathbf{P}_l) = |\alpha_l| / \|\boldsymbol{\alpha}\|_1$. Here, $\|\boldsymbol{\alpha}\|_1$ is the l_1 norm of $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_L)$ as $\|\boldsymbol{\alpha}\|_1 = \sum_{l=1}^{L} |\alpha_l|$. The function f is defined by

$$f_{l_1}(\mathbf{P}, \mathbf{O}_l, \mathcal{K}) = \mathcal{K}(\mathbf{P})^{-1} \delta_{\mathbf{P},\mathbf{O}_l}. \tag{10.21}$$

From Eq. (10.20), the variance of the estimator obtained using importance sampling could be calculated by

$$\mathrm{Var}[\hat{\mathbf{o}}] = \|\boldsymbol{\alpha}\|^2 - \mathrm{Tr}(\rho \mathbf{O})^2, \tag{10.22}$$

which is directly related to the l_1 norm of the coefficients.

10.8.3 Grouping

The key idea of the grouping measurement strategies is that we first allocate observables $\mathbf{O}_l$ to several non-overlapping sets, which satisfies that any two observables $\mathbf{O}_l$ and $\mathbf{O}_{l'}$ in each set are compatible with each other, i.e., $\mathbf{O}_l \triangleright \mathbf{O}_{l'}$ or $\mathbf{O}_{l'} \triangleright \mathbf{O}_l$. Note that when the Pauli observables are compatible with each other, their expectation values can be simultaneously obtained by measuring in one basis. While finding the optimal measurement basis sets for the observables is NP-hard, several heuristic methods for finding a measurement basis have been proposed that run in a polynomial time [87, 198]. Here, I focus on the largest degree first (LDF) grouping method, whereas other grouping methods can be analysed in a similar way. We can divide $\mathcal{O} = \{\mathbf{O}_l\}$ into several groups $\mathcal{S}_j$ such that $\cup_j \mathcal{S}_j = \mathcal{O}$, $\mathcal{S}_j \cap \mathcal{S}_{j'} = \varnothing, \forall j \neq j'$. For each group $\mathcal{S}_j$, a measurement $\mathbf{P}_j$ is assigned such that we can measure any observable $\mathbf{Q}$ in the jth set $\mathcal{S}_j$ with measurement $\mathbf{P}_j$, i.e., $\mathbf{Q} \triangleright \mathbf{P}_j, \forall \mathbf{Q} \in \mathcal{S}_j$. The probability $\mathcal{K}(\mathbf{P}_j)$ can be chosen either uniformly or based on the total weight of the observables in the jth set, i.e. $\mathcal{K}(\mathbf{P}_j) = \|\mathcal{S}_j\|_1 / (\sum_j \|\mathcal{S}_j\|_1)$. The function f of the grouping method integrated with the importance sampling is chosen by

$$f_{\mathrm{group}}(\mathbf{P}_j, \mathbf{Q}, \mathcal{K}) = \mathcal{K}(\mathbf{P}_j)^{-1} \delta_{\mathbf{Q} \in \mathcal{S}_j}. \tag{10.23}$$

From Eq. (10.20), the variance of the estimator obtained using the grouping method could be calculated by

$$\mathrm{Var}[\hat{\mathbf{o}}] = \sum_j \mathcal{K}(\mathbf{P}_j)^{-1} \sum_{l,l' : \mathbf{O}_l, \mathbf{O}_{l'} \in \mathcal{S}_j} \alpha_l \alpha_{l'} \mathrm{Tr}(\rho \mathbf{O}_l \mathbf{O}_{l'}) - \mathrm{Tr}(\rho \mathbf{O})^2. \tag{10.24}$$

This uses the definition of f_{group}, which is nonzero only if $\mathbf{O}_l \in \mathcal{S}_j$.

10.8.4 Classical shadow

The basic idea of the classical shadow (CS) scheme [199, 200] is that we first perform randomised measurements on each qubit and then post process these classical outcomes to estimate the target observables. The probability distribution $\mathcal{K}_i(P_i)$ that performs Pauli measurement P_i on ith qubit is independent on each site, and therefore the probability distribution for one measurement $\mathbf{P}$ is a product of the distribution on each site $\mathcal{K}(\mathbf{P}) = \prod_i \mathcal{K}_i(P_i)$. The uniform CS method consider a

uniform distribution over the Pauli basis as $\mathcal{K}_i(P_i) = 1/3$, irrespective of the target observables. In Ref. [200], the authors proposed that the local probability distribution $\mathcal{K}_i$ could be optimised to reduce the number of samples, termed as locally biased classical shadow method. The function is defined by

$$f_{\mathrm{CS}}(\mathbf{P}, \mathbf{Q}, \mathcal{K}) = \prod_i f_i(P_i, Q_i, \mathcal{K}_i) \tag{10.25}$$

with $f_i(P_i, Q_i, \mathcal{K}_i) = \delta_{Q_i, I_2} + \mathcal{K}_i(P_i)^{-1}\delta_{Q_i, P_i}$. Note that the variance for the CS method can be bounded by

$$\mathrm{Var}[\hat{\mathbf{o}}] \leq \sum_{l,l'} \alpha_l \alpha_{l'} f_{\mathrm{CS}}\left(\mathbf{O}_l, \mathbf{O}_{l'}, \mathcal{K}\right) \mathrm{Tr}\left(\rho \mathbf{O}_l \mathbf{O}_{l'}\right) \leq 3^{\mathrm{supp}(\mathbf{O})} \left(\sum_{l=1}^{L} \alpha_l\right)^2 \tag{10.26}$$

with $\mathrm{supp}(\mathbf{O}) := \max_l \mathrm{supp}(\mathbf{O}_l)$. From Eq. (10.26), the variance for the uniform CS method scales exponentially with the support of the target observable, and the uniform CS method hence could be inefficient for the estimation of nonlocal observables with large support.

Huang *et al.* further proposed a derandomised CS algorithm, in which the measurement basis $\mathbf{P}$ is deterministically selected [317]. The derandomisation algorithm first assigns a collection of N_s completely random n-qubit Pauli measurements, and iteratively identifies the measurement basis that minimises the conditional expectation value over all remaining random measurement assignments. As argued in the Ref. [317], the derandomised measurement procedure has no assurance to be globally optimal, or close to optimal. Nevertheless, it shows significant performance for realistic molecular Hamiltonians, ranging from 8 to 16 qubits in the author's paper [87].

It is worth noting that once the measurement bases are determined, it could be regarded as a special grouping method. The difference is that the observables are permitted to be assigned in different groups, as opposed to conventional grouping methods introduced above. We can find its underlying relations to the grouping method within the measurement framework described by Eq. (10.18).

10.8.5 Comments on quantum state measurements

Several other relevant works that do not introduce entangling gates for measurements were posted (see [87, 318, 319]). Hadfield *et al.* [318] proposed an adaptive Pauli shadow algorithm, and Hillmich *et al.* [319] proposed a decision diagrams method to generate an estimation. The author proposed the overlapped grouping method that exploits the spirit of Pauli grouping and classical shadows. These measurement schemes improve the performance of energy estimation by optimising measurement bases and probability distributions. The overlapped grouping method shows significant measurement cost reductions over the works that are discussed in this chapter, supported by numerical simulation [87].

It is worth noting that these methods can be described within the framework introduced in Section 10.8.1. One can compare the performance of different methods by the variance as

$$\mathrm{Var}[\hat{\mathbf{o}}] = \sum_{j,k} \alpha_l \alpha_{l'} g(\mathbf{O}_l, \mathbf{O}_{l'}) \mathrm{Tr}(\rho \mathbf{O}_l \mathbf{O}_{l'}) - \mathrm{Tr}(\rho \mathbf{O})^2 \tag{10.27}$$

where $g(\mathbf{O}_l, \mathbf{O}_{l'}) = \left(\sum_{\mathbf{P}:\mathbf{O}_l \triangleright \mathbf{P}} \mathcal{K}(\mathbf{P})\right)^{-1} \left(\sum_{\mathbf{P}:\mathbf{O}_{l'} \triangleright \mathbf{P}} \mathcal{K}(\mathbf{P})\right)^{-1} \sum_{\mathbf{P}:\mathbf{O}_l \triangleright \mathbf{P} \wedge \mathbf{O}_{l'} \triangleright \mathbf{P}} \mathcal{K}(\mathbf{P})$. Here, we use the definition in Eq. (10.18), and the denominator in Eq. (10.27) indeed represents the probability that the observable $\mathbf{O}_l$ ($\mathbf{O}_{l'}$) is effectively measured with the measurement basis $\mathbf{P}$ (See Lemma 1 in Ref. [87]). One can check that it reduces the conventional grouping method in Eq. (10.24) if we restrict that each observable can only be assigned into one group. Moreover, if the measurement bases are deterministically selected from the probability distribution and then fixed, it has the same spirit as that in derandomisation. One can similarly compare the performance of these methods using the experimental data and the corresponding post-processing method. I refer to [98] for the experimental demonstration of different quantum state measurement strategies.

10.9 Discussion

In this chapter, I first introduce stochastic and hybrid quantum error mitigation schemes to mitigate noise in a continuous process. While previous error mitigation methods designed for DQS regard noise as an error channel before or after the gate, such a description becomes inadequate with multi-qubit quantum gates, and when the noise is inherently mixed in the realisation of the quantum gate. By regarding the implementation of each quantum gate as a continuous process, our error mitigation methods can thus be applied to mitigate errors for realisation of multi-qubit gates (which generally have large errors). Since the dominant noise in NISQ devices is from implementing multi-qubit operations or inherent noise with finite coherence time, our scheme can effectively suppress these noise sources, and thus extend the computation capability of analogue quantum simulators and digital gate-based quantum computers in solving practical problems [8]. Our QEM methods are numerically tested for mitigating errors occurring in analogue simulators for several Hamiltonian simulations under incoherent errors (energy relaxing and dephasing errors) and a parameterised quantum circuit under both incoherent errors and coherent crosstalk errors. In addition, the requirements for mitigating errors in general NISQ devices with up to 100 qubits are discussed. The proposed QEM schemes work for all digital and many analogue quantum simulators with accurate single-qubit controls.

Furthermore, resolving the drift or temporal fluctuations of noise is challenging for conventional QEM methods. Our hybrid scheme incorporating extrapolation can mitigate model estimation error and the error of recovery operations, which alleviates the requirement of precise tomography of the error model and precise control of the quantum simulators. Our method is tested to be robust to the drift of noise. Although the discussion focused on local time-independent

noise, our scheme can be potentially generalised to general non-local time-dependent noise by employing the time dependent recovery operation $\mathcal{E}_Q(t)$.

Efficient quantum measurement is crucial for many quantum algorithms and quantum processing. To alleviate statistical errors in quantum measurements, I propose efficient quantum state measurement scheme that significantly reduces the measurement cost and speeds up the quantum computation, especially when we aim to realise quantum advantage for realistic problems. I start by introducing a unified framework for quantum state measurements. The underlying mechanism of the advanced measurement strategies, including importance sampling, grouping, and classical shadows, which are seemingly distinct from each other, is revealed. The numerical and experimental results in the author's papers [87, 98] suggest a significant measurement cost reduction when applying more advanced measurement methods.

Author contributions. The error mitigation part is relevant to a published work [97]. S.E., X.Y., and J.S. initiated the project. S.E., V.V., and X.Y. supervised the project. S.E. and J.S. conceived the idea of stochastic error mitigation. S.E. conceived the idea of hybrid error mitigation. J.S. carried out numerical verification of QEM with input from X.Y., S.E., and V.V.. All the authors contributed to the discussion and writing the manuscript.

The state measurement part is relevant to published works [87, 98]. J.S. was responsible for the theoretical aspects of the work [98]. X.Y. and H.L. supervised the project. X.Y. and J.S. proposed the experiment. T.Z., X.F., and H.L. performed the experiments. J.S. and T.Z. performed analysis of the experimental data. J.S. proposed the framework for quantum state measurements in [87]. J.S. and X.Y. initiated the project. X.Y. supervised the project. B.W. developed OGM with input from J.S., and carried out numerical simulation with input from Q.H..

Chapter 11

Conclusion and outlook

Interesting phenomena usually emerge from nontrivial quantum correlations, which make it challenging to compute a quantum system and predict its behaviour using classical computing. Quantum computing holds the promise of solving quantum many-body problems and assisting in the discovery of new and interesting phenomena. The focus of this thesis has been on quantum many-body problems along with quantum computing and spectroscopy. A unified and useful framework is established for these topics.

I first give a summary of the contents of this thesis. Part I provided a background introduction to quantum computing, quantum simulation, and quantum many-body problems, as well as an overview of the key ideas at the heart of the methods developed in this thesis. The focus of Part II was on two general classes of quantum many-body problems: static and dynamic problems. I proposed a series of quantum computing approaches to address these two types of problems. In Part III, I expanded the scope of these methods to explore interesting quantum phenomena of many-body systems, such as interacting bosons, fermions, and quantum spin systems. In addition, I investigated the out-of-equilibrium dynamics and spectroscopic properties of model quantum systems. In Part IV, I developed a framework for estimating the quantum resources requirements of quantum algorithms, involving both eigenstate property estimation and Hamiltonian simulation. This forms a good application test bed for the methods developed in this thesis. With this framework, I estimated the resources required for typical applications in respect of both near-term noisy quantum processors and error-corrected quantum computers in the spin systems, condensed-phase electrons, and molecular systems. In Chapter 10, I addressed the two challenges when it comes to the practical implementation of quantum algorithms with noisy quantum devices. I demonstrated how errors can be suppressed in practical applications. More specifically, I showed how to mitigate errors of quantum processes due to inevitable interactions with the environment and how to suppress statistical errors in quantum state measurements with an efficient grouping measurement scheme.

It is arguably constructive to view spectroscopy, analogue quantum simulation, and programmable digital quantum simulation as techniques that can complement the power of quan-

tum computing of quantum many-body systems. From a theoretical perspective, spectroscopic methods provide valuable insights into quantum computing. Drawing on the spectroscopic methods introduced in Chapter 7, I demonstrated how to probe spectroscopic features of quantum many-body systems, and obtain the interacting quasiparticle spectra using the out-of-equilibrium dynamics information. In Chapter 8, I applied the spectroscopy and perturbative approaches developed in this thesis to explore excitation spectra and dynamical properties of several typical many-body systems.

The best scheme for a given physical problem depends on the nature of the task itself as well as what we expect to learn from the results. Quantum computing not only plays a role of computing, but more importantly, is able to uncover the structure of quantum problems. It essentially tells us whether the problems are quantum and what the role of quantum correlation plays in its behaviour. Within the realm of physics, the ultimate success of quantum computing is built upon a comprehensive, deep understanding of the quantum aspects of quantum problems. With the ongoing investigation in quantum systems, I believe that the development of quantum computing methods will ultimately advance the understanding of the underlying relation between different descriptions of correlation effects in many-body systems at a fundamental level, and vice versa.

To conclude, I will briefly touch on some potentially promising future research directions generated by this thesis.

Understanding quantum correlations. Quantum many-body systems normally incorporate both weak and strong correlations at different levels of segments of the system [320]. A typical method of achieving this is through partitioning the quantum system into a fragment and a bath, and treating each part using different levels of theories, also known as quantum embedding theories [210, 211]. The effective Hamiltonian of the original system can be obtained through the projection of the original many-body system onto the reduced fragment-bath quantum state basis. However, embedding theories rely on the knowledge of the target materials and in general they are an inadequate way to represent strongly correlated systems. It is desirable to develop a framework which is able to represent the quantum correlations faithfully. The hybrid tensor network framework that has been established in Chapter 3 could accommodate such a procedure in a natural way. In contrast to the suppression of quantum fluctuations through imposing classical mean-field descriptions in conventional quantum embedding theories, the hybrid tensor network approach could better account for quantum fluctuation effects beyond the mean-field approximation. Moreover, this concatenation of tensor network representation, either using classical and quantum tensors, could significantly save quantum resources in terms of the qubit number and circuit depth.

We can further compare different embedding approaches such as density matrix embedding theory (DMET) [216–218, 321–326], dynamical mean-field theory [212, 214, 215, 327–329],

DFT [163, 330, 331], and other classical methods such as tensor networks and perturbation approaches, which differ in their approximation for correlations between the fragment and the bath. The outcomes could advance the understanding of underlying relations between different descriptions of the correlation effects in quantum many-body systems.

Quantum circuit expressivity and quantum solver. At a fundamental level, the quantum advantage flows from the assumption that quantum circuits may be sufficiently powerful to express quantum states. Therefore, the power of shallow (noisy) quantum circuits, i.e., expressivity, is an important criterion for realising quantum advantages [332]. Previous methods, in which the circuit expressivity is defined as the distance to random unitaries [333] or the covering number of the hypothesis space of quantum circuits [334], might be inappropriate for understanding the essential power of quantum circuits. Here, we can consider entanglement as a measure for the circuit expressivity [335, 336]. Advanced quantum information and complexity theories will be used to define the measure for noisy circuit expressivity.

With the established measure, we can study varied circuit structures, such as coupled-cluster ansatz [119, 337–341], commonly used in quantum chemistry, or Hamiltonian ansatz for the Hubbard model. These results provide insights to understand the superiority of quantum circuits in practical computation. For solving the eigenstates of the effective Hamiltonian, I conceive of applying advanced quantum algorithms such as variational algorithms, adaptive Hamiltonian simulation to further improve the efficiency and accuracy of the quantum-solver. Alternative approaches that have been developed and tested in this thesis, such as quantum algorithmic cooling and quantum subspace expansion, could supplement the power of the quantum-solver.

Implementation on quantum devices. Several critical challenges still need to be addressed when it comes to the practical implementation on near-term quantum devices. For that, one can apply the error mitigation protocols developed in Chapter 10 to ensure simulation accuracy, and employ the classical shadow approach [87, 199, 200] to reduce the measurement complexity during the measurement and post-selection processes in order to preserve quantum advantage along the whole quantum computing pipeline. Task-specific circuit depth reduction methods could additionally be applied along these lines to improve the feasibility of practical implementation. For instance, the quantum resources could be reduced by considering the symmetry of the target problems [342–344]. With these advanced methods, together with my collaborators, we have demonstrated the feasibility of performing quantum simulation of complex quantum systems and we have achieved the largest scale quantum simulation of chemical systems [321] and periodic solids [345] to date. In these works, we numerically simulated this quantum process (to be more precise, we simulated variational quantum eigensolvers) by a state-vector form, and the simulation of a large-scale problem is enabled by density matrix embedding theories, which greatly reduces the quantum resources required.

Applicable domains of this quantum computing pipeline are broad, potentially encompassing chemistry, condensed matter physics, quantum field theories, and general many-body problems. These include solving electronic structures and predicting transport properties of functional materials, such as high-temperature superconductivity and topological materials, which face an inherent computational complexity that must be overcome. The computational tasks which currently appear intractable could potentially be addressed by my compendium of quantum protocols. The outcome of my proposed projects enables us to scale up quantum simulations of materials, and potentially demonstrates a clear quantum advantage in quantum many-body problems.

Appendices

Appendix A

Entanglement and correlation in hybrid tensor networks

In this section, I discuss entanglement and correlation properties of hybrid tree tensor networks with two layers. I consider three cases with local or non-local correlations, each being represented by either a classical or quantum tensor. The discussion can be generalised to trees with multiple layers.

A.1 Local quantum correlation and non-local classical correlation

I will first discuss hybrid tensor networks with local quantum correlation and non-local classical correlation, which were introduced in Section 3.2.3. To expand on the terminology used here, in such a hybrid tensor network, quantum tensors are used to represent quantum correlations of local subsystems, and classical tensors are used to represent correlations between subsystems. In this context, 'local' concerns correlations in each subsystem, whilst 'non-local' relates to correlations between different subsystems. By way of illustration, let us take the classical tensor as MPS. As shown in Figure A.1(a), we separate the whole kn qubits into two subsystems with a blue boundary line, and denote the left and the right parts to be A and $\bar{A}$, respectively. Without loss of generality, we choose two kinds of boundaries — one is in the bulk of the classical tensor at the first layer, where the boundary is between indices i_l and i_{l+1} (in the same time between j_n of $\psi_{\vec{j}}^{i_l}$ and j_1 from $\psi_{\vec{j}}^{i_{l+1}}$); the other is in the bulk of the quantum tensor at the second layer, where the boundary is between indices $j_{l'}$ and $j_{l'+1}$.

For the first type of bipartition between i_l and i_{l+1}, the entanglement of a subsystem is upper bounded by the bond dimension of the index s_l of the MPS, which is normally a constant number that is independent of subsystem size. As a result, the correlation between these n-qubit clusters is weak. The second type of bipartition between $j_{l'}$ and $j_{l'+1}$ is inside the quantum tensor $\psi_{\vec{j}}^{i_1}$, which is represented by a general quantum state. The entanglement entropy can in principle be proportional to $\min\{l', n-l'\}$, that is, proportional to the subsystem size of the quantum tensor.

The entanglement can also be revealed by the correlation functions. If we select two local observables O_1 and O_2 inside the bulk of any local quantum tensors, the correlation function $\langle O_1 O_2 \rangle$ could even be a constant. However, if they are located in different local quantum tensors, the correlation suffers exponential decay $\langle O_1 O_2 \rangle \sim \exp(-a|l_2 - l_1|)$, where a is some constant depending on the chosen MPS, and $|l_2 - l_1|$ labels the distance between the two quantum tensors where O_1 and O_2 are inside. As shown in Figure A.2, this scaling can be obtained by first contracting the local observables O_1 and O_2 with the quantum tensors, and then the total result becomes two new observables, say M_1 and M_2 in the MPS tensor network.

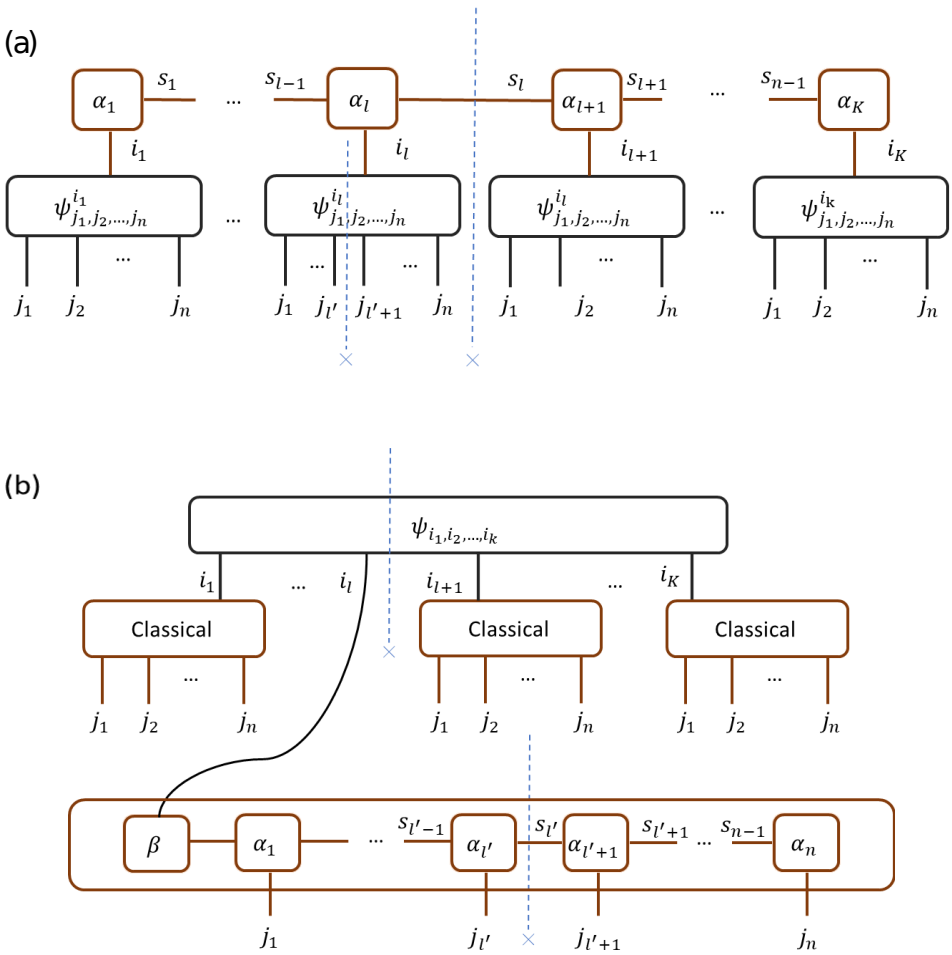


Figure A.1: Illustration of hybrid tensor networks with bipartition. (a) quantum-classical tensor stricture (ansatz) with the classical tensor being MPS introduced in Section 3.2.3; (b) classical-quantum tensor stricture introduced in Section 3.2.4. Here we separate the whole Kn-qubit system into two parts, that is, the left and the right subsystems, with the boundary being denoted by the blue dotted-line. Without loss of generality, we choose two kinds of boundaries: one is in the bulk of the (classical or quantum) tensor at the first level, where the boundary is between indices i_l and i_{l+1}; the other is in the bulk of the tensor at the second level, where the boundary is between indices $j_{l'}$ and $j_{l'+1}$.

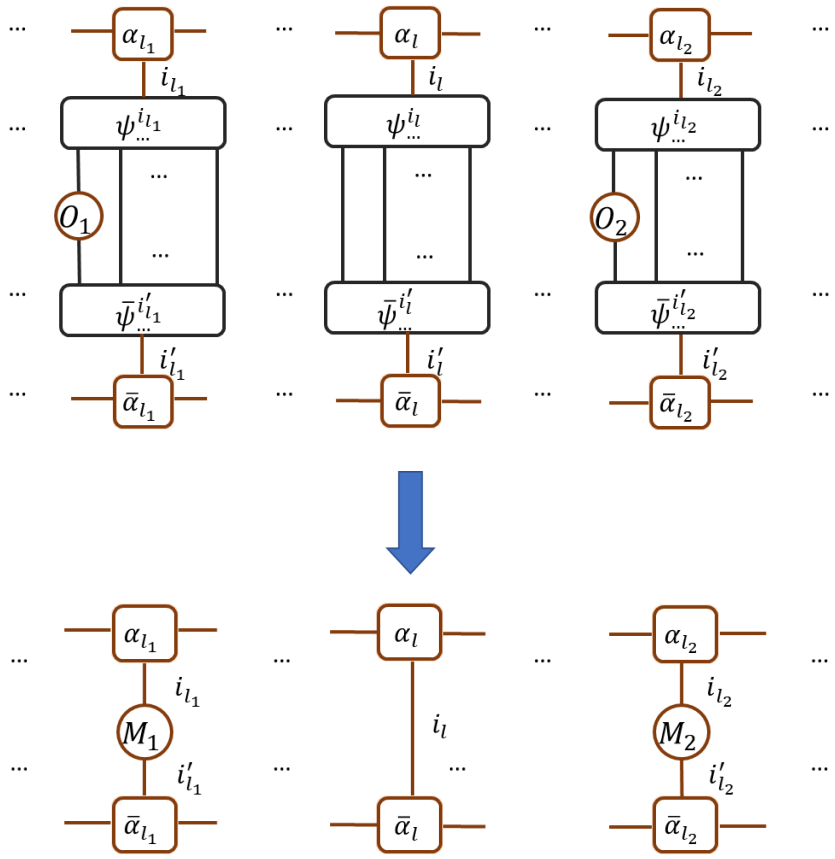


Figure A.2: The contraction process to extract the correlation function of the quantum-classical hybrid tensor network introduced in 3.2.3. Here, the classical tensor is an MPS. The two observables O_1 and O_2 are in two local quantum tensors $\psi^{i_{l_1}}_{\cdots}$ and $\psi^{i_{l_2}}_{\cdots}$, respectively, with i_{l_1} and i_{l_2} being two classical indices. Similar to the extraction of expectation values in Figure 3.2 (a), we first contract O_1 and O_2 with the local quantum tensors to get two new observables M_1 and M_2 for the classical MPS. Here we assume that $\langle \psi^{i'_l}_{\cdots} | \psi^{i_l}_{\cdots} \rangle = \delta_{\{i_l i'_l\}}$ for simplicity, that is, the quantum states indexed by i_l are orthogonal to each other. In this way, the contraction result is the identity for other quantum tensors which do not contain O_1 or O_2. As a result, the correlation function shows $\langle O_1 O_2 \rangle = \langle M_1 M_2 \rangle_{\{MPS\}}$, where the second expectation value is on the MPS and shows an exponential decay.

A.2 Local classical correlation and non-local quantum correlation

Next, let us examine the hybrid tensor network introduced in Section 3.2.4, which uses classical tensors to represent local correlations in the subsystems and uses a quantum tensor to represent non-local quantum correlation between the subsystems. As shown in Figure A.1 (b), the whole kn-qubit is separated into two parts by the blue boundary line. We also choose two kinds of boundaries — one is in the bulk of the quantum tensor between indices i_l and i_{l+1}; the other is in the bulk of the classical tensor between indices $j_{l'}$ and $j_{l'+1}$.

For the bipartition between i_l and i_{l+1} inside the global quantum tensor $\psi_{i_1,i_2,\cdots i_k}$, the en-

tanglement can be proportional to $\min\{l, k-l\}$, with the subsystem size being $|A| = nl$ and $|\bar{A}| = n(k-l)$. As a result, when $k \gg 1$ and $n = \mathcal{O}(1)$, the hybrid tensor network could have a volume law of entanglement, due to the contribution from the quantum tensor. In the regime $n \sim k$, the entanglement is weaker but still stronger than the area law of entanglement. From this point of view, the hybrid tree tensor network could represent more complicated entanglement than pure classical tensor networks. For the second kind of bipartition between $j_{l'}$ and $j_{l'+1}$, the entanglement of the subsystem is upper bounded by the bond dimension of the index $s_{l'}$ of the local MPS, which is normally a constant number independent of the subsystem dimension. Consequently, the correlations inside these n-qubit clusters are weak.

If two local observables O_1 and O_2 are inside the bulk of any local classical tensor, the correlation function $\langle O_1 O_2 \rangle$ decays exponentially. However, if they are located in different local classical tensors, the correlation could be some constant or decay polynomially. One can first contract the local observables O_1 and O_2 with the classical tensors, and then the total result becomes two new observables for the general quantum tensor.

A.3 Local quantum correlation and non-local quantum correlation

Let us now turn to the hybrid tensor network introduced in Section 3.2.5, in which both local and non-local correlations are represented by quantum tensors. It could possess a strong correlation both for local and global correlation. The entanglement entropy and correlation function can be analysed in the same way as detailed above.

For the entanglement of a general hybrid tree structure, one just needs to check if the boundary is in the bulk of a quantum tensor or a classical tensor. The decaying behaviour of the correlation function can be obtained by iterative contractions and identifying whether the final new observables are in a quantum or classical tensor.

A.4 Comments on resources

The classical MPS tensors in the hybrid TTN may be replaced with various classical tensors, such as, a g-qubit classical tree tensor, Multi-scale Entanglement Renormalisation Ansatz (MERA) [346], or Projected Entangled Pair State (PEPS) [347]. For example, one can substitute the MPS block for a (g', D') tree classical tensor [348]. By contracting the local observable from the deepest layer, one layer at a time, the classical cost will be $\frac{g'^{D'}-1}{g'-1}(g'+1)\kappa^{(g'+2)}$. In this instance, this tree tensor is a vertex in the whole hybrid TTN, and consequently, we have $g = g'^{D'}$ and the cost for contraction is $\mathcal{O}(g\kappa^{(g'+2)})$. The classical costs for different types of classical tensors are depicted in Table 3.1. Considering the MERA, the classical contraction cost is $\mathcal{O}(\log(g)\kappa^9)$ for the 1D MERA with ternary indices [349]. The cost of MERA for 2D systems scales as κ^{16}

in some cases [350]. For the generalisation of MPS to the 2D case, PEPS, there is no efficient way of exact tensor contraction. There is approximate algorithm running in $\mathcal{O}(g\kappa^8\tilde{\kappa}^2)$ based on the contraction of 1D MPS inside the 2D PEPS, where $\tilde{\kappa}$ is the bond dimension used to approximately truncate the original tensor. The comparison is shown in Table 3.1.

A.5 Additional numerical tests

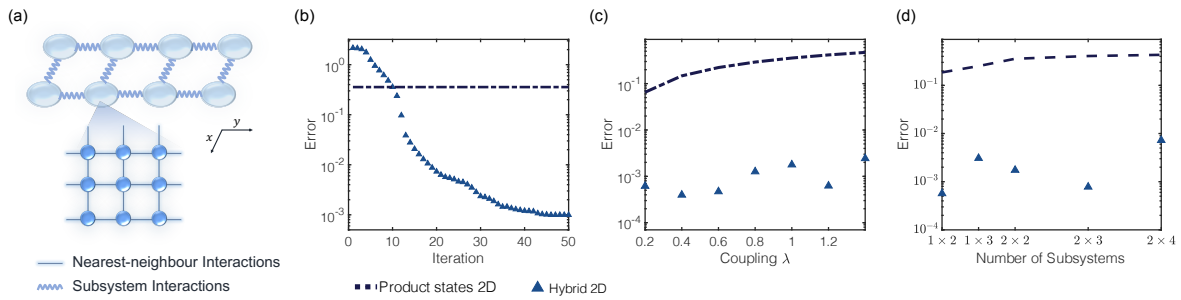


Figure A.3: Numerical simulation for a 2D spin system on a square lattice using hybrid TTNs. (a) Sketch for the 2D spin lattice with nearest-neighbour couplings on the boundary. The interactions between subsystems are represented by the thick jagged lines. 3×3 qubits are grouped on a square sublattice as subsystems. The grouping strategy is the same as that in Figure 3.6. (b)-(d) Simulation results of the ground state energy compared to the results from PEPS E_{PEPS}. The relative error $1 - E/E_{\text{PEPS}}$ is used to characterise the calculation accuracy. The blue dash-dotted line corresponds to the energy when using a tensor product of the ground state of each local subsystem. The blue triangles correspond to the results obtained with hybrid TNs. (b) Convergence towards the ground state energy for the 2D 9×4 systems with $\lambda = 1$. (c) Error versus different coupling strength of subsystems λ for 9×4 systems. (d) Errors with different numbers of local subsystems with $9 \times k$ qubits and $\lambda = 1$. Here $k = N_x \times N_y$ is considered for the 2D system.

To further test the validity of our method, I numerically test a spin lattice model with uniform nearest-neighbour interactions on the boundary, as opposed to the model with random interactions between subsystems discussed in Chapter 3. I consider uniform interactions between subsystems by setting the coupling strengths as a constant $f_{j,i} = \lambda$. The parameters of the local interactions and external fields, and the quantum circuits for the simulation are used the same as those in Figure 3.6. The lattice model and the partitioning strategy are shown in Figure A.3(a). Figure A.3(b) shows the convergence of the ground state energy, both with coupling strength $\lambda = 1$ on 9×4 qubits. Figure A.3(c, d) show how the coupling strength or the number of subsystems affect the efficacy of hybrid TTN, respectively. The simulation results again support the effectiveness of our method.

Appendix B

Eigenstate energy and property estimation

In this chapter, I first show the comparison of ground state energy estimation and observable estimation on the ground state with other representative methods discussed in Chapter 4 and Chapter 9. Then, I provide complexity analysis of the eigenstate energy and property estimation to support the results in Chapter 4 and Chapter 9. The projection methods and phase estimation methods may be further improved in terms of the initial state dependence.

Estimation	Max time	Repetition	Extra qubits	Total time complexity
This work	$\mathcal{O}(\log(p_0^{-1})\epsilon^{-1})$	$\mathcal{O}(p_0^{-2})$	1	$\mathcal{O}(\log(p_0^{-1})p_0^{-2}\epsilon^{-1})$
Lin & Tong 21 [48]	$\mathcal{O}(\log(p_0^{-1}\epsilon^{-1})\epsilon^{-1})$	$\mathcal{O}(p_0^{-2})$	1	$\mathcal{O}(\log(p_0^{-1}\epsilon^{-1})p_0^{-2}\epsilon^{-1})$
Time series [182]	$\mathcal{O}(\mathrm{Polylog}(p_0^{-1}\epsilon^{-1})\epsilon^{-1})$	$\mathcal{O}(\epsilon^{-3}p_0^{-2})$	1	$\tilde{\mathcal{O}}(p_0^{-2}\epsilon^{-4})$
Phase estimation	$\tilde{\mathcal{O}}(p_0^{-1}\epsilon^{-1})$	$\mathcal{O}(p_0^{-1})$	$\log(\epsilon^{-1})+\log(p_0^{-1})$	$\tilde{\mathcal{O}}(p_0^{-2}\epsilon^{-1})$
Projection [29]	$\tilde{\mathcal{O}}(p_0^{-1/2}\epsilon^{-3/2})$	$\tilde{\mathcal{O}}(p_0^{-1/2})$	$\log(\epsilon^{-1})$	$\tilde{\mathcal{O}}(p_0^{-1}\epsilon^{-3/2})$

Table B.1: Comparison of ground state energy estimation. The result of this work in the table is based on Theorem 1 using the Gaussian cooling function. Our work is mainly compared with the oracle-free methods.

Estimation	Max time	Repetition	Extra qubits	Total time complexity
This work	$\mathcal{O}(\log(p_0^{-1}\epsilon^{-1})\Delta^{-1})$	$\mathcal{O}(p_0^{-2}\epsilon^{-2})$	1	$\tilde{\mathcal{O}}(p_0^{-2}\epsilon^{-2}\Delta^{-1})$
Phase estimation (known E_0)	$\tilde{\mathcal{O}}(p_0^{-1}\epsilon^{-1}\Delta^{-1})$	$\mathcal{O}(p_0^{-1}\epsilon^{-2})$	$\log(\epsilon^{-1})+\log(\Delta^{-1})$	$\tilde{\mathcal{O}}(p_0^{-2}\epsilon^{-3}\Delta^{-1})$
Projection (known E_0) [29]	$\tilde{\mathcal{O}}(p_0^{-1/2}\Delta^{-1})$	$\tilde{\mathcal{O}}(p_0^{-1/2}\epsilon^{-2})$	$\log\log(\epsilon^{-1})+\log(\Delta^{-1})$	$\tilde{\mathcal{O}}(p_0^{-1}\epsilon^{-2}\Delta^{-1})$
Projection (Unknown E_0) [29]	$\tilde{\mathcal{O}}(p_0^{-1/2}\Delta^{-3/2})$	$\tilde{\mathcal{O}}(p_0^{-1/2}\epsilon^{-2}\Delta^{-1/2})$	$\log\log(\epsilon^{-1})+\log(\Delta^{-1})$	$\tilde{\mathcal{O}}(p_0^{-1}\epsilon^{-2}\Delta^{-2})$

Table B.2: Comparison of observable estimation on the ground state. The result of this work in the table is based on Theorem 2 using Gaussian cooling function. The methods are mainly compared to those without the block encoding of Hamiltonian.

B.1 Eigenenergy estimation

In Section B.1.1, the distance between $D_\tau(E)$ and $p_j g(\tau(E-E_j))^2$ caused by finite τ when $E \in [E_j - \frac{\Delta}{2}, E_j + \frac{\Delta}{2}]$ is bounded. In Section B.1.2, the estimation error of the normalisation factor $D_\tau(E)$ caused by a normalised cutoff time x_m is bounded. In Section B.1.3, based on the

measurement using Hadamard test, the statistical error of the estimation of $D_\tau^{(x_m)}(E)$ caused by the finite sampling error N_M is bounded. Finally, in Section B.1.4, I consider the peak-value search problem using the estimator and analyse the eigenenergy accuracy dependence to the circuit depth and sample complexity. We shall see how the cooling bandwidth $g^{-1}(1-\varepsilon)$ affects the accuracy of the peak-value search.

B.1.1 Finite imaginary evolution time

I first demonstrate how the distance between $D_\tau(E)$ and $p_j g(\tau(E-E_j))^2$ caused by finite τ when $E \in [E_j - \frac{\Delta}{2}, E_j + \frac{\Delta}{2}]$ can be bounded.

Proposition 13 (Error of the normalisation function introduced by finite imaginary time)**.** *When $E \in [E_j - \frac{\Delta}{2}, E_j + \frac{\Delta}{2}]$, we have*

$$|D_\tau(E) - p_j g(\tau(E-E_j))^2| \leq \varepsilon_\tau^{(1)}, \tag{B.1}$$

when $\tau \geq \frac{2}{\Delta} g^{-1}(\frac{\varepsilon_\tau^{(1)}}{2})$. When $E = E_j$, we can improve the requirement to $\tau \geq \frac{1}{\Delta} g^{-1}(\frac{\varepsilon_\tau^{(1)}}{2})$.

The basic idea is that the imaginary time evolution drives the initial state exponentially faster to the target state, as that have been shown in Section 2.3. Here, I simply give an intuition for the proof and refer to the author's paper [93] for the formal proof.

B.1.2 Finite normalised cutoff time

I then demonstrate how the estimation error of $D_\tau(E)$ caused by a normalised cutoff time x_m can be bounded. The estimation formula is given in Eq. (4.14),

$$D_\tau^{(x_m)}(E) = \int_{-x_m}^{x_m} dy \tilde{p}(y) e^{-i\tau y E_j} \langle \psi_0 | e^{i\tau y H} | \psi_0 \rangle . \tag{B.2}$$

Here, $\tilde{p}(y) = \int_{-\infty}^{\infty} p(t-y)p(t)dt = [p \star p](y)$.

Proposition 14 (Error of the normalisation factor introduced by finite normalised cutoff time)**.** *The truncated estimation $D_\tau^{(x_m)}(E)$ defined in Eq. (B.2) is related to the normalisation factor defined in Eq. (4.13) by*

$$|D_\tau^{(x_m)}(E) - D_\tau(E)| \leq \varepsilon_x^{(1)}, \forall E \in \mathbb{R}, \tag{B.3}$$

when the normalised cutoff time $x_m \geq \sqrt{2} L(\frac{\varepsilon_x^{(1)}}{2})$. Here, $L(\varepsilon)$ is the tail function of the cooling function $g(h)$ defined in Definition 2.

The basic idea is to analyse the error dependence on the cutoff of the tail function. We can find that the error exponentially decays with an increasing cutoff x_m for Gaussian cooling function. Again, I refer to [93] for the formal proof.

B.1.3 Finite sampling number

In this section, I consider the statistical fluctuation when estimating the values of $D_\tau^{(x_m)}(E)$ using Eq. (B.2). In the single-shot version of Hadamard test in Algorithm 4, we can describe the independent and identical distribution (i.i.d.) of single-round estimators for $D_\tau^{(x_m)}(E)$ as $\{\hat{d}_p(E)\}_{p=1}^{N_M}$. Each single-round estimator $\hat{d}_p(E)$ is a random variable defined as follows,

$$\hat{d}_p = \begin{cases} \mathrm{Re}(2\, e^{-i\tau y E}), & (b,a) = (0,0), y : |y| \leq x_m \\ \mathrm{Re}(-2\, e^{-i\tau y E}), & (b,a) = (0,1), y : |y| \leq x_m \\ \mathrm{Re}(2i\, e^{-i\tau y E}), & (b,a) = (1,0), y : |y| \leq x_m \\ \mathrm{Re}(-2i\, e^{-i\tau y E}), & (b,a) = (1,1), y : |y| \leq x_m \\ 0, & y : |y| > x_m. \end{cases} \tag{B.4}$$

The final estimation of $D_\tau^{(x_m)}(E)$ is given by

$$\hat{D}_\tau^{(x_m)}(E) = \frac{1}{N_M} \sum_{p=1}^{N_M} \hat{d}_p(E). \tag{B.5}$$

Based on Eq. (B.2), we know that

$$\mathbb{E}_{(y,b,a)} \left(\hat{D}_\tau^{(x_m)} \right) = D_\tau^{(x_m)}. \tag{B.6}$$

To analyse the statistical fluctuation, we apply the Hoeffding bound.

Lemma 2 (Hoeffding bound)**.** *For n independent random variables $\{\hat{X}_i\}_{i=1}^n$ which are bounded by $[a,b]$, the average value*

$$\bar{X} := \frac{1}{n} \sum_{i=1}^{n} \hat{X}_i, \tag{B.7}$$

satisfies

$$\Pr(\left|\bar{X} - \mathbb{E}(\bar{X})\right| \geq \varepsilon) \leq 2 \exp\left(-\frac{2n\varepsilon^2}{(b-a)^2} \right). \tag{B.8}$$

Proposition 15 (Error of the normalisation factor introduced by finite sampling number)**.** *The estimator $\hat{D}_\tau^{(x_m)}(E)$ defined in Eq. (B.5) is related to the truncated estimation $\hat{D}_\tau^{(x_m)}(E)$ defined in Eq. (B.2) by*

$$|\hat{D}_\tau^{(x_m)}(E) - D_\tau^{(x_m)}(E)| \leq \varepsilon_n^{(1)}, \forall E \in \mathbb{R}, \tag{B.9}$$

with a failure probability $\delta^{(1)} := 2 \exp\left(-2K(\varepsilon_n^{(1)})^2/16\right)$, when the sample number $N_M \geq K(\varepsilon_n^{(1)})^{-2}$.

Proof. Note that the estimators $\{\hat{d}_p(E)\}_{p=1}^{N_M}$ are independent, whose values are bounded by $[-2,2]$. Using the Hoeffding bound in Lemma 2 with the bound $[-2,2]$, we finish the proof. □

B.1.4 Accuracy of the eigenenergy estimation

Ideally, the eigenenergy E_j satisfies,

$$E_j = \text{argmax}_{E\in[E_j^L,E_j^U]} D_\tau(E). \tag{B.10}$$

In practice, what we can solve is the following problem,

$$\hat{E}_j = \text{argmax}_{E\in[E_j^L,E_j^U]} \hat{D}_\tau^{(x_m)}(E). \tag{B.11}$$

We now can show that when $[E_j^L, E_j^U] \subset [E_j - \frac{\Delta}{2}, E_j + \frac{\Delta}{2}]$, $\hat{E}_j$ is a good estimation of E_j with precision κ when the finite imaginary time $\tau = \mathcal{O}(\kappa^{-1})$, finite normalised cutoff time $x_m = \mathcal{O}(L(p_j))$, and sample number $N_M = \mathcal{O}(p_j^{-2})$.

Theorem 8 (Accuracy of the eigenenergy estimation)**.** *When $[E_j^L, E_j^U] \subset [E_j - \frac{\Delta}{2}, E_j + \frac{\Delta}{2}]$, the eigenenergy estimation $\hat{E}_j$ defined in Eq. (B.11) is related to the eigenenergy E_j by*

$$|\hat{E}_j - E_j| \leq \kappa, \tag{B.12}$$

with a failure probability of $2\delta^{(1)} = 4\exp(-K/8)$, when the finite imaginary time $\tau \geq \frac{1}{\kappa} g^{-1}(\frac{1-g(1)}{6} p_j)$, the normalised cutoff time $x_m \geq \sqrt{2}L\left(\frac{1-g(1)}{6} p_j\right)$, and the sample number $N_M \geq \frac{9K}{(1-g(1))^2} p_j^{-2}$.

Proof. From Eq. (B.11) we know that,

$$\hat{D}_\tau^{(x_m)}(\hat{E}_j) \geq \hat{D}_\tau^{(x_m)}(E_j). \tag{B.13}$$

Now, we take $x_m = \sqrt{2}L(\frac{\varepsilon_x p_j}{2})$ and $N_M = (\varepsilon_n p_j)^{-2} K$. Using Proposition 14 and 15, we have

$$\begin{aligned} |\hat{D}_\tau^{(x_m)}(\hat{E}_j) - D_\tau(\hat{E}_j)| &\leq |D_\tau^{(x_m)}(\hat{E}_j) - D_\tau(\hat{E}_j)| + |\hat{D}_\tau^{(x_m)}(\hat{E}_j) - D_\tau^{(x_m)}(\hat{E}_j)| \leq p_j(\varepsilon_x + \varepsilon_n), \\ |\hat{D}_\tau^{(x_m)}(E_j) - D_\tau(E_j)| &\leq |D_\tau^{(x_m)}(E_j) - D_\tau(E_j)| + |\hat{D}_\tau^{(x_m)}(E_j) - D_\tau^{(x_m)}(E_j)| \leq p_j(\varepsilon_x + \varepsilon_n) \end{aligned}, \tag{B.14}$$

each with a failure probability $\delta^{(1)}$. Furthermore, when $\tau = \frac{2}{\Delta} g^{-1}(\frac{p_j \varepsilon_\tau}{2})$, from Proposition 13 we have

$$\begin{aligned} |D_\tau(\hat{E}_j) - p_j g(\tau(E_j - \hat{E}_j))^2| &\leq p_j \varepsilon_\tau, \\ |D_\tau(E_j) - p_j| &\leq p_j \varepsilon_\tau. \end{aligned} \tag{B.15}$$

Combine Eq. (B.13), Eq. (B.14), and Eq. (B.15), we have

$$\begin{aligned} p_j &\leq D_\tau(E_j) + \varepsilon_\tau \leq \hat{D}_\tau^{(x_m)}(E_j) + p_j(\varepsilon_x + \varepsilon_n) + p_j \varepsilon_\tau \\ &\leq \hat{D}_\tau^{(x_m)}(\hat{E}_j) + p_j(\varepsilon_x + \varepsilon_n) + p_j \varepsilon_\tau \\ &\leq D_\tau(\hat{E}_j) + 2p_j(\varepsilon_x + \varepsilon_n) + p_j \varepsilon_\tau, \\ &\leq p_j g(\tau(E_j - \hat{E}_j))^2 + 2p_j(\varepsilon_\tau + \varepsilon_x + \varepsilon_n), \end{aligned} \tag{B.16}$$

with a failure pribability $2\delta^{(1)}$. Therefore,

$$g(\tau(\hat{E}_j - E_j))^2 \geq 1 - 2(\varepsilon_\tau + \varepsilon_x + \varepsilon_n). \tag{B.17}$$

Denote $\kappa := \hat{E}_j - E_j$, we further simplify the expression above,

$$\begin{aligned}
\kappa &\leq \frac{1}{\tau} g^{-1}(\sqrt{1 - 2(\varepsilon_\tau + \varepsilon_x + \varepsilon_n)}) \\
&\leq \frac{1}{\tau} g^{-1}(1 - (\varepsilon_\tau + \varepsilon_x + \varepsilon_n)) \\
&= \frac{1}{\tau} g^{-1}\left(1 - \frac{2}{p_j} g(\frac{\Delta}{2}\tau) - \frac{2}{p_j} L^{-1}(\frac{x_m}{\sqrt{2}}) - \sqrt{\frac{K}{p_j^2 N_M}}\right)
\end{aligned} \tag{B.18}$$

In the second inequality, we use the property that $g^{-1}(p)$ is a decreasing function when $p > 0$. Therefore, the following requirement is sufficient to make sure the eigenenergy error is smaller than κ,

$$\begin{aligned}
\tau &\geq \frac{1}{\kappa}, \\
\frac{2}{p_j} g(\frac{\Delta}{2}\tau) + \frac{2}{p_j} L^{-1}(\frac{x_m}{\sqrt{2}}) + \sqrt{\frac{K}{p_j^2 N_M}} &\leq 1 - g(1).
\end{aligned} \tag{B.19}$$

Let $\frac{2}{p_j} g(\frac{\Delta}{2}\tau) = \frac{2}{p_j} L^{-1}(\frac{x_m}{\sqrt{2}}) = \sqrt{\frac{K}{p_j^2 N_M}}$, we have

$$\begin{aligned}
\tau &\geq \frac{2}{\Delta} g^{-1}(\frac{1 - g(1)}{6} p_j), \\
x_m &\geq \sqrt{2} L\left(\frac{1 - g(1)}{6} p_j\right) \\
N_M &\geq K\left(\frac{3}{1 - g(1)}\right)^2 \frac{1}{p_j^2}.
\end{aligned} \tag{B.20}$$

In the above derivation, we require $\tau \geq \frac{1}{\kappa}$ and $\tau \geq \frac{2}{\Delta} g^{-1}(\frac{1-g(1)}{6} p_j)$. In practice, $\kappa \ll \frac{\Delta}{2}$. Then it suffices to have $\tau \geq \frac{1}{\kappa} g^{-1}(\frac{1-g(1)}{6} p_j)$. □

From Theorem 8 we can see that, the (maximal) circuit depth and sample complexity of Algorithm 4 are $\tau x_m = \mathcal{O}(\kappa^{-1} g^{-1}(p_j) L(p_j))$ and $N_M = \mathcal{O}(p_j^{-2})$, respectively, for a given accuracy κ and initial state overlap p_j. This achieves the Heisenberg limit $\frac{1}{\kappa}$ for the eigenenergy searching.

B.2 Algorithms

In this section, I show the algorithms for eigenenergy estimation and normalisation factor estimation in Algorithm 4. Other algorithms can be found in [93].

Algorithm 4 Eigenenergy and normalisation factor estimation

Require: An n-qubit Hamiltonian H; initial state $|\psi_0\rangle$ with nonzero overlap with jth eigenstate of H: $p_j = |\langle u_j|\psi_0\rangle|^2$; the energy interval $[E_j^L, E_j^U]$ for the jth eigenstate; cooling function $g(h)$ and the corresponding sampling probability $p(x)$; proper choice of imaginary time τ, normalised cutoff time x_m.

Ensure: Estimation of the eigenenergy E_j and the corresponding normalisation factor $\hat{D}_\tau^{(x_m)}$.

for $p = 1$ **to** N_M **do**
  Sample a normalised evolution time y_p from $\tilde{p}(y) := \frac{1}{2}\int_{-\infty}^{\infty} p(\frac{z+y}{2})p(\frac{z-y}{2})dz$.
  if $y_p > x_m$ **then** ▷ Exceed the preset cutoff value
    Set the estimation $\hat{d}_p = 0$.
  else ▷ Normal single-shot quantum sampling by Hadamard test
    Prepare an ancillary qubit on $|+\rangle$ state and the initial state $|\psi_0\rangle$.
    Implement a controlled-unitary C-U from the ancillary qubit to the state $|\psi_0\rangle$. Here, $U = e^{i\tau y_p H}$.
    Generate a random bit b_p with a uniform distribution of values $\{0, 1\}$. Perform a gate W^{b_p} on the ancillary qubit. Here, $S = \mathrm{diag}(1, -i)$ is a $\frac{\pi}{4}$-rotation gate.
    Measure the ancillary qubit on X-basis, and record the binary result a_p. Then record the values $\{y_p, b_p, a_p\}$ and set the temporal estimation value $\hat{d}_p := 2(-1)^{a_p} i^{b_p}$.
  end if
end for
Set the estimated normalisation factor $\hat{D}_\tau^{(x_m)'} = 0$.
for E_j' **in** $[E_j^L, E_j^U]$ **do** ▷ Try different possible eigenenergy value
  Calculate the estimated normalisation factor $\hat{D}_\tau^{(x_m)'} := \mathrm{Re}\left(\frac{1}{N_M}\sum_{p=1}^{N_M} e^{-i\tau y_p E_j'}\hat{d}_p\right)$.
  if $\hat{D}_\tau^{(x_m)'}$ **then**
    Set $\hat{D}_\tau^{(x_m)} = \hat{D}_\tau^{(x_m)'}$ and $E_j = E_j'$.
  end if
end for

B.3 Eigenstate property estimation

In this section, I analyse the complexity of eigenstate property estimation in a similar vein to eigenenergy estimation in Section B.1.

In Section B.3.1, the effect of finite imaginary time τ is considered; in Section B.3.2, the estimation error caused by a normalised cutoff time x_m is bounded. The two factors above determine the actual maximum evolution time. In Section B.3.3, based on the measurement using Hadamard test, we determine the statistical error caused by finite sampling number N_M. Finally, I summarise the finite sampling effect in Section B.3.4, and show the dependence of circuit depth t_m and sampling number N_M with respect to the initial state $|\psi_0\rangle$, Hamiltonian H, and observable O.

B.3.1 Finite imaginary evolution time

Recall that

$$\langle O\rangle = \langle u_j|O|u_j\rangle = \frac{N(O)}{D}, \tag{B.21}$$

where

$$D = D(E_j) = \langle\psi_0|\hat{P}_j|\psi_0\rangle = p_j,$$
$$N(O) = \langle\psi|\hat{P}_j O \hat{P}_j|u_j\rangle = p_j \langle O\rangle, \tag{B.22}$$

are, respectively, the normalisation factor and the unnormalised expectation value. Here, $\hat{P}_j := |u_j\rangle\langle u_j|$ is the projector to the jth eigenstate.

When the imaginary time τ is finite, we have the approximated value

$$\langle O\rangle_\tau = \frac{N_\tau(O)}{D_\tau}, \tag{B.23}$$

where

$$D_\tau = \langle\psi_0|g^2(\tau(H-E_j))|\psi_0\rangle,$$
$$N_\tau(O) = \langle\psi_0|g(\tau(H-E_j))\, O\, g(\tau(H-E_j))|\psi_0\rangle. \tag{B.24}$$

are, respectively, the normalisation factor and the unnormalised expectation value with finite τ.

In Proposition 13 in Section B.1.1, the distance between $D_\tau(E)$ and $D(E)$ when $|E-E_j| \leq \frac{\Delta}{2}$ is shown. Similarly, the distance between $N_\tau(O)$ and $N(O)$ can be bounded.

Proposition 16 (Error of the unnormalised expectation value introduced by finite imaginary time)**.** *The unnormalised expectation value $N_\tau(O)$ defined in Eq. (B.24) is related to the ideal $N(O)$ defined in Eq. (B.22) by*

$$|N_\tau(O) - N(O)| \leq \|O\|_\infty \varepsilon_\tau^{(2)}, \tag{B.25}$$

when the imaginary time $\tau \geq \frac{1}{\Delta} g^{-1}(\frac{\varepsilon_\tau^{(2)}}{2})$. Here, $\|O\|_\infty$ is the spectral norm of O.

The idea of bounding the distance between $N_\tau(O)$ and $N(O)$ is that the distance between $\hat{P}_j$ and $g(\tau(H-E_j))$ can be bounded, and then the difference between $N_\tau(O)$ and $N(O)$ can be bounded naturally.

B.3.2 Finite normalised cutoff time

In Proposition 14, I demonstrate how the estimation error of $D_\tau(O)$ caused by a finite x_m can be bounded. Following similar methods in Section B.1, I bound the estimation errors of $N_\tau(O)$ caused by a finite x_m. The estimation formula is given in Eq. (4.14),

$$N_\tau^{(x_m)}(O) = \|O\|_1 \int_{-x_m}^{(x_m)} dx \int_{-x_m}^{x_m} dx' p(x)p(x') \sum_{l\in\mathcal{P}_n} \mathrm{Pr}_O(l) e^{-i\tau(x-x')E_j} \langle\psi_0|e^{-i\tau x' H} P_l e^{i\tau x' H}|\psi_0\rangle, \tag{B.26}$$

where

$$O = \sum_{l\in\mathcal{P}_n} o_l P_l = \|O\|_1 \sum_{l\in\mathcal{P}_n} \mathrm{Pr}_O(l) P_l, \tag{B.27}$$

$\|O\|_1$ is the l_1-norm of Pauli coefficients of O,

$$\|O\|_1 = \sum_{l\in\mathcal{P}_n} o_l. \tag{B.28}$$

Note that the coefficients $\{o_l\}_l$ are all positive, since we put the signs into the Pauli matrices $\{P_l\}$. The probability distribution $\mathrm{Pr}_O(l)$ is defined to be,

$$\mathrm{Pr}_O(l) = \frac{o_l}{\sum_{l\in\mathcal{P}_n} o_l} = \frac{1}{\|O\|_1} o_l. \tag{B.29}$$

Proposition 17 (Error of the unnormalised expectation value introduced by finite normalised cutoff time)**.** *The truncated estimation $N_\tau^{(x_m)}(O)$ defined in Eq. (B.26) is related to the normalisation factor $N_\tau(O)$ defined in Eq. (B.24) by*

$$|N_\tau^{(x_m)}(O) - N_\tau(O)| \leq \|O\|_\infty \varepsilon_x^{(2)}, \tag{B.30}$$

when the normalised cutoff time $x_m \geq L(\frac{\varepsilon_x^{(2)}}{2})$. Here, $L(\varepsilon)$ is the tail function of the cooling function $g(h)$ defined in Definition 2.

Proof. We have,

$$\begin{aligned}
|N_\tau^{(x_m)}(O) - N_\tau(O)| &\leq \iint_{\bar{S}} dxdx'\, p(x)p(x') \left| e^{-i\tau(x-x')E_j} \langle\psi_0|e^{-i\tau x'H} O e^{i\tau x'H}|\psi_0\rangle \right| \\
&\leq \iint_{\bar{S}} dxdx'\, p(x)p(x') \left| \langle\psi_0|e^{-i\tau x'H} O e^{i\tau x'H}|\psi_0\rangle \right| \\
&\leq \|O\|_\infty \iint_{\bar{S}} dxdx'\, p(x)p(x') =: \|O\|_\infty \varepsilon_x^{(2)},
\end{aligned} \tag{B.31}$$

where $\bar{S}$ implies the complement area of $S : \{(x,x')||x| \leq x_m, |x'| \leq x_m\}$.

The error term $\varepsilon_x^{(2)}$ is related to the tail probability by

$$\begin{aligned}
&\varepsilon_x^{(2)} = 1 - \iint_S dxdx'\, p(x)p(x') = 1 - (1 - L^{-1}(x_m))^2 \leq 2L^{-1}(x_m) \\
\Rightarrow \quad & x_m \geq L(\frac{\varepsilon_x^{(2)}}{2}).
\end{aligned} \tag{B.32}$$

Therefore, when we choose the x_m to be $L(\varepsilon_x^{(2)})$, we can achieve the estimation of $N_\tau(O)$ with a precesion of $\|O\|_\infty \varepsilon_x^{(2)}$.

□

B.3.3 Finite sample number

In Proposition 14, I bound the estimation error of $D_\tau^{(x_m)}(O)$ caused by finite sampling error N_M. Following similar methods in Section B.1, I bound the statistical error of $N_\tau^{(x_m)}(O)$.

In the single-shot version of Hadamard test, we can describe these independent and identical distribution (i.i.d.) of single-round estimators for $N_\tau^{(x_m)}(O)$ as $\{\hat{n}_q\}_{q=1}^{N_M}$. Each single-round estimator $\hat{n}_q$ is a random variable defined as follows,

$$\hat{n}_q(O) = \begin{cases}
\mathrm{Re}(2\|O\|_1\, e^{-i\tau(x-x')E_j}), & (b,a) = (0,0), (x,x') : |x| \leq x_m, |x'| \leq x_m \\
\mathrm{Re}(-2\|O\|_1\, e^{-i\tau(x-x')E_j}), & (b,a) = (0,1), (x,x') : |x| \leq x_m, |x'| \leq x_m \\
\mathrm{Re}(2i\|O\|_1\, e^{-i\tau(x-x')E_j}), & (b,a) = (1,0), (x,x') : |x| \leq x_m, |x'| \leq x_m \\
\mathrm{Re}(-2i\|O\|_1\, e^{-i\tau(x-x')E_j}), & (b,a) = (1,1), (x,x') : |x| \leq x_m, |x'| \leq x_m \\
0, & (x,x') : |x| > x_m \text{ or } |x'| > x_m.
\end{cases} \tag{B.33}$$

The final estimation of $N_\tau^{(x_m)}(O)$ is given by

$$\hat{N}_\tau^{(x_m)}(O) = \frac{1}{N_M}\sum_{q=1}^{N_M}\hat{n}_q(O). \tag{B.34}$$

Based on Eq. (B.26), we know that

$$\mathbb{E}_{(x,x',P,b,a)}\left(\hat{N}_\tau^{(x_m)}(O)\right) = N_\tau^{(x_m)}(O). \tag{B.35}$$

Now I bound the statistical error of $\hat{N}_\tau^{(x_m)}(O)$ using Lemma 2.

Proposition 18 (Error of the observable expectation value introduced by finite sampling number)**.** *The estimator $\hat{N}_\tau^{(x_m)}(O)$ defined in Eq. (B.34) is related to the truncated estimation $\hat{N}_\tau^{(x_m)}(O)$ defined in Eq. (B.26) by*

$$|\hat{N}_\tau^{(x_m)}(O) - N_\tau^{(x_m)}(O)| \leq \|O\|_1\varepsilon_n^{(2)}, \tag{B.36}$$

with a failure probability $\delta^{(2)} := 2\exp\left(-2K(\varepsilon_n^{(2)})^2/16\right)$, when the sample number $N_M \geq K(\varepsilon_n^{(2)})^{-2}$. Here, $\|O\|_1 = \sum_l |o_l|$ is the sum of the Pauli coefficients of O.

Proof. Note that the estimators $\{\hat{n}_q(O)\}_{q=1}^{N_M}$ are independent, whose values are bounded by $[-2\|O\|_1, 2\|O\|_1]$. Here, $\|O\|_1 = \sum_l |o_l|$. Using the Hoeffding bound in Lemma 2 with the bound $[-2\|O\|_1, 2\|O\|_1]$, we finish the proof. □

B.3.4 Accuracy of the observable estimation

The final observable estimation is given by

$$\langle\hat{O}\rangle_\tau^{(x_m)} = \frac{\hat{N}_\tau^{(x_m)}}{\hat{O}_\tau^{(x_m)}}. \tag{B.37}$$

In Proposition 13, 14, 15, 16, 17, and 18, we estimate the errors caused by finite τ, x_m, and N_M on D_τ and $N_\tau(O)$, respectively. Combining the results together, we have the following result.

Proposition 19. *For the initial state $|\psi_0\rangle$, target eigenstate $|u_j\rangle$ and cooling function $g(h)$, if we set the imaginary time τ, normalised cutoff time x_m, and sample number N to be*

1. $\tau = \frac{1}{\Delta}\max\{g^{-1}(\frac{\varepsilon_\tau^{(1)}}{2}, g^{-1}(\frac{\varepsilon_\tau^{(2)}}{2})\}$;
2. $x_m = \max\{\sqrt{2}L(\frac{\varepsilon_x^{(1)}}{2}), L(\frac{\varepsilon_x^{(2)}}{2})\}$;
3. $N_M = \max\{K(\varepsilon_n^{(1)})^{-2}, K(\varepsilon_n^{(2)})^{-2}\}$;

then the estimated observable expectation value $\langle\hat{O}\rangle_\tau^{(x_m)}$ defined in Eq. (B.37) is related to the real observable expectation value $\langle O\rangle$ by,

$$|\langle\hat{O}\rangle_\tau^{(x_m)} - \langle O\rangle| \leq p_j^{-1}(\langle O\rangle + 1)(\varepsilon_\tau^{(1)} + \varepsilon_x^{(1)} + \varepsilon_n^{(1)}) + p_j^{-1}(\varepsilon_\tau^{(2)}\|O\|_\infty + \varepsilon_x^{(2)}\|O\|_\infty + \varepsilon_n^{(2)}\|O\|_1), \tag{B.38}$$

with a failure probability $\delta^{(1)} + \delta^{(2)} = 4e^{-\frac{K}{2}}$.

Proof. Following Proposition 13, 14, 15, 16, 17, and 18, when we choose τ, x_m, and N_M to be the values mentioned above, then the following bound holds at the same time

$$\begin{aligned} \left|\hat{D}_\tau^{(x_m)} - D\right| &\leq \varepsilon_\tau^{(1)} + \varepsilon_x^{(1)} + \varepsilon_n^{(1)}, \\ \left|\hat{N}_\tau^{(x_m)}(O) - N(O)\right| &\leq (\varepsilon_\tau^{(2)}\|O\|_\infty + \varepsilon_x^{(2)}\|O\|_\infty + \varepsilon_n^{(2)}\|O\|_1), \end{aligned} \tag{B.39}$$

with a failure probability $\delta(K) = 4e^{-\frac{K}{8}}$. To simplify the notation, denote $\varepsilon^{(1)} := \varepsilon_\tau^{(1)} + \varepsilon_x^{(1)} + \varepsilon_n^{(1)}$ and $\varepsilon^{(2)}(O) := \varepsilon_\tau^{(2)}\|O\|_\infty + \varepsilon_x^{(2)}\|O\|_\infty + \varepsilon_n^{(2)}\|O\|_1$.

Based on Eq. (B.39), difference between the quotients $\langle\hat{O}\rangle_\tau^{(x_m)}$ and $\langle O\rangle$ is bounded as

$$\begin{aligned} |\langle O\rangle - \langle\hat{O}\rangle_\tau^{(x_m)}| &\leq \frac{(N(O)+D)\varepsilon^{(1)} + D\varepsilon^{(2)}(O))}{D^2} \\ &= p_j^{-1}(\langle O\rangle + 1)(\varepsilon_\tau^{(1)} + \varepsilon_x^{(1)} + \varepsilon_n^{(1)}) + p_j^{-1}(\varepsilon_\tau^{(2)}\|O\|_\infty + \varepsilon_x^{(2)}\|O\|_\infty + \varepsilon_n^{(2)}\|O\|_1). \end{aligned} \tag{B.40}$$

□

Based on Proposition 19, we can estimate the circuit depth and sample complexity for the observable estimation.

Theorem 9 (Accuracy of the observable estimation)**.** *The expectation value estimation $\langle\hat{O}\rangle_\tau^{(x_m)}$ defined in Eq. (B.37) is related to the real observable expectation value $\langle O\rangle$ by*

$$\begin{aligned} |\langle\hat{O}\rangle_\tau^{(x_m)} - \langle O\rangle| &\leq \varepsilon\left(\frac{1}{2}(\langle O\rangle + 1) + \frac{1}{3}\|O\|_\infty + \frac{1}{6}\|O\|_1\right) \\ &\leq \varepsilon(\|O\|_1 + 1) \end{aligned} \tag{B.41}$$

with a failure probability of $\delta^{(1)} + \delta^{(2)} = 4\exp(-K/8)$, when the finite imaginary time $\tau \geq \frac{1}{\Delta}g^{-1}\left(\frac{\varepsilon p_j}{12}\right)$, the normalised cutoff time $x_m \geq \sqrt{2}L\left(\frac{\varepsilon p_j}{12}\right)$, and the sample number $N_M \geq K\left(\frac{\varepsilon p_j}{6}\right)^{-2}$. Here, $\|O\|_1 = \sum_l |o_l|$ is the sum of the Pauli coefficients of O, Δ is a known lower bound of $\min\{|E_j - E_{j-1}|, |E_j - E_{j+1}|\}$.

Proof. From Proposition 19 we know that, if we set $\tau = \frac{1}{\Delta}\max\{g^{-1}(\frac{\varepsilon_\tau^{(1)}}{2}), g^{-1}(\frac{\varepsilon_\tau^{(2)}}{2})\}$ $x_m = \max\{\sqrt{2}L(\frac{\varepsilon_x^{(1)}}{2}), L(\frac{\varepsilon_x^{(2)}}{2})\}$ and $N_M = \max\{K(\varepsilon_n^{(1)})^{-2}, K(\varepsilon_n^{(2)})^{-2}\}$, then

$$\begin{aligned} |\langle O\rangle - \langle\hat{O}\rangle_\tau^{(x_m)}| &\leq p_j^{-1}(\langle O\rangle + 1)(\varepsilon_\tau^{(1)} + \varepsilon_x^{(1)} + \varepsilon_n^{(1)}) + p_j^{-1}(\varepsilon_\tau^{(2)}\|O\|_\infty + \varepsilon_x^{(2)}\|O\|_\infty + \varepsilon_n^{(2)}\|O\|_1) \\ &\leq p_j^{-1}(\|O\|_1 + 1)(\varepsilon_\tau^{(1)} + \varepsilon_x^{(1)} + \varepsilon_n^{(1)}) + p_j^{-1}\|O\|_1(\varepsilon_\tau^{(2)} + \varepsilon_x^{(2)} + \varepsilon_n^{(2)})) \end{aligned} \tag{B.42}$$

If we set

$$\varepsilon_\tau^{(1)} = \varepsilon_\tau^{(2)} = \varepsilon_x^{(1)} = \varepsilon_x^{(2)} = \varepsilon_n^{(1)} = \varepsilon_n^{(2)} = \frac{\varepsilon p_j}{6}, \tag{B.43}$$

which corresponds to

$$\begin{aligned} \tau &\geq \frac{1}{\Delta}g^{-1}\left(\frac{\varepsilon p_j}{12}\right) \\ x_m &\geq \sqrt{2}L\left(\frac{\varepsilon p_j}{12}\right), \\ N_M &\geq K\left(\frac{\varepsilon p_j}{6}\right)^{-2}, \end{aligned} \tag{B.44}$$

then Eq. (B.42) becomes,

$$
\begin{aligned}
|\langle \hat{O} \rangle_{\tau}^{(x_m)} - \langle O \rangle| &\leq \varepsilon \left(\frac{1}{2}(\langle O \rangle + 1) + \frac{1}{3}\|O\|_{\infty} + \frac{1}{6}\|O\|_1 \right) \\
&\leq \varepsilon(\|O\|_1 + 1).
\end{aligned}
\tag{B.45}
$$

□

From Theorem 9 we can see that, the circuit depth and sample complexity of the observable estimation are $\tau x_m = \mathcal{O}(\frac{1}{\Delta} g^{-1}(\varepsilon p_j) L(\varepsilon p_j))$ and $N_M = \mathcal{O}(\varepsilon^{-2} p_j^{-2})$, respectively.

For the Gaussian cooling, we have

$$
g^{-1}(\varepsilon) = \mathcal{O}\left(\log(\varepsilon^{-1})\right), \tag{B.46}
$$

which is exponentially better than the phase estimation algorithm.

Appendix C

Error analysis of the adaptive strategy

C.1 Error analysis

Proof of Proposition 4. To start with, we first express the algorithmic error of a circuit with $\vec{O}$ and $\vec{\Lambda}$ in a finite time step δt as

$$\tilde{\varepsilon} = \sqrt{\|e^{-iH\delta t}G(\vec{O},\vec{\Lambda})\,|\Psi_0\rangle - G(\vec{O},\vec{\Lambda}+\vec{\lambda}\delta t)\,|\Psi_0\rangle\,\|^2}. \tag{C.1}$$

Denote the differential operator on Λ_j as $\partial_j \equiv \partial/\partial\Lambda_j$ and $\partial_{jj'} \equiv \partial_j\partial_{j'}$. We can expand $G(\vec{O},\vec{\Lambda}+\vec{\lambda}\delta t)$ as

$$G(\vec{O},\vec{\Lambda}+\vec{\lambda}\delta t) = G(\vec{O},\vec{\Lambda}) + \sum_j \lambda_j\partial_j G(\vec{O},\vec{\Lambda})\delta t + \frac{1}{2}\sum_{jj'}\lambda_j\lambda_{j'}\partial_j\partial_{j'}G(\vec{O},\vec{\Lambda})\delta t^2 + \mathcal{O}(\delta t^3) \tag{C.2}$$

Therefore, we have

$$\begin{aligned}
&e^{-iH\delta t}G(\vec{O},\vec{\Lambda}) - G(\vec{O},\vec{\Lambda}+\vec{\lambda}\delta t)\\
=&-iHG(\vec{O},\vec{\Lambda})\delta t - \sum_j \lambda_j\partial_j G(\vec{O},\vec{\Lambda})\delta t - \frac{1}{2}H^2G(\vec{O},\vec{\Lambda})\delta t^2 - \frac{1}{2}\sum_{jj'}\lambda_j\lambda_{j'}\partial^2_{jj'}G(\vec{O},\vec{\Lambda})\delta t^2 + \mathcal{O}(\delta t^3)
\end{aligned} \tag{C.3}$$

and

$$\begin{aligned}
&\|(e^{-iH\delta t}G(\vec{O},\vec{\Lambda}) - G(\vec{O},\vec{\Lambda}+\vec{\lambda}\delta t))\,|\Psi_0\rangle\,\|^2\\
=&\|-(iH+\sum_j\lambda_j\partial_j)G(\vec{O},\vec{\Lambda})\,|\Psi_0\rangle\,\delta t - \frac{1}{2}(H^2+\sum_{jj'}\lambda_j\lambda_{j'}\partial^2_{jj'})G(\vec{O},\vec{\Lambda})\,|\Psi_0\rangle\,\delta t^2 + \mathcal{O}(\delta t^3)\|^2\\
=&\|\,|\psi_1\rangle\,\delta t + |\psi_2\rangle\,\delta t^2 + \mathcal{O}(\delta t^3)\|^2 = |\psi_1\rangle\,\langle\psi_1|\,\delta t^2 + 2\mathrm{Re}\,(|\psi_1\rangle\,\langle\psi_2|)\,\delta t^3 + \mathcal{O}(\delta t^4)
\end{aligned} \tag{C.4}$$

where

$$|\psi_1\rangle \equiv -(iH+\sum_j\lambda_j\partial_j)G(\vec{O},\vec{\Lambda})\,|\Psi_0\rangle\,, \quad |\psi_2\rangle \equiv -\frac{1}{2}(H^2+\sum_{jj'}\lambda_j\lambda_{j'}\partial^2_{jj'})G(\vec{O},\vec{\Lambda})\,|\Psi_0\rangle \tag{C.5}$$

Substituting the above expansion into Eq. (C.1), we have

$$\tilde{\varepsilon} = \sqrt{\Delta^2 \delta t^2 + \Delta_2^2 \delta t^3 + \mathcal{O}(\delta t^4)} \leq \Delta \delta t + \Delta_2 \delta t^{3/2} + \mathcal{O}(\delta t^2). \tag{C.6}$$

The first-order order error Δ can be expressed as

$$\Delta^2 = |\psi_1\rangle \langle \psi_1| = \langle H^2 \rangle + \sum_{jj'} A_{jj'} \lambda_j \lambda_{j'} - 2 \sum_j C_j \lambda_j, \tag{C.7}$$

where we denote

$$A_{jj'} = \mathrm{Re}\left(\langle \Psi_0 | \, \partial_j G^\dagger(\vec{O}, \vec{\Lambda}) \partial_{j'} G(\vec{O}, \vec{\Lambda}) \, |\Psi_0\rangle \right), \quad C_j = \mathrm{Im}\left(\langle \Psi_0 | \partial_j G^\dagger(\vec{O}, \vec{\Lambda}) H G(\vec{O}, \vec{\Lambda}) | \Psi_0 \rangle \right), \tag{C.8}$$

and

$$\langle H^2 \rangle = \langle \Psi_0 | G^\dagger(\vec{O}, \vec{\Lambda}) H^2 G(\vec{O}, \vec{\Lambda}) | \Psi_0 \rangle . \tag{C.9}$$

□

C.1.1 Error from the finite time step

In this section, the errors due to the finite time step is analysed, which is quantified by Δ_2. According to Eq. (5.12), to suppress the effect of the algorithmic error to $\varepsilon_{\text{total}}$ with $\Delta^{(\max)} T = \varepsilon_{\text{total}}/2$ and $\Delta_2^{(\max)} \sqrt{\delta t} T = \varepsilon_{\text{total}}/2$, we need to ensure $\Delta^{(\max)} \leq \varepsilon_{\text{total}}/(2T)$, and also set the time step $\delta t \leq \varepsilon_{\text{total}}^2/(4\Delta_2^{2(\max)} T^2)$. Hence we require the number of steps to be $N = T/\delta t \geq 4\Delta_2^{2(\max)} T^3/\varepsilon_{\text{total}}^2$.

Proposition 20. *Δ_2^2 is upper bounded as $\Delta_2^2 \leq \|\vec{\lambda}\|_1^2 \|H\| + \|\vec{\lambda}\|_1 \|H\|^2 + \|\vec{\lambda}\|_1^3$.*

Proof. From Eq. (C.4), the error of the finite time step, Δ_2, can be expanded as

$$\begin{aligned}
\Delta_2^2 =& 2\,\mathrm{Re}\left(|\psi_1\rangle \langle \psi_2|\right) \\
=& -\mathrm{Re}\left(\langle \Psi_0 | \left(i G^\dagger(\vec{O}, \vec{\Lambda}) H - \sum_j \lambda_j \partial_j G^\dagger(\vec{O}, \vec{\Lambda}) \right) \left(H^2 + \sum_{jj'} \lambda_j \lambda_{j'} \partial_{jj'}^2 \right) G(\vec{O}, \vec{\Lambda}) \, |\Psi_0\rangle \right) \\
=& \,\mathrm{Im}\left(\langle \Psi_0 | \, G^\dagger(\vec{O}, \vec{\Lambda}) H \sum_{jj'} \lambda_j \lambda_{j'} \partial_{jj'}^2 G(\vec{O}, \vec{\Lambda}) \, |\Psi_0\rangle \right) + \mathrm{Re}\left(\langle \Psi_0 | \, G^\dagger(\vec{O}, \vec{\Lambda}) H^2 \sum_j \lambda_j \partial_j G(\vec{O}, \vec{\Lambda}) \, |\Psi_0\rangle \right) \\
&+ \mathrm{Re}\left(\langle \Psi_0 | \sum_j \lambda_j \partial_j G^\dagger(\vec{O}, \vec{\Lambda}) \sum_{j'j''} \lambda_{j'} \lambda_{j''} \partial_{j'j''}^2 G(\vec{O}, \vec{\Lambda}) \, |\Psi_0\rangle \right).
\end{aligned} \tag{C.10}$$

In the third equality, we use the fact that $\langle \psi | \, H^3 \, |\psi\rangle$ is always real.

In our special case where $G(\vec{O}, \vec{\Lambda})$ is a product formula consisting of e^{-iPt} operators, we have

$$\|\partial_j G(\vec{O}, \vec{\Lambda}) \, |\Psi_0\rangle \, \| = 1. \tag{C.11}$$

Thus, we can simplify Δ_2 as

$$\Delta_2^2 \leq \|\vec{\lambda}\|_1^2 \|H\| + \|\vec{\lambda}\|_1 \|H\|^2 + \|\vec{\lambda}\|_1^3, \tag{C.12}$$

where the norm $\|\cdot\|$ is the operator norm induced by the norm of quantum states, so that

$$\|H\| = \sup_{|\psi\rangle} \|H\,|\psi\rangle\,\|, \tag{C.13}$$

and the 1-norm of the vector is defined as

$$\|\vec{\lambda}\|_1 = \sum_j |\lambda_j|. \tag{C.14}$$

□

It is also worth noting that $\|H\| \leq \|H\|_F = \sqrt{\sum_j a_j^2}$ given that $H = \sum_j a_j P_j$, where $\|\cdot\|_F$ is the Frobenius norm. We can thus estimate the appropriate time step length during the run of the algorithm provided the parameters $\vec{\lambda}$.

Appendix D

Perturbative quantum simulation

D.1 Equivalence between the two Monte Carlo methods

In this section, I prove the equivalence between the stochastic approach and the discrete time Monte Carlo approach with $\delta t \to 0$. Following the above discussion, we can regard the discrete time Monte Carlo approach as a decaying process. Specifically, at each time step, it has probability $p_{\geq 1}$ to apply an additional operation $\tilde{\Phi}$. Starting at time $t = 0$ with the limit of $\delta t \to 0$, the probability that there is no 'decay' event until time t is

$$Q(t) = \lim_{\delta t \to 0} \prod_{i=0}^{t/\delta t} \left(1 - \sum_{k \geq 1} \tilde{p}_k \delta t\right) = e^{-t\alpha}, \tag{D.1}$$

where $\alpha = \sum_k |\alpha_k|$. The probability to have a decay event in the time interval $[t, t + dt]$ is

$$P(t)dt = \alpha e^{-t\alpha} dt. \tag{D.2}$$

For the stochastic method, we generate a uniformly distributed random variable $q \in [0, 1]$ and solve

$$q = e^{-t_{\text{jp}}\alpha}, \tag{D.3}$$

to determine the jump time t_{jp}. Then the probability that jump happens at time t_{jp} or in particular between $[t_{\text{jp}}, t_{\text{jp}} + dt]$ is

$$|dq| = \alpha e^{-t_{\text{jp}}\alpha} dt = P(t_{\text{jp}})dt, \tag{D.4}$$

which agrees with Eq. (D.2). We can thus use the uniformly distributed random variable q to determine the jump time to equivalently simulate the discrete time Monte Carlo approach.

Now, at the jump time t_{jp}, we apply the quantum operations other than the identity operation. We can determine the quantum operation by generating another uniformly distributed random number $q' \in [0, 1]$. If $q' \in [s_{k-1}, s_k]$, we set the quantum operation to $\mathcal{B}_k$, where $s_k(t) = (\sum_{j=1}^{k} \tilde{p}_j)/(\sum_j^{N_{\text{op}}} \tilde{p}_j)$ and N_{op} is the number of the quantum operations during the evolution.

D.2 Expansion by Dyson series

I show in this section that the above explicit expansion method could be reformulated via the Dyson series expansion.

D.2.1 Method description

I first introduce the general method. Consider the time evolution with Hamiltonian $H = H^{\text{loc}} + V^{\text{int}}$,

$$\left[H^{\text{loc}} + V^{\text{int}}\right] |\psi(t)\rangle = i\frac{\partial|\psi(t)\rangle}{\partial t}, \tag{D.5}$$

with $H^{\text{loc}} = \sum_l H_l$ and $V^{\text{int}} = \sum_j \lambda_j V_j^{\text{int}} = \sum_j \lambda_j \prod_l V_{l,j}^{\text{int}}$. It becomes

$$\lambda e^{iH^{\text{loc}}(t-t_0)} V^{\text{int}} e^{-iH^{\text{loc}}(t-t_0)} |\psi_I(t)\rangle = i\frac{\partial |\psi_I(t)\rangle}{\partial t} \tag{D.6}$$

under the interaction picture with

$$|\psi(t)\rangle = e^{-iH^{\text{loc}}(t-t_0)} |\psi_I(t)\rangle . \tag{D.7}$$

A solution with Dyson series is

$$\begin{aligned}|\psi_I(t)\rangle =& \Bigg[1 - i\int_{t_0}^{t} dt_1 e^{iH^{\text{loc}}(t_1-t_0)} V^{\text{int}} e^{-iH^{\text{loc}}(t_1-t_0)} \\ &- \int_{t_0}^{t} dt_1 \int_{t_0}^{t_1} dt_2 e^{iH^{\text{loc}}(t_1-t_0)} V^{\text{int}} e^{-iH^{\text{loc}}(t_1-t_0)} e^{iH^{\text{loc}}(t_2-t_0)} V^{\text{int}} e^{-iH^{\text{loc}}(t_2-t_0)} + \dots\Bigg] |\psi(t_0)\rangle\end{aligned} \tag{D.8}$$

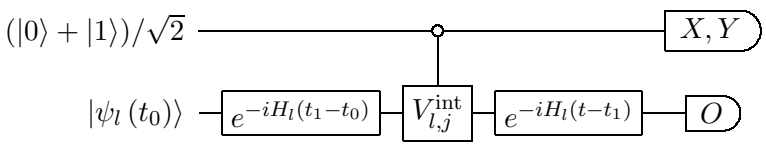
To measure any observable $O = \bigotimes_l O_l$ with $O_I = e^{iH^{\text{loc}}(t-t_0)} O e^{-iH^{\text{loc}}(t-t_0)}$, we have

$$\begin{aligned}\langle\psi_I(t)|O_I|\psi_I(t)\rangle =& \langle\psi(t_0)| e^{iH^{\text{loc}}(t-t_0)} O e^{-iH^{\text{loc}}(t-t_0)} |\psi(t_0)\rangle , \\ &- 2(t-t_0)\int_{t_0}^{t} \frac{dt_1}{t-t_0} \Re\Big[i \langle\psi(t_0)| e^{iH^{\text{loc}}(t-t_0)} O e^{-iH^{\text{loc}}(t-t_1)} V^{\text{int}} e^{-iH^{\text{loc}}(t_1-t_0)} |\psi(t_0)\rangle\Big],\end{aligned} \tag{D.9}$$

up to the first order expansion. Suppose $|\psi(t_0)\rangle = |\psi_1(t_0)\rangle \dots |\psi_L(t_0)\rangle$, the first term is

$$\langle\psi(t_0)| e^{iH^{\text{loc}}(t-t_0)} O e^{-iH^{\text{loc}}(t-t_0)} |\psi(t_0)\rangle = \prod_l \langle\psi_l(t_0)| e^{iH_l(t-t_0)} O e^{-iH_l(t-t_0)} |\psi_l(t_0)\rangle , \tag{D.10}$$

where each term can be easily measured by evolving each subsystem state with the Hamiltonian H_l and measure O_l. To measure the second term, we can uniformly sample t_1 from t_0 to t and use the following circuit



to get $\langle O\rangle_{l,j,t_1} = \langle\psi_l(t_0)|e^{iH_l(t-t_0)}O_l e^{-iH_l(t-t_1)}V^{\text{int}}_{l,j}e^{-iH_l(t_1-t_0)}|\psi(t_0)\rangle$. Then we have

$$\langle\psi(t_0)|e^{iH^{\text{loc}}(t-t_0)}Oe^{-iH^{\text{loc}}(t-t_1)}V^{\text{int}}e^{-iH^{\text{loc}}(t_1-t_0)}|\psi(t_0)\rangle = \sum_j \lambda_j \langle O\rangle_{1,j,t_1}\langle O\rangle_{2,j,t_1}\cdots\langle O\rangle_{L,j,t_1}. \tag{D.11}$$

We can also randomly measure the first term or the second term, as well as each term of the above summation. The cost is now

$$C_1 = 1 + 2T\lambda, \tag{D.12}$$

with $\lambda = \sum_j |\lambda|$. Due to the first order expansion, the approximation error is

$$\varepsilon_1 = \mathcal{O}\left(e^{|V^{\text{int}}|T}(|V^{\text{int}}|T)^2\right). \tag{D.13}$$

We can similarly consider expansion to the kth order, then the cost and the expansion error are

$$\begin{aligned} C_k &= \sum_{n=0}^{k}(2T\lambda)^n/n!, \\ \varepsilon_k &= \mathcal{O}\left(e^{|V^{\text{int}}|T}(|V^{\text{int}}|T)^{k+1}/(k+1)!\right). \end{aligned} \tag{D.14}$$

With the limit of $k\to\infty$, we have

$$\begin{aligned} C_\infty &= e^{2T\lambda}, \\ \varepsilon_\infty &= 0. \end{aligned} \tag{D.15}$$

In this case, note that the cost is the same as the one for the explicit expansion. In the next subsection, I show that they are actually equivalent. Several posted work has proposed the quantum simulation with truncated Dyson series on a universal quantum computer in Refs. [351, 352].

D.2.2 Relation of perturbative expansion and Dyson series

The algorithm using Dyson series implements each expanded term (trajectory) with a quantum computer and sums over the expansion via the average of different trajectories. This is very similar to the above perturbative quantum simulation method. I prove that they are actually equivalent.

Theorem 10. *The infinite-order Dyson series method is equivalent to the perturbative quantum simulation method with the explicit decomposition.*

Proof. To see the equivalence, we first consider a pure state formalism for the perturbative quantum simulation method with the explicit decomposition. Suppose the interaction term $\mathcal{V}^{\text{int}}(\delta t)$ is decomposed as follows,

$$\mathcal{V}^{\text{int}}(\delta t)[\rho] = \mathcal{I}(\rho) - i\delta t\sum_j \lambda_j(V_j^{\text{int}}\rho - \rho V_j^{\text{int}}) + O(\delta t^2). \tag{D.16}$$

Suppose ρ is a pure state $\rho = |\psi\rangle\langle\psi|$, then

$$\begin{aligned}\mathcal{V}^{\text{int}}(\delta t)[|\psi\rangle\langle\psi|] =& \mathcal{I}(\rho) - i\delta t \sum_j \lambda_j (V_j^{\text{int}} |\psi\rangle\langle\psi| - |\psi\rangle\langle\psi| V_j^{\text{int}}) + O(\delta t^2), \\ =& \left(I - i\delta t \sum_j \lambda_j V_j^{\text{int}}\right) |\psi\rangle\langle\psi| \left(I + i\delta t \sum_j \lambda_j V_j^{\text{int}}\right) + O(\delta t^2).\end{aligned} \tag{D.17}$$

Then the whole time evolution with a pure input state $|\psi(t_0)\rangle$ is

$$U(T)|\psi(t_0)\rangle = \left[e^{-iH^{\text{loc}}\delta t}\left(I - i\delta t V^{\text{int}}\right)\right]^{T/\delta t} |\psi(t_0)\rangle + O(\delta t^2), \tag{D.18}$$

and we have

$$\mathcal{U}(T)[\psi(t_0)] = U(T)|\psi(t_0)\rangle\langle\psi(t_0)| U(T)^\dagger. \tag{D.19}$$

Then each expanded term in $\mathcal{U}(T)[\psi(t_0)]$ corresponds to the expanded terms in $U(T)|\psi(t_0)\rangle$ and $\langle\psi(t_0)| U(T)^\dagger$. Now we expand the product of Eq. (D.18) and group the terms according to the number of V^{int} as

$$\begin{aligned}U(T)|\psi(t_0)\rangle =& \Bigg[e^{-iH^{\text{loc}}T} - i\delta t \sum_{i=1}^{T/\delta t} e^{-iH^{\text{loc}}(T-i\delta t)} V^{\text{int}} e^{-iH^{\text{loc}} i\delta t}, \\ & - \delta t^2 \sum_{i_1 \geq i_2 = 1}^{T/\delta t} e^{-iH^{\text{loc}}(T-i_1\delta t)} V^{\text{int}} e^{-iH^{\text{loc}}(i_1-i_2)\delta t} V^{\text{int}} e^{-iH^{\text{loc}} i_2 \delta t}\Bigg] |\psi(t_0)\rangle + O(\delta t^2).\end{aligned} \tag{D.20}$$

Multiplying $e^{iH^{\text{loc}}T}$ and taking the limit of $\delta t \to 0$ we have

$$\begin{aligned}\lim_{\delta t \to 0} e^{iH^{\text{loc}}T} U(T)|\psi(t_0)\rangle =& \Bigg[1 - i\int_{t_0}^{T} dt_1 e^{iH^{\text{loc}}(t_1-t_0)} V^{\text{int}} e^{-iH^{\text{loc}}(t_1-t_0)} \\ & - \int_{t_0}^{T} dt_1 \int_{t_0}^{t_1} dt_2 e^{iH^{\text{loc}}(t_1-t_0)} V^{\text{int}} e^{-iH^{\text{loc}}(t_1-t_0)} e^{iH^{\text{loc}}(t_2-t_0)} V^{\text{int}} e^{-iH^{\text{loc}}(t_2-t_0)} \ldots\Bigg] |\psi(t_0)\rangle,\end{aligned} \tag{D.21}$$

which coincides with the Dyson series expansion. □

Remark that the expansion is universal and avoids the computational cost in diagrammatic perturbation theory. The algorithm with explicit decomposition in Eq. (6.35) effectively implements each expanded term (trajectory) and realise the joint time evolution by summing over the expansion via the average of different trajectories.

D.3 Algorithm

The detailed algorithm of PQS is shown in Algorithm 5.

Algorithm 5 Perturbative quantum simulation. Input: initial state $\rho(0) = \bigotimes_l \rho_l(0)$, number of samples N_s, local evolution $\mathcal{U}_l$, decomposition of the interaction $\mathcal{V}^{\text{int}}(\delta t) = \mathcal{I} + \delta t \sum_k \alpha_k \Phi_{1,k} \otimes \Phi_{2,k} \otimes \cdots \otimes \Phi_{L,k}$ with quantum operations $\Phi_{l,j}$, measurement $O = \bigotimes_l O_l$. Output: $\bar{O}$.

Get C, $\{\alpha_j\}$, $\{\tilde{p}_j = |\alpha_k|\}$, and $\theta_i = -i \ln(\alpha_k/|\alpha_k|)$ from interaction channel $\mathcal{V}$, set $\left\{ s_j = \frac{\sum_{i=1}^{j} \tilde{p}_i}{\sum_i \tilde{p}_i} \right\}$ and $\Gamma(t) = t \sum_k \tilde{p}_k$.
for $m = 1$ to N_s **do**
    Randomly generate $q_0 \in [0, 1]$, set $t = 0$, $n = 0$, $\theta = 0$.
    while $t \leq T$ **do**
        Get t_{jp}^n by solving $\exp\left(-\Gamma(t_{\text{jp}}^n)\right) = q_n$.
        Randomly generate $q_n' \in [0, 1]$.
        Set $j_n = j$ if $q_n' \in [s_{j-1}, s_j]$ and update $\theta = \theta_{j_n} + \theta$.
        Update $t = t + t_{\text{jp}}^n$ and $n = n + 1$.
    end while
    for $l = 1$ to L **do**
        Set $\rho_l = \rho_l(0)$ and $\bar{O} = 0$.
        for $k = 0 : n - 1$ **do**
            Evolve ρ_l under $\mathcal{U}_l$ for time t_{jp}^k and apply Φ_{l,j_k}.
        end for
        Evolve ρ_l under $\mathcal{U}_l$ for time $T - \sum_{k=0}^{n-1} t_{\text{jp}}^k$.
        Measure O of ρ_l to get $O_{l,m}$.
    end for
    Update $\bar{O} = \bar{O} + Ce^{i\theta} \prod_l O_{l,m}/N_s$
end for

D.4 Cost analysis

In this section, I provide the proof of the low bound of the cost in Section 6.2.3.

Proof or Theorem 4. Given the above decomposition, the Choi state of $\mathcal{V}^{\text{int}}(\delta t)$ is

$$\phi_{1,1',\ldots,L,L'}^{\text{int}} = \sum_k \tilde{\alpha}_k \Phi_{1,k}(\phi_{1,1'}) \otimes \Phi_{2,k}(\phi_{2,2'}) \otimes \cdots \otimes \Phi_{L,k}(\phi_{L,L'}). \tag{D.22}$$

Taking $\psi_{1,1'}^{\text{int}}$ as an example, we have

$$\psi_{1,1'}^{\text{int}} = \sum_k \tilde{\alpha}_k \left|\Phi_{1,k}(\phi_{1,1'})\right\rangle \otimes \left\langle \Phi_{2,k}(\phi_{2,2'})^*\right| \otimes \cdots \otimes \left\langle \Phi_{L,k}(\phi_{L,L'})^*\right|, \tag{D.23}$$

where $\left|\Phi_{L,k}(\phi_{l,l'})\right\rangle = S_{l,l'}(\Phi_{L,k}(\phi_{l,l'}))$ and $\left\langle \Phi_{L,k}(\phi_{l,l'})^*\right| = T_{l,l'}(\Phi_{L,k}(\phi_{l,l'}))$. Based on the triangle inequality of the trace norm, we have

$$\|\psi_{1,1'}^{\text{int}}\|_1 \leq \sum_k |\tilde{\alpha}_k| \left\| \left|\Phi_{1,k}(\phi_{1,1'})\right\rangle \otimes \left\langle \Phi_{2,k}(\phi_{2,2'})^*\right| \otimes \cdots \otimes \left\langle \Phi_{L,k}(\phi_{L,L'})^*\right| \right\|_1, \tag{D.24}$$

The trace norm of each term is

$$\begin{aligned} &\left\| \left|\Phi_{1,k}(\phi_{1,1'})\right\rangle \otimes \left\langle \Phi_{2,k}(\phi_{2,2'})^*\right| \otimes \cdots \otimes \left\langle \Phi_{L,k}(\phi_{L,L'})^*\right| \right\|_1 \\ =&\text{Tr}\left[\sqrt{\left\langle \Phi_{1,k}(\phi_{1,1'}) | \Phi_{1,k}(\phi_{1,1'})\right\rangle \left\langle \Phi_{2,k}(\phi_{2,2'})^* | \Phi_{2,k}(\phi_{2,2'})^*\right\rangle \cdots \left\langle \Phi_{L,k}(\phi_{L,L'})^* | \Phi_{L,k}(\phi_{L,L'})^*\right\rangle} \right]. \end{aligned}$$

Note that

$$\langle \Phi_{l,k}(\phi_{l,l'})|\Phi_{l,k}(\phi_{l,l'})\rangle = \langle \Phi_{l,k}(\phi_{l,l'})^*|\Phi_{l,k}(\phi_{l,l'})^*\rangle = \mathrm{Tr}[\Phi_{l,k}(\phi_{l,l'})^\dagger \Phi_{l,k}(\phi_{l,l'})]. \tag{D.25}$$

Here we used the norms of S and T. Based on the property of generalised quantum operations $\Phi_{l,k}$, we have

$$\mathrm{Tr}[\Phi_{l,k}(\phi_{l,l'})^\dagger \Phi_{l,k}(\phi_{l,l'})] = \|\Phi_{l,k}(\phi_{l,l'})]\|_2^2 \leq= 1. \tag{D.26}$$

Combining the above results, we have

$$\|\psi_{1,1'}^{\mathrm{int}}\|_1 \leq \sum_k |\tilde{\alpha}_k| = \tilde{\alpha}. \tag{D.27}$$

Since the inequality holds for any $\psi_{l,l'}^{\mathrm{int}}$, we have

$$\tilde{\alpha} \geq \max_l \|\psi_{l,l'}^{\mathrm{int}}\|_1, \tag{D.28}$$

which completes the proof.

□

D.5 Higher-order moments analysis

In the above discussion, I showed how to use the perturbative quantum simulation (PQS) method to get linear observable measurements. Here, I show that the PQS method applies to measurements of higher-order moments. We take the subsystem purity as an example, while the result applies to general measurements. Without loss of generality, we consider the purity $\mathrm{Tr}[\rho_1(T)^2]$ of the first subsystem, and we denote the set without the first system as $\mathcal{S} = \{2, .., L\}$. Following the PQS method with the explicit expansion in Eq. (6.37), we have

$$\mathcal{U}(T)\left[\bigotimes_l \rho_l\right] = C\sum_{\mathbf{k}} e^{i\theta_{\mathbf{k}}} p_{\mathbf{k}} \bigotimes_l \left[\tilde{U}_{l,k_{T/\delta t}} U_l(\delta t) \ldots \tilde{U}_{l,k_1} U_l(\delta t)\rho_l U_l^\dagger(\delta t)\tilde{V}_{l,k_1} \ldots U_l^\dagger(\delta t)\tilde{V}_{l,k_{T/\delta t}}\right], \tag{D.29}$$

where the input state is $\bigotimes_l \rho_l$, $\tilde{U}_{l,k}$ and $\tilde{V}_{l,k}$ are either I or $V_{l,j}^{\mathrm{int}}$. Now we calculate the reduced density matrix of the first subsystem,

$$\begin{aligned}\rho_1(T) &= \mathrm{Tr}_{\mathcal{S}}[\mathcal{U}(T)\left[\bigotimes_l \rho_l\right]], \\ &= \sum_{\mathbf{k}} \beta_{\mathbf{k}} \tilde{U}_{1,k_{T/\delta t}} U_1(\delta t) \ldots \tilde{U}_{1,k_1} U_1(\delta t)\rho_1 U_1^\dagger(\delta t)\tilde{V}_{1,k_1} \ldots U_1^\dagger(\delta t)\tilde{V}_{1,k_{T/\delta t}},\end{aligned} \tag{D.30}$$

where

$$\beta_{\mathbf{k}} = Ce^{i\theta_{\mathbf{k}}} p_{\mathbf{k}} \prod_{l\in\mathcal{S}} \mathrm{Tr}\left[\tilde{U}_{l,k_{T/\delta t}} U_l(\delta t) \ldots \tilde{U}_{l,k_1} U_l(\delta t)\rho_l U_l^\dagger(\delta t)\tilde{V}_{l,k_1} \ldots U_l^\dagger(\delta t)\tilde{V}_{l,k_{T/\delta t}}\right], \tag{D.31}$$

which could be measured for each $\mathbf{k}$. For the purity of the first subsystem, we have

Therefore, we have

$$\rho_1^2(T) = \sum_{\mathbf{k},\mathbf{k}'} \beta_{\mathbf{k}}\beta_{\mathbf{k}'} \tilde{U}_{1,k_{T/\delta t}} U_1(\delta t) \dots \tilde{U}_{1,k_1} U_1(\delta t) \rho_1 U_1^\dagger(\delta t) \tilde{V}_{1,k_1} \dots U_1^\dagger(\delta t) \tilde{V}_{1,k_{T/\delta t}} \\ \cdot \tilde{U}_{1,k'_{T/\delta t}} U_1(\delta t) \dots \tilde{U}_{1,k'_1} U_1(\delta t) \rho_1 U_1^\dagger(\delta t) \tilde{V}_{1,k'_1} \dots U_1^\dagger(\delta t) \tilde{V}_{1,k'_{T/\delta t}}. \tag{D.32}$$

Suppose the initial state is pure $\rho_1 = |\psi_1\rangle \langle\psi_1|$, we have

$$\mathrm{Tr}[\rho_1^2(T)] = \sum_{\mathbf{k},\mathbf{k}'} \beta_{\mathbf{k}}\beta_{\mathbf{k}'} \langle\psi_1| U_1^\dagger(\delta t) \tilde{V}_{1,k_1} \dots U_1^\dagger(\delta t) \tilde{V}_{1,k_{T/\delta t}} \tilde{U}_{1,k'_{T/\delta t}} U_1(\delta t) \dots \tilde{U}_{1,k'_1} U_1(\delta t) |\psi_1\rangle \\ \times \langle\psi_1| U_1^\dagger(\delta t) \tilde{V}_{1,k'_1} \dots U_1^\dagger(\delta t) \tilde{V}_{1,k'_{T/\delta t}} \tilde{U}_{1,k_{T/\delta t}} U_1(\delta t) \dots \tilde{U}_{1,k_1} U_1(\delta t) |\psi_1\rangle . \tag{D.33}$$

We note that the two overlap terms could be evaluated with circuits that are similar to the ones used for measuring linear observables. In practice, we can use the Monte Carlo method to estimate the purity. The sample complexity for the purity estimation is related to $\sum_{\mathbf{k},\mathbf{k}'} |\beta_{\mathbf{k}}\beta_{\mathbf{k}'}| \propto C^2$.

Other higher order moments can be derived similarly, and I leave it to the dedicated readers.

Appendix E

Response, neutron spectroscopy, and quantum simulation

E.1 Response

In the discussion in Chapter 7, it is assumed that the perturbation is weak, and hence only linear response is considered. This section concerns high-order response under external drive. Denote the original unperturbed Hamiltonian as $\hat{H}_0$, and the perturbation as $\hat{V}$, and the Hamiltonian reads

$$\hat{H} = \hat{H}_0 + \hat{V}. \tag{E.1}$$

In the Heisenberg picture, the operators are time dependent. The source of perturbation can be mathematically written as

$$\hat{V}(t) = \sum_i \hat{O}_i(t) f_i(t), \tag{E.2}$$

where the observable may be local as $\hat{O}(\mathbf{x}, t)$. In engineered spectroscopy, the ground state of the original Hamiltonian is initially prepared. We then apply a perturbation $\hat{V}$ at time $t = 0$, and measure the observable $\hat{O}$ at time $t > 0$. The perturbation drives the system out of equilibrium, and the system will show a dynamic response to the perturbation applied.

In the interaction picture, the time-evolved state is expressed as

$$\rho_I(t) = U_I(t)\rho_{0,I}(t)U_I^\dagger(t), \tag{E.3}$$

with $U(t) = U(t; t_0 \to \infty)$ and

$$U_I(t; t_0) := \mathcal{T} \exp(-i \int_{t_0}^{t} \hat{V}(t_1) dt_1). \tag{E.4}$$

The expectation of observable can be expanded through the generalised susceptibility as

$$
\begin{aligned}
\langle \hat{O}_i(t)\rangle =& \operatorname{Tr} \rho_I(t)\hat{O}_i(t) \\
=& \operatorname{Tr} \rho_{0,I}(t)U_I^\dagger(t)\hat{O}_i(t)U_I(t) \\
=& \langle \hat{O}_i(t)\rangle|_{f=0} - i\int_{-\infty}^{t} dt_1 \langle[\hat{O}_i(t),\hat{V}(t_1)]\rangle + \int_{-\infty}^{t} dt_1 dt_2 \langle[[\hat{O}_i(t),\hat{V}(t_1)],\hat{V}(t_2)]\rangle + \dots. \\
=& \langle \hat{O}_i(t)\rangle|_{f=0} - i\int_{-\infty}^{t} dt_1 \sum_j \langle[\hat{O}_i(t),\hat{O}_j(t_1)]\rangle f_j(t_1) \\
& - \int_{-\infty}^{t} dt_1 dt_2 \sum_{jk} \langle[[\hat{O}_i(t),\hat{O}_j(t_1)],\hat{O}_k(t_2)]\rangle f_j(t_1) f_k(t_2) + \dots. \\
:=& \langle \hat{O}_i(t)\rangle|_{f=0} + \int_{-\infty}^{\infty} dt_1 \sum_j \chi_{ij}^{(1)}(t;t_1) f_j(t_1) + \int_{-\infty}^{\infty} dt_1 \int_{-\infty}^{\infty} dt_2 \sum_{jk} \chi_{ijk}^{(2)}(t;t_1,t_2) f_j(t_1) f_k(t_2) + \\
& + \int_{-\infty}^{\infty} dt_1 \int_{-\infty}^{\infty} dt_2 \int_{-\infty}^{\infty} dt_3 \sum_{jkl} \chi_{ijkl}^{(3)}(t;t_1,t_2,t_3) f_j(t_1) f_k(t_2) f_l(t_3) + \dots
\end{aligned}
\tag{E.5}
$$

Here we have defined

$$
\chi_{ij}^{(1)}(t;t_1) = -i\theta(t-t_1)\langle[\hat{O}_i(t),\hat{O}_j(t_1)], \quad \chi_{ijk}^{(2)}(t;t_1,t_2) = \theta(t-t_1)\theta(t-t_2)\langle[[\hat{O}_i(t),\hat{O}_j(t_1)],\hat{O}_k(t_2)]\rangle. \tag{E.6}
$$

The nested correlation function can be extracted from the dynamics of the observables.

In the weak perturbative regime, where linear response holds, we have

$$
\begin{aligned}
\langle \hat{O}_i(t)\rangle =& \operatorname{Tr} \rho_0 U^{-1}(t)\hat{O}_i(t)U(t) \\
=& \langle \hat{O}_i(t)\rangle|_{f=0} - i\int_{-\infty}^{t} dt_1 \langle[\hat{O}_i(t),\hat{V}(t_1)]\rangle + \mathcal{O}(\|\hat{V}\|^2).
\end{aligned}
\tag{E.7}
$$

Suppose that we have a delta-like pulse as $\hat{V}(t) = -\hat{O}_j\delta(t)$, and then the resulting change of observables is

$$
\begin{aligned}
\delta\langle \hat{O}_i(t)\rangle =& -i\int_{-\infty}^{t} dt_1 \langle[\hat{O}_i(t),\hat{V}(t_1)]\rangle + \mathcal{O}(\|\hat{V}\|^2) \\
=& \, i\langle[\hat{O}_i(t),\hat{O}_j]\rangle + \mathcal{O}(\|\hat{V}\|^2).
\end{aligned}
\tag{E.8}
$$

This indicates that we can infer the unequal-time correlator $[\hat{O}_i(t),\hat{O}_j]$ by directly measuring the local observables. We have

$$
\langle[\hat{O}_i(t),\hat{O}_j]\rangle = \sum_{\mu,\nu} \langle 0|\mu\rangle \langle\mu|\hat{O}_i|\nu\rangle \langle\nu|\hat{O}_j|0\rangle \, e^{i(E_\mu - E_\nu)t}, \tag{E.9}
$$

from which energy differences can be obtained through spectroscopic analysis, which is introduced in Chapter 7. Note that in equilibrium, the density matrix of the quantum state commutes with the Hamiltonian, $[\rho, H] = 0$.

Non-linear response can be used to distinguish some phenomena, such as strong quenched disorder and fractionalisation. Both mechanisms broaden the excitation spectrum and cannot be distinguished from linear response spectroscopy [353].

E.2 Relation to the projection operation

In this section, I show how the spectroscopy methods introduced in Chapter 7 relate to algorithmic cooling introduced in Chapter 4. Recall that we have introduced the function $G(t)$, which can be equivalently written in the Schrödinger picture as

$$G(t) = \langle\psi_0|e^{iHt}\hat{O}e^{-iHt}|\psi_0\rangle\,. \tag{E.10}$$

This definition can be extended as

$$G(t,t') = \langle\psi_0|e^{iHt}\hat{O}e^{-iHt'}|\psi_0\rangle\,. \tag{E.11}$$

Let us define a weighted Fourier transform of $G(t,t')$ as

$$G(\omega,\omega') = \frac{c^2}{(2\pi)^2}\int_{-\infty}^{\infty} G(\tau t,\tau t')g(t)g(t')e^{i\theta_t}e^{-i\theta_{t'}}e^{-i\tau\omega t}e^{i\tau\omega' t'}dtdt'. \tag{E.12}$$

We have

$$G(\omega,\omega') = \sum_{n,=0}\mathrm{Tr}[\hat{O}\,|n\rangle\,\langle n|\psi_0\rangle\,\langle\psi_0|n'\rangle\,\langle n'|\,p(\tau(E_n-\omega))p(\tau(E_{n'}+\omega'))], \tag{E.13}$$

with the function $p(\tau\omega)$ is the dual Fourier transform of $g(t)$ related by $p(\tau\omega) = \frac{c}{2\pi}\int g(t)e^{i\theta_t}e^{-i\tau\omega t}dt$.

Now the energies of $|n\rangle$ and $|n'\rangle$, which are originally connected by the energy selection rule, are decoupled, and thus one can directly evaluate the energies instead of the energy gaps by tuning the parameters ω,ω': A special case is when we consider $\hat{O} = I$ and $t' = 0$, where

$$\begin{aligned} G(\omega,0) &= \frac{c}{2\pi}\int_{-\infty}^{\infty} G(t,0)e^{-i\tau\omega t}dt. \\ &= \frac{c}{2\pi}\int_{-\infty}^{\infty}\langle\psi_0|e^{i\tau Ht}|\psi_0\rangle\,g(t)e^{-i\tau\omega t}dt \\ &= \frac{c}{2\pi}\sum_j |c_j|^2\int_{-\infty}^{\infty} g(t)e^{i\tau(E_j-\omega)t}dt \\ &= \sum_j |c_j|^2 p(\tau(E_j-\omega)). \end{aligned} \tag{E.14}$$

To see why $G(\omega,0)$ could select the eigenvalues, let us consider a matrix function acting on the Hamiltonian as

$$\hat{p}(H) := \sum_{i=0}^{N-1} p(E_i)\,|u_i\rangle\,\langle u_i|\,, \tag{E.15}$$

where $p(h):\mathbb{R}\to\mathbb{C}$ is a generic continuous-variable function determining the transformation of the spectrum of the Hamiltonian. We find that $G(\omega,0)$ effectively realises the projection operator $\hat{p}$ on the initial state as

$$G(\omega,0) = \langle\psi_0|\hat{p}(\tau(H-\omega))|\psi_0\rangle\,. \tag{E.16}$$

For instance, the projection operator could be $\hat{p}(\tau H) = e^{-\tau^2 H^2}$, which projects out the contribution of other eigenstates with an increasing τ. The eigenvalue information could, in addition, be expressed by

$$P(\omega) = \sum_j |c_j|^2 \delta(\omega - E_j), \tag{E.17}$$

and one can find that

$$G(\omega, 0) = (p \star P)(\omega), \tag{E.18}$$

which is because

$$(p \star P)(\omega) = \int_{-\infty}^{\infty} p(t) P(\omega - t) dt = \sum_j |c_j^2| \int_{-\infty}^{\infty} \delta(\omega - E_j - t) dt = \sum_j |c_j|^2 p(E_j - \omega). \tag{E.19}$$

E.3 Learning from neutron scattering spectroscopy

As has been intensively discuss in previous section, scattering spectroscopy is a well-established technique that has been widely used to explore properties of quantum materials. Nevertheless, spectroscopy has inherent limitations; for instance, it can only probe properties of materials in equilibrium phases. In addition, it could be difficult to explain the experimental data produced in spectroscopy experiments due to a lack of effective models, or the computational complexity involved in solving effective models. In this section, I discuss how to learn the effective physics of quantum materials from the measurement outcomes of neutron scattering with the assistance of quantum computing.

E.3.1 Motivation

Understanding quantum many-body phenomena of materials is a key research direction, as has previously been stated in this thesis. However, it is largely limited by the contrast between the finite computational resource and highly demanding complexity in the description and solution of these systems. Neutron scattering provides a powerful tool for probing many-body phenomena, and has achieved great success in the detection of magnetic structure, deepening our understanding in the fields of magnetism and superconductivity. As an example, in magnetic neutron scattering, neutrons interact with spins of electrons, and the collected neutron intensities reflect the magnetic response of electrons in the materials, which contain certain information about the magnetic interaction in the materials under probe. Analysing the scattering data usually requires the construction of a simple effective theory that contains the essential ingredients of the system and makes some assumptions for the interacting models, for instance, static lattice and dipole approximations. By fitting the experimental data, we can come up with a corresponding effective Hamiltonian description of the phenomena of the materials.

However, several issues persist when attempting an accurate description of the materials. First, the fitting of the experimental data relies on the assumptions and the effective model. At a

fundamental level, we require a mean-field description of the model a priori; this usually involves coarse-graining the details of the microscopic interactions, or downfolding the interactions [329], to make it classically computable. We additionally require certain assumptions in order to capture the physics during the scattering process. Indeed, the construction of a universal macroscopic theory generally calls for a large set of parameters; however, experiments can only probe a limited number of observables, thus limiting the flexibility of improving the description of a given system [354]. In the case of a general many-body system, we may consequently fail to accurately describe its essential features. More importantly, simulating the physics with a given model and parameters can be classically hard to achieve. In principle, we may be unable to reproduce the experimental data for a general system described in a quantum mechanical way, even if we know the Hamiltonian. Furthermore, the experimental data may be subjected to non-negligible errors, such as imperfections of samples, or large statistical variances of the intensity from the neutron scattering experiment.

Quantum computing may provide an alternative means through which to address these challenging problems. Quantum computing has long been considered to be a promising solution for solving many-body problems due to its ability to simulate a quantum problem with a quantum system in a quantum mechanical way. In theory, quantum computing renders the possibility to simulate the scattering process in spectroscopy experiments. Given sufficient experimental data, we could learn the effective Hamiltonian by comparing the simulated results of the pre-assumed Hamiltonian and experimental results. We could consequently predict experimental phenomena in other scenarios. Simulated results from controllable quantum computers also offer the possibility of compensating for errors arising from imperfections in experiments, and could. thus potentially provide more accurate results.

In this section, I provide a concrete demonstration of how dynamic correlations and associated magnetic neutron cross-section can be efficiently estimated, and extract the effective physics from neutron spectroscopy data. The tasks here are divided into two parts: simulation of the cross-section that is measurable in neutron scattering, and the learning process. For the first task, I first discuss how the problem can be encoded into a task for quantum computing. This is followed by a discussion on the simulation of the observables that can be detected by neutron scattering. I propose quantum algorithms for quantum spectroscopy simulation, in which several quantum technologies that are compatible to near-term quantum devices are developed to promote its practical implementation.

In the second part, I propose to learn the effective model from experimental data, which is supported by efficient simulation of neutron spectroscopy using quantum computing techniques. One first solves the microscopic dynamics of a system exactly, and computes the the magnetic response. The effective description of the system could be identified by comparing the dynamical response function, and the fitting of experimental results could be demonstrated with the

assistance of a quantum computer. Finally, the reconstruction of the experimental outcomes is demonstrated. It is expected that our protocol could provide some insights in explaining and understanding the experimental results, and provide an additional comparison and solution to spectroscopy.

E.3.2 Methods

The neutron scattering intensity measured at the detector can be written as a convolution of the differential scattering cross-section $\frac{d^2\sigma}{d\Omega' dE'}$ as

$$I(\mathbf{Q}, E) = \int \frac{d^2\sigma}{d\Omega' dE'} R(\mathbf{Q}', E', \mathbf{Q}, E) d\mathbf{Q}' dE' \tag{E.20}$$

where the instrument resolution function $R(\mathbf{Q}', E', \mathbf{Q}, E)$ is mainly determined by parameters related to the instrument set-up, and $\mathbf{Q}$ and E are the wave-vector and energy transfer [355]. When the interaction between a neutron and the sample, V, is weak, standard perturbation theory can be used to show that the partial differential cross-section is given by

$$\frac{d^2\sigma}{d\Omega dE_f} = \frac{k_f}{k_i} \left(\frac{m_n}{2\pi}\right)^2 \sum_{\lambda_i \sigma_i} p_{\lambda_i} \sum_{\lambda_f \sigma_f} |\langle \sigma_f \lambda_f | V(\mathbf{Q}) | \sigma_i \lambda_i \rangle|^2 \, \delta\left(E_{\lambda_f} - E_{\lambda_i} - \omega\right), \tag{E.21}$$

where $V(\mathbf{Q})$ is the Fourier transform of the interaction potential between the neutron and the scattering system, $|\sigma_i \lambda_i\rangle$ represents the incident state before the scattering with energy E_i, momentum k_i and spin polarisation σ_i, $|\sigma_f \lambda_f\rangle$ represents the final state after the scattering with energy E_f, momentum k_f and spin polarisation σ_f, and p_{λ_i} is the thermal distribution.

For magnetic materials, the interaction potential contains the contribution from nuclear interaction and magnetic interaction as

$$V(\mathbf{Q}) = V_N(\mathbf{Q}) + V_M(\mathbf{Q}). \tag{E.22}$$

Here, we mainly focus on pure magnetic scattering with $V_M(\mathbf{Q}) = -\mu_0 \boldsymbol{\mu}_n \cdot \mathbf{M}_\perp(\mathbf{Q})$, $\boldsymbol{\mu}_n = -2\gamma\mu_N \mathbf{s}_n$ and $\mathbf{M}_\perp(\mathbf{Q}) := \hat{\mathbf{Q}} \times \{\mathbf{M}(\mathbf{Q}) \times \hat{\mathbf{Q}}\}$ is the projection of $\mathbf{M}$ onto the plane perpendicular to normalised momentum $\hat{\mathbf{Q}} := \mathbf{Q}/|\mathbf{Q}|$. The cross-section of magnetic scattering has the relation with $S(\mathbf{Q}, \omega)$ as

$$\frac{d^2\sigma}{d\Omega dE_f} \propto \frac{k_f}{k_i} S(\mathbf{Q}, \omega) \tag{E.23}$$

where $S(\mathbf{Q}, \omega)$ is the magnetic response function,

$$S(\mathbf{Q}, \omega) = \sum_{\lambda_i} p_{\lambda_i} \sum_{\lambda_f} |2 \langle \sigma_f | \mathbf{s}_n | \sigma_i \rangle \cdot \langle \lambda_f | \mathbf{M}_\perp(\mathbf{Q}) | \lambda_i \rangle|^2 \, \delta\left(E_{\lambda_f} - E_{\lambda_i} - \omega\right). \tag{E.24}$$

See Eq. (6.6) in Chapter 6 in [15] for details.

As such, we can find that the intensity detected by neutron scattering is closely related to the partial magnetic response function, from which we may be able to detect the magnetic structure

of the sample material and construct an effective model to describe the interesting many-body behaviour. From the experimental data produced by neutron spectroscopy, we can build up an effective model of the magnetic structures of the materials being probed. Essentially, it can be reduced to a learning problem of the magnetic form factor and the effective Hamiltonian, provided the intensity data collected from experiments.

However, it could be challenging to fully learn and compare the effective model in classical methods. The challenges include the following:

1. Certain assumptions have to be made to obtain the form of interactions and hence $\mathbf{M}(\mathbf{Q})$. For instance, under static lattice approximation and dipole approximation with a quenched orbital angular momentum, we have $\mathbf{M}(\mathbf{Q}) \simeq -g\mu_B f(\mathbf{Q})\mathbf{S}$, which is a product of a form factor and a spin operator.

2. The computational complexity of spectroscopy simulation may be considerable. Given a general form of $\mathbf{M}(\mathbf{Q})$, the simulation complexity of neutron spectroscopy scales exponentially with the system size.

3. The experimental data may have large statistical errors, and errors occurring as a result of imperfect samples, and sample mounting[1].

Here, I address these challenges with the assistance of quantum computers. I will first set up a formal description of the problem. To make a connection in a quantum computing setup, let us consider the time-resolved response function. The partial magnetic response functions can be alternatively expressed as the time Fourier transform of a correlation function

$$S_{\alpha\beta}(\mathbf{Q},\omega) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \left\langle M_\alpha^\dagger(\mathbf{Q}) M_\beta(\mathbf{Q},t)\right\rangle \exp(-i\omega t)dt. \tag{E.25}$$

We can see from Eq. (E.25) that the partial response function is the Fourier transform of the dynamic correlation function $G^{\alpha\beta}(\mathbf{Q},t)$ at momentum $\mathbf{Q}$

$$G^{\alpha\beta}(\mathbf{Q},t) = \langle M_\alpha^\dagger(\mathbf{Q}) M_\beta(\mathbf{Q},t)\rangle = \mathrm{Tr}\left[M_\alpha^\dagger(\mathbf{Q}) e^{iHt} M_\beta(\mathbf{Q}) e^{-iHt}\rho\right], \tag{E.26}$$

where we have used the density matrix representation.

The Hamiltonian and the magnetic form factor can be expressed as $H := H(\vec{\theta}_1)$ and $\mathbf{M} := \mathbf{M}(\vec{\theta}_2)$ with parameters $\vec{\theta}_1$ and $\vec{\theta}_2$. From the experimental data, we can derive the measured response function $S_{\alpha\beta}(\mathbf{Q},\omega)$, and hence we can get $\{G^{\alpha\beta}(\mathbf{Q},t)\}$. The problem is now transformed into a learning problem: given a series of $\{G^{\alpha\beta}(\mathbf{Q},t)\}$, how to learn the parameters of the effective model.

Intuitively, we can compare the measured observables, i.e., the dynamic correlation function $G^{\alpha\beta}(\mathbf{Q},t)$, from the neutron scattering with that simulated from the a priori Hamiltonian, which

[1] Here, it is worth pointing out that imperfections of samples can also be regarded as a true property of the sample being studied, and thus are not errors.

we denote as $\bar{G}^{\alpha\beta}(\mathbf{Q}, t)$ and $G^{\alpha\beta}(\mathbf{Q}, t|\theta)$, respectively. Here, $\bar{G}^{\alpha\beta}(\mathbf{Q}, t)$ serves as the reference. Formally, we can minimise the distance between $\bar{G}^{\alpha\beta}(\mathbf{Q}, t)$ and $G^{\alpha\beta}(\mathbf{Q}, t|\theta)$

$$\min_{\theta} |G^{\alpha\beta}(\mathbf{Q}, t|\theta) - \bar{G}^{\alpha\beta}(\mathbf{Q}, t)| \tag{E.27}$$

or the measured spectrum

$$\min_{\theta} |S^{\alpha\beta}(\mathbf{Q}, \omega|\theta) - \bar{S}^{\alpha\beta}(\mathbf{Q}, \omega)|, \tag{E.28}$$

over the parameters.

It is worth noting that the response function in the frequency domain may not be directly simulated on a quantum computer. Instead, we can first simulate the time-resolved G and then Fourier transform it to estimate the response. Since we have to set a cutoff for the maximum evolution time, this will bring about a systematic error in the spectrum, especially in the low energy spectrum. Moreover, there is a limitation of finite size, which may be alleviated by imposing a periodic boundary condition. Quantum simulation of neutron spectroscopy will be discussed in the following section.

E.4 Quantum simulation strategy

E.4.1 Stage setup

To learn the effective model, we must be able to efficiently simulate the dynamic correlation function $G^{\alpha\beta}(\mathbf{Q}, t|\theta)$. Therefore, we first discuss how to simulate the dynamic correlation function $G^{\alpha\beta}(\mathbf{Q}, t)$ provided an effective model. While simulating the neutron scattering spectroscopy may not be efficient for a general effective model, we can leverage quantum computing in this context to make the simulation efficient. The key procedure for the quantum computing can be summarised as follows:

1. Map the target systems of interest to qubits;
2. Prepare the initial state;
3. Evolve the initial state governed by the Hamiltonian for time t;
4. Obtain the time-resolved correlation function by measuring the observable.

This strategy fits within the framework developed in this thesis, and therefore, quantum simulation strategies developed in this thesis naturally applies in this context. I am sure that the reader is now already very familiar with it when coming to this point, so I will not over-elaborate on this specific application. However, I will briefly discuss thermal state preparation, which may also find application in other areas.

To compute the expectation of an observable, the first step is to prepare a thermal state [1], which is QMA-hard in general. There are several proposals to prepare the thermal stats, such

as purification, quantum metropolis sampling, and a minimally entangled typical thermal state (METTS) algorithm [356]. This sections focuses on the METTS and discusses its implementation on a quantum computer. In contrast to purification, we only require operating on a single system as opposed to two copies of systems. More concretely, in the standard purification method, the thermal state can be written as $\rho = \mathrm{Tr}_E \left|\Phi\right\rangle \left\langle\Phi\right|$ with

$$\left|\Phi\right\rangle = \frac{1}{\sqrt{Z}} \left(e^{-\frac{1}{2}\beta H} \otimes \mathbb{I}_E\right) \left|\Phi^+\right\rangle \tag{E.29}$$

where $\left|\Phi^+\right\rangle = 1\sqrt{d}\sum_{AE} \left|i\right\rangle_A \left|i\right\rangle_E$ is the maximal entangled state, A and E represents the system and the environment, respectively. Therefore, we can input a maximally entangled state $\left|\Phi^+\right\rangle$ of the whole system (AE) and evolve the whole system with Hamiltonian $H \otimes \mathbb{I}$ under imaginary-time τ. Then, the state of system A at time $\tau = \beta/2$ will be the thermal state with temperature T. This method requires an ancillary system on which to operate.

The minimally entangled typical thermal state algorithm calculates thermal properties by sampling the state obeying the thermal equilibrium condition. I will first discuss the main algorithm, and then show how to implement the subroutine of the sampling process based on imaginary-time evolution.

The thermal average can be expressed as

$$\langle\hat{O}(t)\rangle = \frac{1}{Z}\,\mathrm{Tr}\left[e^{-\beta H}\hat{O}(t)\right] = \frac{1}{Z}\sum_i \langle i|e^{-\frac{\beta}{2}H}\hat{O}(t)e^{-\frac{\beta}{2}H}|i\rangle \tag{E.30}$$

where $\{|i\rangle\}$ is an orthonormal basis set. Define

$$|\phi_i\rangle = P_i^{-1/2} e^{-\frac{\beta}{2}H}|i\rangle, \tag{E.31}$$

and we have

$$\langle\hat{O}(t)\rangle = \frac{1}{Z}\sum_i P_i\langle\phi_i|\hat{O}(t)|\phi_i\rangle = \frac{1}{Z}\sum_i P_i\langle\phi_i|\mathbf{M}_\alpha e^{iHt}\mathbf{M}_\beta e^{-iHt}|\phi_i\rangle \tag{E.32}$$

where we define $\hat{O}(t) := \mathbf{M}_\alpha e^{iHt}\mathbf{M}_\beta e^{-iHt}$. The summation can be estimated by sampling $|\phi_i\rangle$ with probability P_i/Z and then sum the sampled expectation value $\langle\hat{O}(t)\rangle$ which can be measured by a Hadamard test circuit. The procedure can be summarised as follows.

1. Prepare a product state $|i\rangle$.

2. Prepare $|\phi_i\rangle$ by quantum imaginary-time evolution (QITE) and measure the observables of interest.

3. Collapse the state $|\phi_i\rangle$ to a new state $|i'\rangle$ by quantum measurement.

The central idea of QITE is to approximate the nonunitary operator by a sequence of unitary operators,

$$|\phi_i\rangle = P_I^{-1/2} e^{-\frac{\beta}{2}H} |i\rangle \approx U |i\rangle . \tag{E.33}$$

Various methods can be used for this approximation, such as the variational methods introduced in Section 2.2.3.3 or the QITE methods introduced and developed in [177, 357]. I will elaborate on the QITE method in Section E.4.3. Note that observables of interest are measured on the ancillary qubit, and the state will not collapse after the measurement. We can consequently perform measurements on the state to decide the new initial basis for the next iteration.

E.4.2 Mapping to qubits

The spins of electrons under consideration need to be encoded into qubits. For the case of spin-1/2 particles, this mapping takes a simple form

$$\left|\frac{1}{2}\right\rangle := |0\rangle, \quad \left|-\frac{1}{2}\right\rangle := |1\rangle, \tag{E.34}$$

which leads to a natural mapping of the spin operators to Pauli operators

$$S^\alpha = \frac{1}{2}\sigma^\alpha, \tag{E.35}$$

with $\alpha = x, y, z$. The generalisation to spin-S systems can be realised using qudits or using a symmetric encoding. The latter one is conceptually simple and could be easy to implement, although it requires more qubits for the encoding (see [65] for details).

E.4.3 Thermal state preparation

The subroutine of the METTS relies on the preparation of the state under imaginary-time evolution as $|\phi_i\rangle = P_i^{-1/2} e^{-\frac{\beta}{2}H} |i\rangle$. The following is a brief review of state preparation by quantum imaginary-time evolution. Suppose the Hamiltonian has the decomposition $H = \sum_{l=1}^{L} \hat{h}_l$, where the Hamiltonian contains L local terms and each $\hat{h}_l$ acts on at most k neighbouring sites. The goal here is to simulate quantum states under imaginary-time evolution,

$$|\Psi\rangle = e^{-\beta H} |\Psi_0\rangle \,/\, \langle\Psi_0| e^{-2\beta H} |\Psi_0\rangle \tag{E.36}$$

Using the first-order Trotterization, the evolution operator can be represented as

$$\mathrm{e}^{-\beta H} = \left(\prod_{l=1}^{L} e^{-\delta\tau H} \right)^n + \mathcal{O}(\delta\tau) \tag{E.37}$$

with the total number of steps $n = \frac{\beta}{\delta\tau}$.

The evolved state after applying a nonunitary operator $e^{-\delta\tau \hat{h}_l}$ within imaginary-time $\delta\tau$ is given by

$$|\Psi(\tau + \delta\tau)\rangle = c^{-1/2} e^{-\delta\tau H} |\Psi(\tau)\rangle \approx e^{-i\delta\tau \hat{A}} |\Psi(\tau)\rangle, \tag{E.38}$$

where c is the normalisation factor and $\hat{A}$ is a Hermitian operator that acts on a domain of D qubits around the support of $\hat{h}_l$ as

$$\hat{A} = \sum_{i_1 i_2 \ldots i_D} a_{i_1 i_2 \ldots i_D} \hat{\sigma}_{i_1} \hat{\sigma}_{i_2} \ldots \hat{\sigma}_{i_D} = \sum_I a_I \hat{\sigma}_I, \tag{E.39}$$

where I denotes a set of indices $i_1, i_2, ..., i_D$. The unitary operator $e^{-i\delta\tau\hat{A}}$ can be determined by minimising the approximation error in Eq. (E.38) as

$$\min \|c^{-1/2} e^{-\delta\tau H} |\Psi(\tau)\rangle - e^{-i\delta\tau\hat{A}} |\Psi(\tau)\rangle\|_2. \tag{E.40}$$

Up to $\mathcal{O}(\delta\tau)$, the coefficients a are defined by the linear system

$$S\mathbf{a} = \mathbf{b}, \tag{E.41}$$

where the elements of S and b are expectation values over k qubits, defined by

$$S_{I,J} = \mathrm{Re}(\langle\Psi|\hat{\sigma}_I^\dagger \hat{\sigma}_J|\Psi\rangle), b_I = -c^{-1/2}\,\mathrm{Im}(\langle\Psi|H\hat{\sigma}_I|\Psi\rangle). \tag{E.42}$$

For a local Hamiltonian with nearest-neighbour interactions on a d-dimensional cubic lattice, the domain size D is bounded by $\mathcal{O}(C^d)$, where C is the correlation length. The circuit complexity is shown in the following result. More details about the algorithm complexity can be found in Ref. [177].

Theorem 11 (in [177])**.** *For a d-dimensional system, for every $\varepsilon > 0$, there are unitaries U each acting on $N_q = k(2C)^d \ln^d \left(2\sqrt{2}nL\varepsilon^{-1}\right)$ qubits, such that $\| \, |\Psi_{Ln}\rangle - |\Phi_{Ln}\rangle \, \| \leq \varepsilon$, where $|\Psi_{Ln}\rangle$ represents the exact state while $|\Phi_{Ln}\rangle$ represents the approximated state after the nonunitary operations, and C is the correlation strength.*

Instead of realising the nonunitary operator $e^{-\beta H}$ in the Trotter decomposition one-by-one, we can utilise the representation of the quantum circuit, and train the parameters to effectively realise the imaginary-time evolution. It is assumed that the time-evolved state can be approximated by a parameterised trial state $|\psi(\vec{\theta}(\tau))\rangle$, with variational parameters $\vec{\theta}(\tau)$. As mentioned in [120], by minimising the distance between the ideal evolution and the evolution of the parameterised trial state, the evolution of the target state $|\psi(\tau)\rangle$ under the Schrödinger equation can be mapped to the trial state manifold as the evolution of parameters $\vec{\theta}$. This has been discussed in Section 2.2.

The next step is to realise the time evolution $\hat{O}(t) := \mathbf{M}_\alpha e^{iHt} \mathbf{M}_\beta e^{-iHt}$. Various Hamiltonian simulation algorithms which are introduced in Chapter 2 can be used; for instance, product formulae, and the methods developed in Chapter 5, Chapter 6, and Chapter 9. In addition, variational algorithms introduced in Section 2.2 can also be used.

E.4.4 Low-temperature and high-temperature limit

In the experiment, we are interested in the phases at low-temperature or high-temperature, which correspond to the ground state and the normal state, respectively. For these two phases, we can prepare the initial state more efficiently. The following elaborates and expands on the implementation.

In the high-temperature limit, the density matrix is roughly a maximally mixed state

$$\rho \approx \frac{I}{D}, \tag{E.43}$$

with $D = \mathrm{Tr}[I] = 2^n$ being the system size. In this case, the observable dynamics is thus given by

$$\langle \hat{O}(t) \rangle = \frac{\mathrm{Tr}[\mathbf{M}_\alpha \mathbf{M}_\beta(t)]}{D}. \tag{E.44}$$

This indeed contains some information of the Hamiltonian. However, high-temperature spectroscopy in neutron scattering is usually used as the background, and is subtracted by the low-temperature data.

In the low-temperature limit, the initial state is the ground state, which can be obtained by many means. For instance, we can use the variational approach

$$|\Psi\rangle = U(\theta) |\Psi_0\rangle, \tag{E.45}$$

or by QITE as

$$|\Psi\rangle = \lim_{\beta\to\infty} c^{-1/2} e^{\beta H} |\Psi_0\rangle, \tag{E.46}$$

as extensively discussed in the above section.

E.4.5 Phase correction by sum rules

Some useful relations are imposed by the spectral properties, which can be used to calibrate the results from noisy quantum computers. Here, I have listed some examples of useful relations as a reference for interested readers, and would direct the interested reader to [15] for more details on this topic.

The thermal average of spins evolution

$$\left\langle S_i^2 \right\rangle = S_i \left(S_i + 1\right) = \sum_\alpha G_{ii}^{\alpha\alpha}(0). \tag{E.47}$$

The dynamic part of the response function is

$$\tilde{S}(\mathbf{Q}, \omega) = S(\mathbf{Q}, \omega) - |\langle \mathbf{M}_\perp(\mathbf{Q}) \rangle|^2 \delta(\omega), \tag{E.48}$$

which is related to the generalised susceptibility by the fluctuation-dissipation theorem as

$$\widetilde{S}_{\alpha\beta}(\mathbf{Q}, \omega) = \{1 + n(\omega)\} \frac{1}{\pi} \chi''_{\beta\alpha}(\mathbf{Q}, \omega). \tag{E.49}$$

The general sum rule related to the uniform static susceptibility can be expressed by

$$\widetilde{S}_{\alpha\beta}(0,\omega) = \omega\{1 + n(\omega)\}\chi'_{\beta\alpha}(0,0)F_{\beta\alpha}(\omega) \tag{E.50}$$

where $F_{\alpha\beta}(\omega)$ is a spectral-weight function that has unit normalisation,

$$\int_{-\infty}^{\infty} F_{\alpha\beta}(\omega)d\omega = 1. \tag{E.51}$$

Appendix F

Ground state property estimation

In this chapter, I first demonstrate an efficient measurement scheme that is ancilla-free in Section F.1. There follows a demonstration of ground state property estimation by using quantum signal processing and phase estimation.

F.1 Ancilla-free measurement

An important component of many quantum algorithms is the ability to measure the expectation value of a unitary operation, such as those discussed in Chapter 4, Chapter 6 and Chapter 9. This quantity can be measured by using a Hadamard test circuit, which, however, introduces an additional ancillary qubit to control the unitary operation, and hence involves more controlled gates.

In many practical cases, the target problem has certain symmetries $\hat{S}$ satisfying $[H, \hat{S}] = 0$. Consequently, the state can be divided into state space with different symmetry sectors. For instance, for fermionic problems, the Hamiltonian has particle number symmetry, $[H, \hat{N}] = 0$ with $N = \sum_i a_i^\dagger a_i$. In this case, the state can be divided into different sectors in the Hilbert space with different particle numbers, $\mathcal{H} : \text{span}\{|i\rangle\}$ with $|i\rangle$ representing the state with i particles.

Suppose we aim to evaluate $\langle\psi_0|U|\psi_0\rangle$, and in our scenario the unitary is either $U = e^{-iHt}$ or $U = e^{-iHt_1}\hat{O}e^{-iHt_2}$. The expectation value $\langle\psi_0|U|\psi_0\rangle$ can be written in complex polar coordinates as $\langle\psi_0|U|\psi_0\rangle = re^{i\theta}$. In general, fidelity estimation for a general overlap $\langle\phi_i|\phi_j\rangle$ can be implemented by a SWAP test or destructive SWAP test. However, if ψ_0 is a product state, such as a mean-field state, the amplitude of the expectation value $r = |\langle\psi_0|U|\psi_0\rangle|$ can be obtained by measuring in the computational basis. The next step is to obtain the phase θ, for which we can make use of the fact that the unitary operation conserves the symmetry of $\hat{S}$, $[U, \hat{S}] = 0$. It is assumed that we can find a reference state $|R\rangle$, which lies in a different sector of the initial state, and thus we have $\langle R|U|\psi_0\rangle = 0$ [358]. If we take the reference state as the vacuum state with zero number of particle, $|R\rangle = |\text{vac}\rangle$, then the reference amplitude $r_R = 1$, and the phase $\theta_R = -t\langle R|H|R\rangle$.

Suppose the following state can be prepared

$$|\phi_0\rangle = U_s\,|\psi_0\rangle = \frac{1}{\sqrt{2}}(|R\rangle + |\psi_0\rangle). \tag{F.1}$$

Denote the amplitude and phase of $\langle R|U|R\rangle$ as r_R and θ_R, respectively, and we have $\langle R|U|R\rangle := r_R e^{i\theta_R}$. We can readily find that

$$\theta = \arccos\left(\frac{4r_s^2 - r^2 - r_R^2}{2r_R r}\right) + \theta_R \tag{F.2}$$

where

$$r_s = |\,\langle\psi_0|U_s^\dagger U U_s|\psi_0\rangle\,| \tag{F.3}$$

can be similarly measured in the computational basis.

One of the key requirements of this protocol is the preparation of the superposition state. Taking the fermionic problem as an example, the Hartree-Fock state takes the simple product-state form $|\psi_0\rangle = |0\rangle^{\otimes N-N_e}\,|0\rangle^{\otimes N_e} \in \{|N_e\rangle\}$, where the first N_e qubits are prepared in the one state and the rest of the qubits remain in the zero state. To prepare the target superposition state, $|\phi_0\rangle$, the Hadamard gate is applied to the first qubit, followed by a ladder of CNOT gates applied up to the N_eth qubit, resulting in a total of $N_e - 1$ CNOT gate operations.

To further reduce the gate count in the state preparation, we may choose $|R\rangle \in \{|N_e \pm 1\rangle\}$, so that only one additional CNOT gate is required.

F.2 Ground state property estimation by quantum signal processing

In this section, I discuss ground state preparation by quantum signal processing, proposed in [31], which achieves near-optimal asymptotic scaling in this task. The following is an overview of the resource requirements for each step in their protocol, with the quantum circuit compiled at an elementary gate level. Denote the resource as $(\cdot,\cdot,\cdot,\cdot)$ with the three elements representing the ancillary qubits, the number of CNOT gates, the number of T gates, and single-qubit R_z rotation.

1. Block encoding: $(n_L, S_{\text{CNOT}} + 2P_{\text{CNOT}},\, S_{\text{T}} + 2P_{\text{T}}, 0)$

2. Controlled select(H) and two PREP operations: $(n_L + 1, S_{\text{CNOT}} + 2P_{\text{CNOT}},\, S_{\text{T}} + 2P_{\text{T}}, 2)$

3. QSP of the sign function and hence the R operator: $(n_L{+}2, d(S_{\text{CNOT}}{+}2P_{\text{CNOT}}){+}2d, d(S_{\text{T}}{+}2P_{\text{T}}), 3d)$ with $d = [\frac{2e\lambda}{\Delta}\ln(32\pi^{-1/2}\varepsilon^{-1})]$

4. Projector, which is a controlled version of R: $(n_L + 3, d(4 + 6S_{\text{CNOT}} + 2P_{\text{CNOT}} + 2S_{\text{T}} + 2L), d(7S_{\text{CNOT}} + 5S_{\text{T}} + 2P_{\text{T}} + 4L), 4d)$.

5. Amplitude amplification. $(n_L + 3 + \lceil\frac{n-3}{2}\rceil, d\gamma^{-1}(S_{\text{CNOT}} + 2P_{\text{CNOT}} + 6n - 10), d\gamma^{-1}(S_{\text{T}} + 2P_{\text{T}} + 8n - 17), 3d\gamma^{-1})$ with $d = [\frac{2e\lambda}{\Delta}\ln(32\pi^{-1/2}\gamma^{-1}\varepsilon^{-1})]$.

The gate cost is elaborated upon in the 4th step. The controlled gates in each block are compiled as follows:

1. Controlled phase iterate operations: 2 CNOT + 2 single-qubit rotation. Thus, the total single-qubit Pauli rotation gate is $4d$.

2. Controlled C-select(H) and 2 PREP. CNOT: $6S_{\text{CNOT}} + 2P_{\text{CNOT}} + 2S_{\text{T}}$. The thrid $2S_{\text{T}}$ is from that one controlled T gate can be synthesised by 2 CNOT and 2 $\sqrt{\text{T}}$ gates.

   T gate: $7S_{\text{CNOT}} + 5S_{\text{T}} + 2P_{\text{T}}$, where the first 7 comes from Toffolis gates, the second 5 is from that one controlled T gate can be synthesised by 2 CNOT and 2 $\sqrt{\text{T}}$ gates, and we simply assume that 2 $\sqrt{\text{T}}$ may be catalysed by 5 T gats using the Hamming weight strategy introduced in [187].

   Note that the select(H) has the Hadamard gates: each controlled Hadamard gives 2 T gates and 1 CNOT gate. At least, we have $2L + n_L$ Hadamard gates in select(H).

3. The other operations are symmetric.

Step 5 accrues an additional cost from the reflection operation, whose cost can be estimated from Proposition 12. Here, since only circuit depth in a single run is compared, the cost of amplitude amplification is not taken into account.

Ground state preparation errors are composed of two parts: an error from the approximation of the sign function, and a block-encoding error of the PREP operation

$$\varepsilon = \varepsilon_{sgn} + \varepsilon_{tot,\text{PREP}}. \tag{F.4}$$

Here, one may take $\varepsilon_{sgn} = \varepsilon_{tot,\text{PREP}} = \varepsilon/2$. Considering $2d$ repetitions of the PREP operation and the relation between $\varepsilon_{\text{PREP}}$ and ϵ by $\varepsilon_{\text{PREP}} = \frac{\varepsilon}{4d\lambda}$, one may choose to set the amplitude encoding error as

$$\varepsilon_{AE} = \frac{\varepsilon}{4\lambda L d}, \tag{F.5}$$

and $n_{AE} = \lceil -\log_2 n_{AE} \rceil$.

F.2.1 Methods

In this section, I review the methods used in ground state preparation by QSP in [31]. The QSP theorem is presented below.

Lemma 3 (QSP for polynomials of definite parity [31,33])**.** *Let U be an $(\alpha, n_L, 0)$-block-encoding of a Hermitian matrix A. Let $P \in \mathbb{R}[x]$ be a degree-d even or odd real polynomial and $|P(x)| \leq 1$ for any $x \in [-1, 1]$. Then there exists a $(1, n_L+1, 0)$-block-encoding $\widetilde{U}$ of $P(A/\alpha)$ using d queries of U, $U^\dagger$, and $\mathcal{O}((n_L+1)d)$ other primitive quantum gates.*

A key component of this method is a polynomial approximation of the sign function in the domain $[-1,-\delta]\cup[\delta,1]$. Similarly to [31], one can use an explicit construction of a polynomial with the same error scaling, as shown in Lemma 9. The construction is provided in [359] based on an approximation by the Gaussian error function. A more detailed complexity analysis is presented in Lemma 7.

Lemma 4 (Polynomial approximation to the sign function $\mathrm{sgn}(x)$)**.** *For any $\delta<1$, $\varepsilon\le\sqrt{2/\pi e}$, the polynomial $S(x,\delta,\varepsilon)=p_{sgn,\delta,n}(x)=p_{erf,k,n}(x)$ of odd degree $d=\left[\frac{e}{2\delta}\ln(32\pi^{-1/2}\varepsilon^{-1})\right]=\mathcal{O}(\delta^{-1}\log(\varepsilon^{-1}))$ satisfies*

$$\varepsilon_{sgn,\delta,n}=\max_{x\in[-1,-\delta]\cup[\delta,1]}|p_{sgn,\delta,n}(x)-sgn(x)|\le\varepsilon. \tag{F.6}$$

Below, I reviewed the key results of ground state preparation in [31].

When we have the $(\lambda,n_L,0)$-block-encoding of a Hermitian matrix $H=\sum_k E_k\left|u_k\right\rangle\left\langle u_k\right|\in\mathbb{C}^{N\times N}$, $N=2^n$, $E_k\le E_{k+1}$, we can construct a $(\lambda+|\mu|,n_L+1,0)$-block-encoding of matrix $H-\mu I$ using [34, Lemma 29] for any $\mu\in\mathbb{R}$.

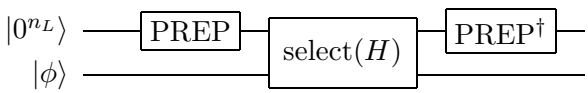


Figure F.1: Block encoding of H.

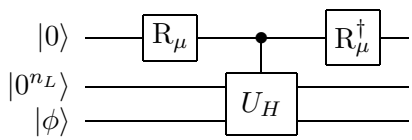


Figure F.2: Block encoding of $H-\mu I$. $\mathrm{R}_\mu\left|0\right\rangle=\sqrt{\frac{\lambda}{\lambda+|\mu|}}\left|0\right\rangle+\sqrt{\frac{|\mu|}{\lambda+|\mu|}}\left|1\right\rangle$

Then using the QSP theorem in Lemma 3, we can obtain a $(1,n_L+2,0)$-block-encoding of $-S(\frac{H-\mu I}{\lambda+|\mu|};\delta,\varepsilon)$ for any δ and ε. If it is assumed that $\Delta/2\le\min_k|\mu-E_k|$, then let $\delta=\frac{\Delta}{4\lambda}$. We know that all the eigenvalues of $-S(\frac{H-\mu I}{\lambda+|\mu|};\delta,\varepsilon)$ are ε-close to either 0 or 1, and thus $-S(\frac{H-\mu I}{\lambda+|\mu|};\delta,\varepsilon)$ is ε-close, in operator norm, to the reflector about the direct sum of eigen-subspaces corresponding to eigenvalues smaller than μ:

$$R_{<\mu}=\sum_{k:E_k<\mu}\left|u_k\right\rangle\left\langle u_k\right|-\sum_{k:E_k>\mu}\left|u_k\right\rangle\left\langle u_k\right|,$$

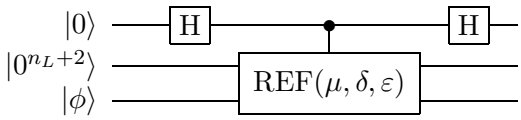
and thus the block-encoding is also a $(1,n_L+2,\varepsilon)$-block-encoding of $R_{<\mu}$. Denote this block-encoding by $\mathrm{REF}(\mu,\delta,\varepsilon)$.

Lemma 5 (Reflector and projector [31]). *Given a Hermitian matrix H with its $(\lambda, n_L, 0)$-block-encoding U_H, with the guarantee that $\mu \in \mathbb{R}$ is separated from the spectrum of H by a gap of at least $\Delta/2$, we can construct a $(1, n_L+2, \varepsilon)$-block-encoding of $R_{<\mu}$, and a $(1, n_L+3, \varepsilon/2)$-block-encoding of $P_{<\mu}$, both using $\mathcal{O}(\frac{\lambda}{\Delta}\log(\frac{1}{\varepsilon}))$ applications of U_H and $U_H^\dagger$, and $\mathcal{O}(\frac{n_L\lambda}{\Delta}\log(\frac{1}{\varepsilon}))$ other one- and two-qubit gates.*

The essential component of the QSP protocol here is projector $P_{<\mu}$, which projects the original space into the direct sum of eigen-subspaces corresponding to eigenvalues smaller than μ

$$P_{<\mu} = \sum_{k:E_k<\mu} |u_k\rangle\langle u_k| = \frac{1}{2}(I + R_{<\mu}).$$

A block-encoding of the projector can be constructed by using the controlled $\mathrm{REF}(\mu, \delta, \varepsilon)$ as



(F.7)

where H is the Hadamard gate, and denote this circuit as $\mathrm{PROJ}(\mu, \delta, \varepsilon)$. Using the relation of $P_{<\mu} = \frac{1}{2}(R_{<\mu} + I)$, $\mathrm{PROJ}(\mu, \delta, \varepsilon)$ is a $(1, n_L+3, \varepsilon/2)$-block-encoding of $P_{<\mu}$.

The result for ground state preparation is summarised in Lemma 6.

Lemma 6 (Ground state preparation with known ground energy bound [31]). *Suppose we have Hamiltonian $H = \sum_k E_k |u_k\rangle\langle u_k| \in \mathbb{C}^{N\times N}$, where $E_k \le E_{k+1}$, given through its $(\lambda, n_L, 0)$-block-encoding U_H. Also suppose we have an initial state $|\phi_0\rangle$ prepared by the circuit U_I, as well as the promises of a nonvanishing overlap and a nonvanishing gap. Then the ground state $|\psi_0\rangle$ can be prepared to fidelity $1-\varepsilon$ with the following costs:*

1. *Query complexity: $\mathcal{O}(\frac{\lambda}{\eta^{1/2}\Delta}\log(\frac{1}{\eta^{1/2}\varepsilon}))$ queries to U_H and $\mathcal{O}(\frac{1}{\eta^{1/2}})$ queries to U_I,*
2. *Number of qubits: $\mathcal{O}(n + n_L)$,*
3. *Other one- and two- qubit gates: $\mathcal{O}(\frac{n_L\lambda}{\eta^{1/2}\Delta}\log(\frac{1}{\eta^{1/2}\varepsilon}))$.*

Lemma 7 (Polynomial approximation to the sign function $\mathrm{sgn}(x)$). *For $\forall\ \kappa > 0$ there exists the polynomial $p_{sgn,\kappa,n}(x) = p_{erf,k,n}(x)$ of odd degree $n = \left\lceil \frac{e}{\kappa}\ln\frac{32}{\sqrt{\pi}\varepsilon}\right\rceil = \mathcal{O}(\frac{1}{\kappa}\log(1/\varepsilon))$ such that*

$$\varepsilon_{sgn,\kappa,n} = \max_{x\in[-1,-\kappa/2]\cup[\kappa/2,1]} |p_{erf,k,n}(x) - sgn(x)| \le \varepsilon, \tag{F.8}$$

where $p_{erf,k,n}(x)$ is the Gaussian error function $erf(kx)$ truncated at order n defined in Eq. (F.9), and $k = \frac{1}{\kappa}\sqrt{2W(\frac{8}{\pi\varepsilon^2})}$ with W being the Lambert-W function.

Following the convention in [359], $p_{\mathrm{erf},k,n}(x)$ is the Gaussian error function $erf(kx)$ truncated at order n defined as

$$p_{\mathrm{erf},k,n}(x) = \frac{2ke^{-k^2/2}}{\sqrt{\pi}}\left(I_0(k^2/2)x + \sum_{j=1}^{(n-1)/2} I_j(k^2/2)(-1)^j\left(\frac{T_{2j+1}(x)}{2j+1} - \frac{T_{2j-1}(x)}{2j-1}\right)\right), \tag{F.9}$$

where $I_j(\cdot)$ are modified Bessel functions of the first kind. In what follows, we show the explicit form of the degree, following [359] rather closely.

Proof of Lemma 7. As shown in Ref. [359], the starting point is to express the exponential decay function using the Jacobi-Anger expansion as

$$f_{\exp,\beta}(x) = e^{-\beta(x+1)} = e^{-\beta}\left(I_0(\beta) + 2\sum_{j=1}^{\infty} I_j(\beta)T_j(-x)\right), \tag{F.10}$$

where $I_j(\beta)$ are modified Bessel functions of the first kind, T_j is the Chebyshev function, and the domain of this function is assumed to be $x \in [-1, 1]$. By truncating the expansion at a finite order, $j > n$, we can obtain a degree n polynomial approximation $p_{\exp,\beta,n}(x)$ with truncation error $\varepsilon_{\exp,\beta,n}$:

$$p_{\exp,\beta,n}(x) = e^{-\beta}\left(I_0(\beta) + 2\sum_{j=1}^{n} I_j(\beta)T_j(-x)\right), \tag{F.11}$$

$$\varepsilon_{\exp,\beta,n} = \max_{x\in[-1,1]} |p_{\exp,\beta,n} - f_{\exp,\beta}| = 2e^{-\beta}\sum_{j=n+1}^{\infty} |I_j(\beta)|. \tag{F.12}$$

By solving $\varepsilon_{\exp,\beta,n}$, one can in principle obtain the required degree n as a function of β, ε.

As proven in [359, Lemma 13], $\forall \beta > 0, \varepsilon \in (0, 1/2]$, there exists a polynomial p_n of degree

$$n = \lceil \sqrt{2\lceil \max[\beta e^2, \log(2/\varepsilon)]\rceil \log(4/\varepsilon)}\rceil \tag{F.13}$$

such that $\max_{x\in[-1,1]} |p_n(x) - e^{-\beta(x+1)}| \leq \varepsilon$. A polynomial approximation to the Gaussian function follows by noticing its relation to the exponential decay function

$$p_{\text{gauss},\gamma,n}(x) = p_{\exp,\gamma^2/2,n/2}(2x^2 - 1) = e^{-\gamma^2/2}\left(I_0(\gamma^2/2) + 2\sum_{j=1}^{n/2} I_j(\gamma^2/2)(-1)^j T_{2j}(x)\right), \tag{F.14}$$

For $\forall \gamma \geq 0, \varepsilon \in (0, 1/2]$ by choosing even degree $n = \lceil \sqrt{\lceil \max[\gamma^2 e^2/2, \log(2/\varepsilon)]\rceil \log(4/\varepsilon)/2}\rceil$, the approximation error is bounded by

$$\varepsilon_{\text{gauss},\gamma,n} = \max_{x\in[-1,1]} |p_{\text{gauss},\gamma,n}(x) - e^{-(\gamma x)^2}| \leq \varepsilon. \tag{F.15}$$

which uses the relation $\varepsilon_{\text{gauss},\gamma,n} = \varepsilon_{\exp,\gamma^2/2,n/2}$.

From the definition of the error function $\text{erf}(kx) = \frac{2}{\pi}\int_0^{kx} e^{-x^2} = \frac{2k}{\sqrt{\pi}}\int_0^x e^{-(kx)^2}dx$, the polynomial $p_{\text{erf},k,n}(x) = k\int_0^x p_{\text{gauss},k,n-1}(x)dx$ follows from integrating Eq. (F.14) term-by-term. The error of the remaining terms is bounded as

$$\varepsilon_{\text{erf},k,n} \leq \frac{4ke^{-k^2/2}}{\sqrt{\pi}n}\sum_{j=(n+1)/2}^{\infty} |I_j(k^2/2)| = \frac{4k}{\sqrt{\pi}n}\varepsilon_{\text{gauss},k,n-1}. \tag{F.16}$$

We thus have the following result [359, Corollary 4]. For $\forall k > 0, \varepsilon \in (0, \mathcal{O}(1)]$ the odd polynomial $p_{\mathrm{erf},k,n}$,

$$p_{\mathrm{erf},k,n}(x) = \frac{2ke^{-k^2/2}}{\sqrt{\pi}} \left(I_0(k^2/2)x + \sum_{j=1}^{(n-1)/2} I_j(k^2/2)(-1)^j \left(\frac{T_{2j+1}(x)}{2j+1} - \frac{T_{2j-1}(x)}{2j-1} \right) \right), \quad \text{(F.17)}$$

of odd degree $n = \left\lceil \sqrt{\frac{t}{2} \ln \frac{16r}{\sqrt{\pi}\varepsilon}} \right\rceil$ and $t = \lceil \max\{\frac{k^2}{2}e^2, \ln \frac{8r}{\sqrt{\pi}\varepsilon}\} \rceil$, suffices to bound the error up to ε as

$$\varepsilon_{\mathrm{erf},k,n} = \max_{x \in [-1,1]} |p_{\mathrm{erf},k,n}(x) - \mathrm{erf}(kx)| \leq \varepsilon. \quad \text{(F.18)}$$

which uses $\varepsilon_{\mathrm{erf},k,n} \leq \frac{4k}{\sqrt{\pi}n} \varepsilon_{\mathrm{gauss},k,n-1} \leq \varepsilon$.

The equation for k comes from Lemma 8, which shows the relation of the sign function and the error function. We have $k = \frac{1}{\kappa}\sqrt{2W(\frac{2}{\pi\varepsilon^2})} \geq \frac{1}{\kappa}\sqrt{2\ln(\frac{2}{\pi\varepsilon^2})}$, with W being the Lambert-W function.

In the above equation of n, we defined $r = k/n$ which is dependent of n. The degree is of the order $n = \Omega(k \ln^{1/2}(1/\varepsilon))$, and hence $r = k/n = \mathcal{O}(\ln^{-1/2}(1/\varepsilon)) = \mathcal{O}(1)$, which does not change the asymptotic scaling. Since n has a logarithmic dependence of r, we set $r = 1$ for the sake of simplification. To obtain a more accurate estimation, one may calculate n self consistently. Under this simplification, the polynomial approximation of odd degree $n = \left\lceil \frac{e}{\kappa} \ln \frac{16}{\sqrt{\pi}\varepsilon} \right\rceil$ suffices to bound the approximation error of the error function to ε.

Lemma 8 relates the error function and the sign function. To suppress the error of approximation to the sign function, $\varepsilon_{\mathrm{sgn},\kappa,n} = \max_{|x| \geq \kappa/2} |p_{\mathrm{erf},\kappa,n}(x) - \mathrm{sgn}(x)| \leq \varepsilon$, we can split the error into two parts; the polynomial $p_{\mathrm{erf},\kappa,n}$ of odd degree

$$n = \left\lceil \frac{e}{\kappa} \ln \frac{32}{\sqrt{\pi}\varepsilon} \right\rceil \quad \text{(F.19)}$$

and $k \geq \frac{1}{\kappa}\sqrt{2\ln(\frac{8}{\pi\varepsilon^2})}$ suffices to suppress the approximation error less than ε.

□

Lemma 8 (Approximation to the sign function $\mathrm{sgn}(x)$, Lemma 10 in [359])**.** *$\forall\, \kappa > 0, x \in \mathbb{R}, \varepsilon \in (0, \sqrt{2/e\pi}]$, let $k = \frac{1}{\kappa}\sqrt{2\log(\frac{2}{\pi\varepsilon^2})}$. Then the function $f_{sgn,\kappa,\varepsilon}(x) = erf(kx)$, defined by the error function, satisfies*

$$\max_{|x| \geq \kappa/2} |f_{sgn,\kappa,\varepsilon}(x) - sgn(x)| \leq \varepsilon, \quad |f_{sgn,\kappa,\varepsilon}(x)| \leq 1. \quad \text{(F.20)}$$

Lemma 9 (Time complexity of quantum eigenvalue transformation of unitary matrices)**.** *The polynomial approximation to the shifted sign function of degree $d = \left\lceil \frac{e}{\delta} \ln(32\pi^{-1/2}\varepsilon^{-1}) \right\rceil = \mathcal{O}(\delta^{-1}\log(\varepsilon^{-1}))$ satisfies*

$$\varepsilon_{sgn,\delta,n} = \max_{x \in [-1,-\delta] \cup [\delta,1]} |p_{sgn,\delta,n}(x) - sgn(x)| \leq \varepsilon. \quad \text{(F.21)}$$

The time complexity of quantum eigenvalue transformation of unitary matrices with real polynomials, termed QET-U, is determined by d.

F.3 Ground state property estimation with phase estimation

In this section, ground state property estimation with phase estimation is reviewed. For the canonical QPE algorithm, we apply a series controlled U, U^2, ..., U^{2^k-1} and an inverse quantum Fourier transform on the ancillary k qubits, such that the state becomes

$$\sum_i c_i \left|0^{\otimes k}\right\rangle \left|E_i\right\rangle \to \sum_i p_i \left|\mathrm{bin}(E_i)\right\rangle \left|E_i\right\rangle . \tag{F.22}$$

I discuss ground state property estimation with phase estimation using Hamiltonian simulation as a subroutine.

F.3.1 Complexity of phase estimation

Observation 1 (Ground state energy estimation with phase estimation)**.** *To obtain a binary estimate of the energy precise to $n = \lceil \log_2 \varepsilon^{-1} \rceil$ bits, we require $k = \mathcal{O}(\log_2 \varepsilon^{-1} + \log_2 c_0^{-1})$ ancillary qubits. The coherent runtime for each phase estimation is $2^{k+1}\pi = \mathcal{O}(\varepsilon^{-1} c_0^{-2})$, and the number of calls to phase estimation is $\mathcal{O}(c_0^{-1})$. The total gate complexity is*

$$\mathcal{O}(\mathcal{C}_{gate} c_0^{-3} \varepsilon^{-1}). \tag{F.23}$$

To obtain a binary estimate of the energy precise to $n = \lceil \log_2 \varepsilon^{-1} \rceil$ bits, we require $k = \mathcal{O}(\log_2 \varepsilon^{-1} + \log_2 c_0^{-1})$ ancillary qubits.

The total error is composed of three parts: the error of phase estimation, Hamiltonian simulation, and circuit synthesis,

$$\varepsilon_{\mathrm{tot}} = \varepsilon_{PE} + \varepsilon_{HS} + \varepsilon_{CS}.$$

The coherent runtime for each phase estimation is lower bounded by

$$t_{PE}^{En} = \frac{\pi}{2 c_0^2 \varepsilon_{PE}}. \tag{F.24}$$

Observation 2 (Ground state preparation with phase estimation for known ground energy)**.** *Using the canonical phase estimation, the state can be prepared ε close to ground state using $k = \mathcal{O}(\log_2 \varepsilon^{-1} + \log_2 \Delta^{-1} + \log_2 c_0^{-1})$. The runtime for each phase estimation is $2^{k+1}\pi = \mathcal{O}(\varepsilon^{-1} \Delta^{-1} c_0^{-1})$, and the number of calls to phase estimation is $\mathcal{O}(c_0^{-1})$ using fixed point search. The total gate complexity is*

$$\mathcal{O}(\frac{\mathcal{C}_{gate}}{c_0^2 \Delta \varepsilon}) \tag{F.25}$$

The coherent runtime for each phase estimation is lower bounded by

$$t_{PE}^{prep} = \frac{\pi}{2c_0 \Delta \varepsilon_{PE}}. \tag{F.26}$$

The error of estimation of an observable $\hat{O}$ consists of three components:

$$\varepsilon = \varepsilon_{PE} + \varepsilon_{HS} + \varepsilon_{CS} + \varepsilon_{observ}.$$

The error of estimating observables using N_s samples is given by

$$\varepsilon_{observ} = \frac{C_{observ}}{\sqrt{N_s}}. \tag{F.27}$$

Suppose we use importance sampling to estimate the observable $\hat{O} = \sum_l o_l P_l$. The measurement overhead is $C_{observ} = \|\mathbf{o}\|_1$, and we may apply other methods to reduce C_{observ}, such as Pauli grouping or the classical shadow methods introduced in Chapter 10.

In our resource analysis, to streamline the comparison, we exclude the observable estimation error, and solely consider the state preparation error, guaranteed by Proposition 7.

A key element in phase estimation is to realise controlled operations $C - U^{2^{m-1}}$ with $m = 1, ..., k$. These operations can be realised using Hamiltonian simulation methods, such as Trotter formulae, quantum signal processing, and others. In the following, I give an optimistic estimation by only considering the dominant cost from the last controlled operations $C - U^{2^{k-1}}$. The gate cost for QFT is neglected as well, which scales as $\mathcal{O}(k)$.

F.3.2 Hamiltonian simulation by Trotterisation

Now, I discuss resource cost of Hamiltonian simulation by Trotterisation which is introduced in Section 2.4. The overall circuit complexity for $2k$ order achieves minima when $\varepsilon_{PE} = \varepsilon_{HS} = \varepsilon/2$

$$\frac{\left(\pi \cdot 5^{k-1} L \Lambda c_0^{-2}\right)^{1+\frac{1}{2k}}}{((2k+1)!)^{\frac{1}{2k}} \varepsilon_{PE}^{1+\frac{1}{2k}} \varepsilon_{HS}^{\frac{1}{2k}}}. \tag{F.28}$$

Its minimum is obtained at

$$\varepsilon_{PE} = \frac{2k+1}{2(k+1)}\varepsilon, \ \varepsilon_{HS} = \frac{1}{2(k+1)}\varepsilon.$$

The total gate complexity using higher-order Trotter is determined by the following procedure

Gate count for eigenenergy estimation.

1. Get the runtime $t_{PE}/2$ with $\varepsilon_{PE} = \varepsilon/2$ in Eq. (F.24).

2. Determine the number of segment ν using Lemma 1.

3. CNOT gates: $2 \cdot 5^{k-1} \nu c_0^{-1} (2\,\mathrm{wt}(H) - L + 2)$. Single-qubit z-axis Pauli rotation gate: $4 \cdot 5^{k-1} \nu c_0^{-1} L$.

Gate count for observable estimation.

1. Get the measurement overhead C_{observ}, runtime $t_{PE}/2$ as a function of ε_{observ} and ε_{PE}, respectively.

2. Determine the number of segment ν using Lemma 1 as a function of ε_{HS}. An approximation of ν is given by Eq. (9.37).

3. CNOT gates: $2 \cdot 5^{k-1} \nu c_0^{-1} C_{observ} \varepsilon_{observ}^{-2} (2\,\mathrm{wt}(H) - L + 2)$. Single-qubit z-axis Pauli rotation gate: $4 \cdot 5^{k-1} \nu C_{observ} \varepsilon_{observ}^{-2} c_0^{-1} L$.

4. Get the gate count by optimising over the distribution of ε.

F.3.3 Hamiltonian simulation by qubitised quantum walk

Phase estimation combining the qubitised quantum walk methods have been discussed in [101, 239]. The key idea is that the spectrum of H can be obtained by performing phase estimation on the Szegedy quantum walk operator, defined as

$$\mathcal{W} := (2\,|G\rangle\,\langle G| \otimes I - I) \cdot \mathrm{select}(H), \tag{F.29}$$

with $|G\rangle = \mathrm{PREP}\,|\bar{0}\rangle$. The spectrum has the relation

$$\mathrm{spectrum}(H) = \lambda \cos(\arg[\mathrm{spectrum}(\mathcal{W})]), \tag{F.30}$$

with $\arg(e^{i\phi}) = \phi$.

Their results suggest that we can estimate the phase to a number of bits given by

$$k = \left\lceil \log\left(\frac{\sqrt{2}\pi\lambda}{2\varepsilon_{PE}}\right)\right\rceil \tag{F.31}$$

with k extra ancillary qubits. Here, we further assume a small error of gate synthesis in PREP and QFT. Using phase estimation, the query number is

$$d := 2^k \leq \frac{\sqrt{2}\pi\lambda}{2\varepsilon_{PE}} + 1. \tag{F.32}$$

We set $\varepsilon_{\mathrm{PREP}} = \varepsilon_{PE} = \varepsilon/2$. The state preparation error $\varepsilon_{\mathrm{PREP}}$ for a single application of $\mathcal{W}$ is

$$\varepsilon_{tot,\mathrm{PREP}} \leq \|e^{i\arccos(H/\lambda)} - e^{i\arccos(\tilde{H}/\lambda)}\| \tag{F.33}$$

It is related to the amplitude encoding error ε_{AE} by

$$\varepsilon_{\mathrm{PREP}} \leq \frac{L\varepsilon_{AE}}{\lambda}\left(1 - \left(\frac{\|H\| + L\varepsilon_{AE}}{\lambda}\right)^2\right)^{-1/2} \tag{F.34}$$

Suppose that we require the preparation error to be $\varepsilon_{\mathrm{PREP}}$. The preparation error is set to be

$$\varepsilon_{\mathrm{PREP}} \leq \frac{\sqrt{2}}{2\lambda}\frac{\varepsilon_{PE}}{2^k} = \frac{\sqrt{2}\varepsilon_{PE}}{\lambda 2^{k+1}} \tag{F.35}$$

and hence for a single block

$$\varepsilon_{AE} = \frac{\sqrt{2}\varepsilon}{4L\lambda d} \tag{F.36}$$

and we have 2 preparation in one block. Again, the cost from the QFT is ignored, which scales as $\mathcal{O}(k \log k)$.

The overall gate complexity of the eigenenergy estimation is

$$\mathcal{O}(\frac{\lambda L}{\varepsilon}). \tag{F.37}$$

The total gate count can be estimated by using Proposition 11, Proposition 12, and the Hamiltonian dependent select(H), given by Corollary 3.

Each block requires: 1 controlled select(H), 2 PREP and 1 Reflection on n_L qubits, which has the gate count

$$(n_L + \max(k, n_L + 2n_{AE} + 1), S_{\text{CNOT}} + 2P_{\text{CNOT}} + 6(n_L - 2), S_{\text{T}} + 2P_{\text{T}} + (8n_L - 17), 0). \tag{F.38}$$

and additional k repetition of controlled reflection, each block (2 preparation) has the cost:

$$(0, 2P_{\text{CNOT}} + 6(n_L - 1), 2P_{\text{T}} + (8n_L - 9), 0) \tag{F.39}$$

We require d queries and $\eta^{-1}\Delta^{-1}$ repetitions

$$\begin{aligned}(n_L + \max(k, (n_L + 2n_{AE} + 1), &\eta^{-1}\Delta^{-1}\left(d(S_{\text{CNOT}} + 2P_{\text{CNOT}} + 6(n_L - 2)) + k(2P_{\text{CNOT}} + 6(n_L - 1))\right), \\ &\eta^{-1}\Delta^{-1}\left(d(S_{\text{T}} + 2P_{\text{T}} + (8n_L - 17)) + k(2P_{\text{T}} + (8n_L - 9))\right), 0).\end{aligned} \tag{F.40}$$

F.4 Paired Taylor-series sampling with zeroth- and higher-order Trotter formulae

This section discusses Hamiltonian simulation by paired Taylor-series sampling with zeroth and higher-order Trotter formulae, which were introduced in Chapter 9, in more detail.

F.4.1 Leading-order pairing

We can expand the 0th and 1st order terms in Eq. (9.9) and group them as,

$$\begin{aligned} U(x) &= \sum_{s=0}^{\infty} F_0^{(s)}(x) \\ &= I - i\lambda x \sum_{l=1}^{L} p_l P_l + \sum_{s=2}^{\infty} F_0^{(s)}(x) \\ &= \sum_{l=1}^{L} p_l (I - i\lambda x P_l) + \sum_{s=2}^{\infty} F_0^{(s)}(x). \end{aligned} \tag{F.41}$$

Then, we can apply Eq. (9.12) on Eq. (F.41) to convert the leading-order terms to Pauli rotation unitaries

$$U(x) = \sqrt{1+(\lambda x)^2}\sum_{l=1}^{L} p_l e^{i\theta_0 P_l} + \sum_{s=2}^{\infty} F_0^{(s)}(x), \tag{F.42}$$

where $\theta_0 = \tan^{-1}(-\lambda x)$. Taking out the normalisation factor, we can write Eq. (F.42) as a RLCU formula,

$$U(x) = \mu_0^{(L)}(x)\left(\mathrm{Pr}_0^{(L)}(1)\sum_{l=1}^{L} p_l e^{i\theta_0 P_l} + \sum_{s=2}^{\infty}\mathrm{Pr}_0^{(L)}(s)U_{0,L}^{(s)}\right), \tag{F.43}$$

where

$$\begin{aligned}
\mu_0^{(L)}(x) &:= \sqrt{1+(\lambda x)^2} + \sum_{s=2}^{\infty}\mu(F_0^{(s)}(x)) \\
&= \sqrt{1+(\lambda x)^2} + \left(e^{\lambda x} - 1 - \lambda x\right), \\
\mathrm{Pr}_0^{(L)}(s) &:= \frac{1}{\mu_0^{(L)}(x)}\begin{cases}\sqrt{1+(\lambda x)^2}, & s=1, \\ \mu(F_0^{(s)}(x)) = \frac{(\lambda x)^s}{s!}, & s\geq 2,\end{cases} \\
U_{0,L}^{(s)} &:= \sum_{l_1,l_2,\ldots,l_s} p_{l_1}p_{l_2}\ldots p_{l_s}(-i)^s P_{l_1}P_{l_2}\ldots P_{l_s}.
\end{aligned} \tag{F.44}$$

After "pairing" the terms with $s=0$ and $s=1$, we obtain Eq. (F.43) which is a new RLCU formula with the normalisation factor $\mu_0^{(L)}(x)$.

In practice, to limit the number of gates in $U_{0,L}^{(s)}$ as well as avoiding sampling over infinite orders, we set a truncation integer $s_c > 2$ to the expansion order s. When $s > s_c$, we do not implement any gate. The truncated RLCU formula is then,

$$\tilde{U}_0^{(L)}(x) = \mu_0^{(L)}(x)\left(\mathrm{Pr}_0^{(L)}(1)\sum_{l=1}^{L} p_l e^{i\theta_0 P_l} + \sum_{s=2}^{s_c}\mathrm{Pr}_0^{(L)}(s)U_{0,L}^{(s)} + \sum_{s>s_c}\mathrm{Pr}_0^{(L)}(s)I\right). \tag{F.45}$$

After the introduction of the truncation, we can equivalently regard the probability distribution $\mathrm{Pr}_0^{(L)}(s)$ as a distribution with finite number of possible values,

$$\mathrm{Pr}_0^{(L)}(s) := \begin{cases}\frac{1}{\mu_0^{(L)}(x)}\sqrt{1+(\lambda x)^2}, & s=1, \\ \frac{1}{\mu_0^{(L)}(x)}\frac{(\lambda x)^s}{s!}, & 2\leq s\leq s_c, \\ 1-\sum_{r=1}^{s_c}\mathrm{Pr}_0^{(L)}(r), & s=s_c+1.\end{cases} \tag{F.46}$$

Sampling s from Eq. (F.46), the same $\tilde{U}_0^{(L)}(x)$ will be obtained as the one in Eq. (F.45).

The following proposition gives the performance characterization of $\tilde{U}_0^{(L)}(x)$ to approximate $U(x)$.

Proposition 21 (Random-sampling LCU using the zeroth-order leading-order-pairing formula)**.** *For $x>0$ and $s_c \geq 1$, $\tilde{U}_0^{(L)}(x)$ in Eq. (F.45) is a $(\mu_0^{(L)}(x), \varepsilon_0^{(L)}(x))$-RLCU formula of $U(x)$ with*

$$\begin{aligned}
\mu_0^{(L)}(x) &\leq e^{\frac{3}{2}(\lambda x)^2}, \\
\varepsilon_0^{(L)}(x) &\leq 2\left(\frac{e\lambda x}{s_c+1}\right)^{s_c+1}.
\end{aligned} \tag{F.47}$$

The sampling procedure of $\tilde{U}_0^{(L)}(x)$ in Eq. (F.45) is summarised in Algorithm 6. This algorithm is called leading-order-rotation algorithm in Ref. [242]; Hereafter, it is referred to as the leading-order-pairing algorithm.

Algorithm 6 Random gate-sequence sampling using zeroth-order leading-order-pairing formula in Eq. (F.45) [242]

Require: an n-qubit Hamiltonian $H = \lambda \sum_{l=1}^{L} p_l P_l$ with $p_l \geq 0 \, \forall l$ and $\sum_l p_l = 1$, evolution time $x > 0$, Taylor-series expansion truncation order $s_c \in \mathbb{N}$.
Ensure: A random unitary sequence U_w.

```
Set the gate sequence U_w = I.
Calculate the rotation angle θ_0 := tan^{-1}(−λx).
Sample a positive integer s with the probability Pr_0^{(L)}(s) defined in Eq. (F.44).
if s = 1 then                                                    ▷ Leading order case
    Sample an index l with the probability {p_l}_{l=1}^L.
    Append the gate V_{0,L}^{(1)} := e^{iθ_0 P_l} to U_w.
else if s ≤ s_c then                                             ▷ Normal high-order case
    Independently sample s indices l_1, ..., l_s with the probability {p_l}_{l=1}^L.
    Append the Pauli gates P_{l_s}, ..., P_{l_1} to U_w.
    Append (−i)^s I to U_w.
else                                                             ▷ Truncated high-order case
    Append the gate I to U_w.
end if
return U_w.
```

The overall RLCU formula for $U(t)$ is to repeat the sampling procedure in Algorithm 6 for ν times, $\tilde{U}_0^{(L)}(t) = \tilde{U}_0^{(L)}(x)^\nu$. Using Proposition 21, we conclude that $\tilde{U}_0^{(L)}(t)$ is a $(\mu_{0,tot}^{(L)}(t), \varepsilon_{0,tot}^{(L)}(t))$-RLCU formula of $U(t)$ with

$$\begin{aligned} \mu_{0,tot}^{(L)}(t) &= \mu_0^{(L)}(x)^\nu \leq e^{\frac{3}{2}\frac{(\lambda t)^2}{\nu}}, \\ \varepsilon_{0,tot}^{(L)}(t) &\leq \nu \mu_{0,tot}^{(L)}(t) \varepsilon_0^{(L)}(x) \leq 2\nu e^{\frac{3}{2}\frac{(\lambda t)^2}{\nu}} \left(\frac{e\lambda x}{s_c + 1} \right)^{s_c+1}. \end{aligned} \tag{F.48}$$

F.4.2 Higher-order pairing

One can consider an alternative way to suppress the normalisation factor by inserting the Trotter formula before the random sampling. That is, instead of sampling the time evolution $U(x)$ directly, one now considers to sample the (multiplicative) remainder $V_k(x)$ of the Trotter formulae defined in Eq. (2.50). If we construct a good RLCU formula $\tilde{V}_k(x)$ for the remainder $V_k(x)$, then the overall circuit with k-th Trotter formula will be a RLCU for $U(x)$. The gate complexity of inserting the higher-order Trotter formula is summarised in Proposition 9 and Table 9.2.

Appendix G

Quantum error mitigation

G.1 Stochastic error mitigation

In this section, I show the intuition of continuous error mitigation and stochastic error mitigation.

G.1.1 Continuous error mitigation scheme

I first illustrate the detailed procedure of continuous error mitigation. We can rewrite the evolution of noisy and ideal quantum states by using infinitesimal δt as

$$\begin{aligned} \rho_N(t+\delta t) &= \rho_N(t) + \delta t\big\{ -i[H(t), \rho_N(t)] + \lambda\mathcal{L}\big[\rho_N(t)\big]\big\} \\ \rho_I(t+\delta t) &= \rho_I(t) + \delta t\big\{ -i[H(t), \rho_I(t)]\big\}, \end{aligned} \tag{G.1}$$

where $H(t)$ denotes the ideal Hamiltonian with $\mathcal{L}$ corresponding to the noisy evolution. In the presence of Markovian stochastic noise involved with environment,

$$\mathcal{L}[\rho] = \frac{1}{2}\sum_k (2L_k\rho L_k^\dagger - L_k^\dagger L_k\rho - \rho L_k^\dagger L_k), \tag{G.2}$$

while the dynamics induced with the undesired Hamiltonian $H_C(t)$ which causes coherent errors can be described as

$$\mathcal{L}[\rho] = -i[H_C(t), \rho]. \tag{G.3}$$

The latter case occurs due to the imperfection of the analogue quantum simulators and implementation of quantum logic gates from physical Hamiltonians [192,286]. For systems with finite-range interactions, Bairey *et al.* and Silva *et al.* proposed methods that use only local measurements to reconstruct local Markovian dynamical process [304,305]. I show how to eliminate these errors by using a continuous error mitigation method.

By using the Pauli transfer matrix representation, Eq. (G.1) is mapped to $|\rho_\alpha(t+\delta t)\rangle\rangle = (I + E_\alpha(t)\delta t)\,|\rho_\alpha(t)\rangle\rangle$ where $|\rho_\alpha(t)\rangle\rangle$ $(\alpha = N, I)$ is the vectorised density matrix of $\rho_\alpha(t)$ and $E_\alpha(t)$ corresponds to the second term of Eq. (G.1). Equivalently, the superoperartor representation of the evolution gives $\rho_\alpha(t+\delta t) = \mathcal{E}_\alpha(t)\rho_\alpha(t)$. In the following, we will use these two equivalent

representations interchangeably. Note that the evolution induced by $\mathcal{E}_\alpha$ in the main text becomes $I + E_\alpha \delta t$ in the Pauli transfer representation. We introduce the recovery operation $I + E_Q \delta t$ to obtain the ideal dynamics, which can be expressed as

$$(I + E_Q \delta t)(I + E_N \delta t) = I + E_I \delta t + \mathcal{O}(\delta t^2) \tag{G.4}$$

such that $E_Q = E_I - E_N$. Note that $(I + E_Q \delta t)$ corresponds to $\mathcal{E}_Q$ in the main text. Due to the linearity of the representation, we can see that E_Q corresponds to the Pauli transfer matrix representation of $-\lambda \mathcal{L}\big[\rho_N(t)\big]$. In this framework, $\mathcal{E}_I(t) \approx \mathcal{E}_N(t)$ holds within a sufficiently small timestep δt.

The experimental errors including the interactions with the open environment, undesired couplings and imperfections in the quantum simulators are generally local. We therefore assume $\mathcal{E}_Q$ can be decomposed into local operators as $\mathcal{E}_Q = \sum_{S=1}^{N_S} \mathcal{E}_Q^{(S)}$, where $\mathcal{E}_Q^{(S)}$ operates on polynomial subsystems of the N-qubit quantum system. We now decompose the operation $\mathcal{E}_Q$ into the set of basis operations as

$$\mathcal{E}_Q^{(S)} = \sum_{j \geq 0} q_j^{(S)} \mathcal{B}_j^{(S)}, \tag{G.5}$$

where $q_j^{(S)}$ is the quasi-probability and $\mathcal{B}_j^{(S)}$ is the basis operation for compensating the errors. Note that $\mathcal{B}_j^{(S)}$ only acts on the same small subsystem as $\mathcal{E}_Q^{(S)}$. By performing basis operations for $\mathcal{E}_Q^{(S)}$ with corresponding quasi-probability distributions in Eq. (G.5), we can implement the overall quasi-probability operations corresponding to $\mathcal{E}_Q$ as shown below. Therefore, we can extend the quasi-probability operations into a large-scale system. In addition, this argument can be naturally applied to multi-level systems when we can prepare basis operations for them.

In particular, the quasi-probability operation at time t takes the form of

$$\begin{aligned} \mathcal{E}_Q &= (1 + q_0 \delta t)\mathcal{I} + \sum_{i \geq 1} q_i \delta t \mathcal{B}_i, \\ &= c \left(p_0 \mathcal{I} + \sum_{i \geq 1} \alpha_i p_i \mathcal{B}_i \right) \end{aligned} \tag{G.6}$$

where $\mathcal{B}_0$ is an identity operation and the superscript (S) is omitted for simplicity. The probability to generate the identity operation $\mathcal{I}$ and $\mathcal{B}_i$ $(i \geq 1)$ is $p_0 = 1 - \sum_{i \geq 1} p_i$ and $p_i = |q_i| \delta t / c$ $(i \geq 1)$, where $c = \sum_{i \geq 0} p_i = 1 + (q_0 + \sum_{i \geq 1} |q_i|)\delta t$. In addition, the parity α_0 for $\mathcal{B}_0 = \mathcal{I}$ is always unity, and the parity α_i corresponding to $\mathcal{B}_i$ $(i \geq 1)$ equals to $\mathrm{sign}(q_i)$.

The overhead coefficient c corresponding to $\mathcal{E}_Q^{(S)}$ is given by $c = 1 + C_1^{(S)} \delta t$, with $C_1^{(S)} := (q_0^{(S)} + \sum_{i \geq 1} |q_i^{(S)}|)$. As discussed above, this coefficient introduces a sampling overhead. The overhead coefficient from $t = 0$ to $t = T$ within infinitely small discretisation δt is

$$C(T) = \lim_{\delta t \to 0} \prod_S \prod_{k=0}^{T/\delta t} (1 + C_1^{(S)} \delta t) = \prod_S \exp\left(C_1^{(S)} T \right). \tag{G.7}$$

Note that $|q_i| \propto \lambda$, therefore we have $C_1 \propto \lambda$, and the overall overhead is

$$C(T) = \exp(\mathcal{O}(\lambda T)). \tag{G.8}$$

A proper normalisation λ is chosen so that the contribution of $\mathcal{L}$ is bounded by a constant l: $\|\mathcal{L}_{\text{exp}}\|_1 \leq l$. Here, we define the super-operator norm by $\|\Phi\|_1 = \sup_A\{\|\Phi(A)\|_1/\|A\|_1 : A \neq 0\}$ with $\|A\|_1 = \text{Tr}|A|$. Therefore, given a finite number of samples in experiments, the condition that the scheme works efficiently with a constant resource cost is $\lambda T = \mathcal{O}(1)$. By interpreting λT as the total noise strength, the requirement is thus consistent with the case of DQS.

It is also possible to consider time-dependent recovery operation for suppressing time-dependent noise. In this case, the quasi-probability becomes time-dependent and can be obtained by Eq. (G.5). Therefore, the overall overhead for time-dependent noise is

$$C(T) = \lim_{\delta t \to 0} \prod_S \prod_{k=0}^{T/\delta t} (1 + C_1^{(S)}(k\delta t)\delta t) = \prod_S \exp\left(\int_0^T C_1^{(S)}(t)dt\right). \tag{G.9}$$

G.1.2 Algorithm

I show the stochastic error mitigation as follows.

Algorithm 7 Stochastic error mitigation. Input: initial state $\rho(0)$, number of samples N_s, noisy evolution $\mathcal{E}_N$, basis operations $\mathcal{B}_j$. Output: $\bar{O}$.

```
Get C, {α_j}, and {p̃_j}, set { s_j = (Σ_{i=1}^{j} p̃_i) / (Σ_{i=1}^{N_op} p̃_i) }.
for m = 1 to N_s do
    Randomly generate q_0 ∈ [0,1], set t = 0, n = 0, α = 1.
    while t ≤ T do
        Get t_jp^n by solving exp(−Γ(t_jp^n)) = q_n.
        Randomly generate q'_n ∈ [0,1].
        Set j_n = j if q'_n ∈ [s_{j−1}, s_j] and update α = α_{j_n} · α.
        Update t = t + t_jp^n and n = n + 1.
    end while
    Set ρ_Q = ρ(0) and Ō = 0.
    for k = 0 : n − 1 do
        Evolve ρ_Q under E_N for time t_jp^k and apply B_{j_k}.
    end for
    Evolve ρ_Q under E_N for time T − Σ_{k=0}^{n−1} t_jp^k.
    Measure O of ρ_Q to get O_m.
    Update Ō = Ō + CαO_m/N_s
end for
```

G.1.3 Comparison with conventional error mitigation

Errors, occurring in the continuous time evolution, can inherently mix and propagate along time evolution, leading to highly nonlocal correlated effects. For instance, dominant errors in superconducting qubits are inherent system dephasing or relaxation, and coherent errors (or crosstalk)

when applying entangling gates. analogue quantum simulators may not even implement discretised quantum gates. Therefore, conventional quantum error mitigation methods fail to work for realistic errors and general continuous quantum processes. This problem is addressed by first considering a more general scenario of a continuous process with realistic noise models. More concretely, I consider the time-independent Lindblad master equation

$$\frac{d\rho}{dt} = (\mathcal{H} + \mathcal{L})(\rho) \tag{G.10}$$

with dynamics of Hamiltonian (including coherent errors) and incoherent Markovian process

$$\begin{aligned} \mathcal{H}(\rho) &= -i[H + \delta H, \rho] \\ \mathcal{L}(\rho) &= \frac{1}{2}\sum_k (2L_k \rho L_k^\dagger - L_k^\dagger L_k \rho - \rho L_k^\dagger L_k), \end{aligned} \tag{G.11}$$

which describes either gate synthesis in digital quantum computing or the continuous evolution of an analogue quantum simulator. Here δH and $\mathcal{L}$ describe coherent errors (such as crosstalk or imperfections of Hamiltonian) and inherent coupling with the environment (such as dephasing and damping), respectively. Even though the coherent error δH and the Lindblad operators L_k act locally on the quantum system, the effect of errors propagates to the entire system after the evolution. Therefore, such global effects of noise cannot be effectively mitigated using the conventional quasi-probability method, which assumes a simple gate-independent error model described by single- or two-qubit error channels before or after each gate.

G.1.4 Decomposition of the recovery operation and optimisation

In this section, I discuss the decomposition of the recovery operation into local basis operations. Denote the complete basis operations as $\{\mathcal{B}_i\}$. For multiple qubit systems, tensor products of single-qubit operations, e.g., $\mathcal{B}_i \otimes \mathcal{B}_j$, also form a complete basis set for composite systems. Therefore, if we can implement the complete basis operations for a single-qubit, we can also emulate arbitrary operations for multiple qubits systems. In Section 2.5.3, it is shown that every single-qubit operation can be emulated by using 16 basis operations. Table 2.1 of Section 2.5.3 shows one efficient set of basis operations for a single-qubit.

I show in the Section G.1.1 that the recovery operations without Hamiltonian error can be analytically expressed as

$$\mathcal{E}_Q = \mathcal{I} - \lambda \mathcal{L} \delta t, \tag{G.12}$$

where $\mathcal{L}$ represents the noise superoperator and λ is the noise strength. From Eq. (G.12), we can analytically decompose the general noise into local basis operations. In the Section 2.5.3, we provide the recovery operations for several typical Markovian processes, including depolarising, dephasing and amplitude damping, during the quantum simulation. It is worth mentioning that by using only observables within the spatial domain, we can recover the Lindbladian acting on this domain and reconstruct the local Markovian dynamics [304, 305].

Over-complete basis can be used to further reduce the resource cost for the stochastic error mitigation scheme. In general, the target quasi-probability operation $\mathcal{E}_Q$ can be decomposed as a linear combination of unitary channels and projective measurements by using Pauli transfer matrix representation. The quasi-probability operation $\mathcal{E}_Q$ can be decomposed into a complete basis $\{\mathcal{B}_i\}$ as

$$\mathcal{E}_Q = \sum_i q_i \mathcal{B}_i, \tag{G.13}$$

where we set $\mathcal{B}_0 = \mathcal{I}$. Given the target quasi-probability operation $\mathcal{E}_Q$, the overall resource cost for the quasi-probability scheme is given by $C(T) = \exp(C_1 T)$ with $C_1 = q_0 + \sum_{i\geq 1} |q_i|$.

In order to minimise the resource cost, we aim to reduce C_1. Consider an over-complete basis, $\{\mathcal{B}'_i\}$ which includes the complete basis $\{\mathcal{B}_i\}$ and also other randomly generated unitary operators and projective measurements. Then the quasi-probability operation E_Q is decomposed into this over-complete basis $\{\mathcal{B}'_i\}$ as

$$\mathcal{E}_Q = \sum_i q'_i \mathcal{B}'_i. \tag{G.14}$$

Minimising $C_1 = q_0 + \sum_{i\geq 1} |q_i|$ can be further rewritten as a linear programming as follows,

$$\begin{aligned} \min C_1 &= q_0 + \sum_{i\geq 1} (q_i^+ - q_i^-), \\ s.t.\ \mathcal{E}_Q &= \sum_i (q_i^+ - q_i^-) \mathcal{B}'_i, \\ & q_i^+, q_i^- \geq 0. \end{aligned} \tag{G.15}$$

The overall resource cost $C(T)$ for stochastic error mitigation scheme can therefore be reduced by this linear programming optimisation method.

G.1.5 Decomposition of Lindbladian

By using only observables within the spatial domain, we can recover the Lindbladian acting on this domain and reconstruct the local Markovian dynamics [304]. Here, I provide the recovery operations for several typical Markovian processes during the quantum simulation and coherent errors in implementing CNOT gates.

The recovery operations can be analytically expressed as $\mathcal{E}_Q = \mathcal{I} - \lambda \mathcal{L} \delta t$, where $\mathcal{L}$ represents the noise superoperator and λ is the noise strength. For depolarising, dephasing and amplitude damping, the recovery operations $\mathcal{E}_Q$ can be, respectively, decomposed as

$$\begin{aligned} \mathcal{E}_Q^{\text{depolarise}} &= (1 + \frac{3}{4}\lambda\delta t)\mathcal{I} - \frac{\lambda}{4}(\mathcal{X} + \mathcal{Y} + \mathcal{Z})\delta t \\ \mathcal{E}_Q^{\text{dephase}} &= (1 + \lambda\delta t)\mathcal{I} - \lambda\mathcal{Z}\delta t \\ \mathcal{E}_Q^{\text{amp}} &= (1 + \frac{1}{4}\lambda\delta t)\mathcal{I} + \lambda(-\frac{1}{2}\mathcal{X} - \frac{1}{2}\mathcal{Y} - \frac{1}{4}\mathcal{Z} + [R_{xy}] + [\pi_{xy}])\delta t. \end{aligned} \tag{G.16}$$

In the parameterised quantum circuits, the CNOT gates or more general entangling gates are prepared by cross-resonance drive, with the drive Hamiltonian

$$H = \Omega(\sigma_z^{(c)}\sigma_x^{(t)} + \gamma\mathbb{I}^{(c)}\sigma_x^{(t)} + H_\Delta) \tag{G.17}$$

where Ω is the effective qubit-qubit coupling, γ represents the effect of crosstalk between qubits and H_Δ corresponds to additional errors whose strengths can be revealed by Hamiltonian tomography. On the IBM's quantum devices, for example, H_Δ includes $\mu\sigma_z^{(c)}\mathbb{I}^{(t)}$ with μ corresponding to the drive-induced Stark-shift. In the cross-resonance drive, one dominant error is from crosstalk, and the corresponding recovery operation is

$$\mathcal{E}_Q = (1 + \lambda\delta t)I + \lambda\delta t\mathcal{X} - 2\lambda\delta t[R_x] \tag{G.18}$$

with $\lambda = \gamma\Omega$. The additional error, for example, the drive-induced Stark-shift can be mitigated by the recovery operation $\mathcal{E}_Q = (1 + \lambda\delta t)I + \lambda\delta t\mathcal{Z} - 2\lambda\delta t[R_z]$.

G.2 Hybrid error mitigation

In this section, I show how to apply the extrapolation method to mitigate model estimation error and the errors associated with imperfect recovery operations. Combined with stochastic error mitigation, I propose a hybrid error mitigation method for errors in practical NISQ devices.

G.2.1 Boosting model estimation error

I first show how to boost model estimation error, which will be used for its mitigation. Assume that the evolution of the quantum system is described by the open-system master equation

$$\frac{d}{dt}\rho_\lambda = -i[H(t), \rho_\lambda] + \lambda\mathcal{L}_{\text{exp}}[\rho_\lambda]\,. \tag{G.19}$$

The evolution of the system under a scaled Hamiltonian drive $\frac{1}{r}H\left(\frac{t}{r}\right)$ takes the form of

$$\frac{d}{dt}\rho'_\lambda = -i\left[\frac{1}{r}H\left(\frac{t}{r}\right), \rho'_\lambda\right] + \lambda\mathcal{L}_{\text{exp}}\left[\rho'_\lambda\right]. \tag{G.20}$$

Assuming the noise superoperator $\mathcal{L}$ is invariant under rescaling, we have

$$\begin{aligned}\frac{d}{dt}\rho'_\lambda(rt) &= \frac{dt'}{dt}\frac{\partial}{\partial t'}\rho'_\lambda\left(t'\right)\Big|_{t'=rt} \\ &= r\left\{-i\left[\frac{1}{r}H\left(\frac{t'}{r}\right), \rho'_\lambda\left(t'\right)\right] + \lambda\mathcal{L}\left[\rho'_\lambda\left(t'\right)\right]\right\}\Big|_{t'=rt} \\ &= -i\left[H(t), \rho'_\lambda(rt)\right] + r\lambda\mathcal{L}\left[\rho'_\lambda(rt)\right].\end{aligned} \tag{G.21}$$

On the other hand, the density matrix $\rho_{r\lambda}(t)$ with enhanced noise strength $r\lambda$ is given by

$$\frac{d}{dt}\rho_{r\lambda}(t) = -i\left[H(t), \rho_{r\lambda}(t)\right] + r\lambda\mathcal{L}\left[\rho_{r\lambda}(t)\right]. \tag{G.22}$$

Comparing Eq. (G.21) and Eq. (G.22), one finds that $\rho'_\lambda(rt)$ and $\rho_{r\lambda}(t)$ follow the same differential equation, and thus with the initial conditions $\rho'_\lambda(0) = \rho_{r\lambda}(0)$ we prove $\rho'_\lambda(rt) = \rho_{r\lambda}(t)$. This indicates that we can effectively boost physical errors of quantum systems by evolving the re-scaled Hamiltonian for time rt.

Now, I discuss how to boost the model estimation error. By applying stochastic error mitigation, we obtain

$$\frac{d}{dt}\rho_\lambda^{(Q)}(t) = -i[H(t), \rho_\lambda^{(Q)}(t)] + \lambda\Delta\mathcal{L}\big[\rho_\lambda^{(Q)}(t)\big], \tag{G.23}$$

where $\rho_\lambda^{(Q)}(t)$ is the error mitigated effective density matrix after stochastic error mitigation. Assuming $\Delta\mathcal{L} = \mathcal{L}_{\text{exp}} - \mathcal{L}_{\text{est}}$ is invariant under re-scaling of the Hamiltonian, we can similarly obtain

$$\frac{d}{dt}\rho_{r\lambda}^{(Q)}(t) = -i[H(t), \rho_{r\lambda}^{(Q)}(t)] + r\lambda\Delta\mathcal{L}\big[\rho_{r\lambda}^{(Q)}(t)\big]. \tag{G.24}$$

This can be experimentally achieved by applying stochastic error mitigation for a re-scaled time rt under the re-scaled Hamiltonian.

It is worth noting that even if the noise model is time dependent, this method can still work as long as the evolution can be described by a Lindblad equation and its dependence on time is known.

For example, consider a time dependent noisy process with stochastic error mitigation described by

$$\frac{d\rho_\lambda^{(Q)}(t)}{dt} = -i[H(t), \rho_\lambda^{(Q)}(t)] + \lambda t\Delta\mathcal{L}_0[\rho_\lambda^{(Q)}(t)], \tag{G.25}$$

where $\Delta\mathcal{L}_0$ is time independent. Then, the re-scaled dynamical equation becomes

$$\frac{d\rho_\lambda'^{(Q)}(rt)}{dt} = -i[H(t), \rho_\lambda'^{(Q)}(rt)] + r^2\lambda t\Delta\mathcal{L}_0[\rho_\lambda'^{(Q)}(rt)]. \tag{G.26}$$

In this case, we can interpret that the noise rate is boosted by a factor of r^2.

G.2.2 Richardson's extrapolation for physical errors and model estimation errors

In this section, I briefly review the extrapolation method proposed in Ref. [116, 150]. I assume the open system evolution is described by

$$\frac{d\rho_N(t)}{dt} = -i[H_{\text{sim}}(t), \rho_N(t)] + \lambda\mathcal{L}_{\text{exp}}\big[\rho_N(t)\big]. \tag{G.27}$$

In Ref. [150], it is shown that the expectation value of an observable O can be expressed as

$$\langle O(\lambda)\rangle = \langle O(0)\rangle + \sum_{k=1}^{n}\alpha_k\lambda^k + B_{n+1}(\lambda, \mathcal{L}, T), \tag{G.28}$$

where $\alpha_k \approx O(N^k T^k)$ and $B_{n+1}(\lambda, \mathcal{L}, T)$ is upper bounded by

$$B_{n+1}(\lambda, \mathcal{L}, T) \leq \|O\| a_{n+1} \frac{\lambda^{n+1} T^{n+1}}{(n+1)!}, \tag{G.29}$$

where $\|O\| = \max_\psi \langle \psi | O | \psi \rangle$ is the spectra norm of O. Here, in the case that $\mathcal{L}$ is a Lindblad type operator, one can have the bound for a_{n+1} as

$$a_{n+1} \leq \|\mathcal{L}_{\text{exp}}\|_1^{n+1}. \tag{G.30}$$

Now, we have

$$B_{n+1}(\lambda, \mathcal{L}, T) \leq \|O\| \frac{(\lambda T \|\mathcal{L}_{\text{exp}}\|_1)^{n+1}}{(n+1)!}. \tag{G.31}$$

In order to employ the extrapolation method, we need to obtain the expectation value of observable $\langle O(r_j \lambda) \rangle$ ($j = 0, 1, .., n$, $r_0 = 1$) at time $t = T$ corresponding to the equation,

$$\frac{d}{dt} \rho_\lambda(t) = -i \left[H(t), \rho_\lambda(t) \right] + r_j \lambda \mathcal{L} \left(\rho_\lambda(t) \right), \tag{G.32}$$

which can be obtained by using the re-scaling of the Hamiltonian as described in section G.2.1. Then we can obtain the approximation of the noise free expectation value of the observable O as

$$\langle O(0) \rangle_n^* = \sum_{j=0}^{n} \beta_j \langle O \rangle'_{r_j \lambda}, \tag{G.33}$$

where $\langle O(0) \rangle^*$ is the estimated noise free expectation value up to an error of order $O(\lambda^{n+1})$, and $\langle O \rangle'_{r\lambda}$ are the measurement outcome corresponding to the state $\rho_{r\lambda}(T)$. Here, the coefficients $\beta_j = \prod_{l \neq j} r_l (r_l - r_j)^{-1}$ are defined by the solution of the following equations

$$\sum_{j=0}^{n} \beta_j = 1, \sum_{j=0}^{n} \beta_j r_j^k = 0, k = 1, ..., n. \tag{G.34}$$

In Ref. [150], it has been shown that the difference between the estimator and the error free expectation value is bounded by

$$| \langle O(0) \rangle_n^* - \langle O \rangle_I | \leq \gamma_n \left(\frac{r_{\max}^{n+1} \Delta_{\max}}{\sqrt{N_{\text{sample}}}} + \|O\| \frac{(r_{\max} \lambda T \|\mathcal{L}_{\text{exp}}\|_1)^{n+1}}{(n+1)!} \right), \tag{G.35}$$

where $\gamma_n = \sum_{j=0}^{n} |\beta_j|$, $r_{\max} = \max_j r_j$, and $\Delta_{\max}/\sqrt{N_{\text{sample}}}$ is the largest experimental errors due to shot noises with N_{sample} being the number of samples. From Eq. (G.35), we can see that extrapolation methods requires

$$r_{\max} \lambda T \|\mathcal{L}_{\text{exp}}\|_1 = \mathcal{O}(1). \tag{G.36}$$

Now, under the stochastic error mitigation for a continuous process, Eq. (G.27) is modified to

$$\frac{d}{dt} \rho_\lambda^{(Q)}(t) = -i[H(t), \rho_\lambda^{(Q)}(t)] + \lambda \Delta \mathcal{L}[\rho_\lambda^{(Q)}(t)], \tag{G.37}$$

where $\Delta\mathcal{L} = \mathcal{L}_{\text{exp}} - \mathcal{L}_{\text{est}}$. Similar to the mitigation of physical errors via Richardson's extrapolation, we can obtain the approximation of the noise free expectation value of the observable O as

$$\langle O(0)\rangle_n^* = \sum_{j=0}^{n} \beta_j \langle O\rangle_{r_j\lambda}\,, \tag{G.38}$$

where $\langle O(0)\rangle^*$ is the estimated noise free expectation value up to an error of order $O(\lambda^{n+1})$, and $\langle O\rangle_{r\lambda}$ is the measurement outcome after stochastic error mitigation, corresponding to $\rho_{r\lambda}^{(Q)}(T)$.

Hence, under stochastic error mitigation, the inequality of Eq. (G.35) is modified to

$$|\langle O(0)\rangle_n^* - \langle O\rangle_I| \leq \gamma_n \left(\frac{C(r_{\max}T) r_{\max}^{n+1} \Delta_{\max}}{\sqrt{N_{\text{sample}}}} + \|O\| \frac{(r_{\max}\lambda T \|\Delta\mathcal{L}\|_1)^{n+1}}{(n+1)!} \right), \tag{G.39}$$

with Eq. (G.36) changed into

$$r_{\max}\lambda T \|\Delta\mathcal{L}\|_1 = \mathcal{O}(1). \tag{G.40}$$

Here, we used the fact that the variance of the error-mitigated expectation value of the observable is amplified with the overhead coefficient C.

From Eq. (G.39), the deviation between the ideal measurement outcome and the error-mitigated one is bounded independently with the Hamiltonian, i.e., the to-be-simulated problem. The bound only relies on the noise model, the evolution time, the number of samples, and the parameters used in extrapolation.

References

[1] F. Barahona, On the computational complexity of ising spin glass models. *Journal of Physics A: Mathematical and General* **15**, 3241 (1982).

[2] J. Kempe, A. Kitaev, O. Regev, The complexity of the local hamiltonian problem. *Siam journal on computing* **35**, 1070–1097 (2006).

[3] P. Wocjan, T. Beth, The 2-local hamiltonian problem encompasses np. *International Journal of Quantum Information* **1**, 349–357 (2003).

[4] A. Y. Kitaev, A. Shen, M. N. Vyalyi, M. N. Vyalyi, *Classical and quantum computation*, no. 47 (American Mathematical Soc., 2002).

[5] R. Oliveira, B. M. Terhal, The complexity of quantum spin systems on a two-dimensional square lattice. *arXiv preprint quant-ph/0504050* (2005).

[6] R. P. Feynman, Simulating physics with computers. *International journal of theoretical physics* **21**, 467–488 (1982).

[7] I. M. Georgescu, S. Ashhab, F. Nori, Quantum simulation. *Reviews of Modern Physics* **86**, 153 (2014).

[8] M. Kjaergaard, M. E. Schwartz, J. Braumüller, P. Krantz, J. I.-J. Wang, S. Gustavsson, W. D. Oliver, Superconducting qubits: Current state of play. *arXiv preprint arXiv:1905.13641* (2019).

[9] R. Blatt, C. F. Roos, Quantum simulations with trapped ions. *Nature Physics* **8**, 277 (2012).

[10] G. A. Quantum, *et al.*, Hartree-fock on a superconducting qubit quantum computer. *Science* **369**, 1084–1089 (2020).

[11] F. Arute, K. Arya, R. Babbush, D. Bacon, J. C. Bardin, R. Barends, R. Biswas, S. Boixo, F. G. Brandao, D. A. Buell, *et al.*, Quantum supremacy using a programmable superconducting processor. *Nature* **574**, 505–510 (2019).

[12] S. McArdle, S. Endo, A. Aspuru-Guzik, S. Benjamin, X. Yuan, Quantum computational chemistry. *arXiv preprint arXiv:1808.10402* (2018).

[13] S. Lloyd, Universal quantum simulators. *Science* pp. 1073–1078 (1996).

[14] S. W. Lovesey, Theory of neutron scattering from condensed matter (1984).

[15] A. T. Boothroyd, *Principles of Neutron Scattering from Condensed Matter* (Oxford University Press, 2020).

[16] P. Jurcevic, P. Hauke, C. Maier, C. Hempel, B. Lanyon, R. Blatt, C. F. Roos, Spectroscopy of interacting quasiparticles in trapped ions. *Physical Review Letters* **115**, 100501 (2015).

[17] C. Senko, J. Smith, P. Richerme, A. Lee, W. Campbell, C. Monroe, Coherent imaging spectroscopy of a quantum many-body spin system. *Science* **345**, 430–433 (2014).

[18] B. Yoshimura, W. Campbell, J. Freericks, Diabatic-ramping spectroscopy of many-body excited states. *Physical Review A* **90**, 062334 (2014).

[19] L. Villa, J. Despres, L. Sanchez-Palencia, Unraveling the excitation spectrum of many-body systems from quantum quenches. *Physical Review A* **100**, 063632 (2019).

[20] L. Villa, J. Despres, S. Thomson, L. Sanchez-Palencia, Local quench spectroscopy of many-body quantum systems. *Physical Review A* **102**, 033337 (2020).

[21] L. Villa, Out-of-equilibrium dynamics and quench spectroscopy of ultracold many-body quantum systems, Ph.D. thesis, Institut polytechnique de Paris (2021).

[22] S. Lloyd, M. Mohseni, P. Rebentrost, Quantum principal component analysis. *Nature Physics* **10**, 631 (2014).

[23] X. Xu, J. Sun, S. Endo, Y. Li, S. C. Benjamin, X. Yuan, Variational algorithms for linear algebra. *Science Bulletin* (2021).

[24] H.-Y. Huang, K. Bharti, P. Rebentrost, Near-term quantum algorithms for linear systems of equations. *arXiv preprint arXiv:1909.07344* (2019).

[25] K. Bharti, A. Cervera-Lierta, T. H. Kyaw, T. Haug, S. Alperin-Lea, A. Anand, M. Degroote, H. Heimonen, J. S. Kottmann, T. Menke, *et al.*, Noisy intermediate-scale quantum (nisq) algorithms. *arXiv preprint arXiv:2101.08448* (2021).

[26] J. Watrous, *Proceedings 41st Annual Symposium on Foundations of Computer Science* (IEEE, 2000), pp. 537–546.

[27] M. A. Nielsen, I. Chuang, Quantum computation and quantum information (2002).

[28] A. Y. Kitaev, Quantum measurements and the abelian stabilizer problem. *arXiv preprint quant-ph/9511026* (1995).

[29] Y. Ge, J. Tura, J. I. Cirac, Faster ground state preparation and high-precision ground energy estimation with fewer qubits. *Journal of Mathematical Physics* **60**, 022202 (2019).

[30] P. A. Casares, R. Campos, M. A. Martin-Delgado, Tfermion: A non-clifford gate cost assessment library of quantum phase estimation algorithms for quantum chemistry. *Quantum* **6**, 768 (2022).

[31] L. Lin, Y. Tong, Near-optimal ground state preparation. *Quantum* **4**, 372 (2020).

[32] G. H. Low, I. L. Chuang, Optimal hamiltonian simulation by quantum signal processing. *Phys. Rev. Lett.* **118**, 010501 (2017).

[33] G. H. Low, I. L. Chuang, Hamiltonian simulation by qubitization. *Quantum* **3**, 163 (2019).

[34] A. Gilyén, Y. Su, G. H. Low, N. Wiebe, *Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing* (2019), pp. 193–204.

[35] D. W. Berry, A. M. Childs, R. Kothari, *2015 IEEE 56th Annual Symposium on Foundations of Computer Science* (2015), pp. 792–809.

[36] A. M. Childs, N. Wiebe, Hamiltonian simulation using linear combinations of unitary operations. *Quantum Information and Computation* **12** (2012).

[37] A. M. Childs, D. Maslov, Y. Nam, N. J. Ross, Y. Su, Toward the first quantum simulation with quantum speedup. *Proceedings of the National Academy of Sciences* **115**, 9456–9461 (2018).

[38] D. W. Berry, A. M. Childs, R. Cleve, R. Kothari, R. D. Somma, Simulating hamiltonian dynamics with a truncated taylor series. *Physical review letters* **114**, 090502 (2015).

[39] T. Keen, E. Dumitrescu, Y. Wang, Quantum algorithms for ground-state preparation and green's function calculation. *arXiv preprint arXiv:2112.05731* (2021).

[40] R. O. Jones, Density functional theory: Its origins, rise to prominence, and future. *Reviews of modern physics* **87**, 897 (2015).

[41] S. P. Jordan, K. S. Lee, J. Preskill, Quantum algorithms for quantum field theories. *Science* **336**, 1130–1133 (2012).

[42] J. Liu, Z. Li, H. Zheng, X. Yuan, J. Sun, Towards a variational jordan–lee–preskill quantum algorithm. *Machine Learning: Science and Technology* **3**, 045030 (2022).

[43] R. Babbush, N. Wiebe, J. McClean, J. McClain, H. Neven, G. K.-L. Chan, Low-depth quantum simulation of materials. *Physical Review X* **8**, 011044 (2018).

[44] L. Clinton, T. Cubitt, B. Flynn, F. M. Gambetta, J. Klassen, A. Montanaro, S. Piddock, R. A. Santos, E. Sheridan, Towards near-term quantum simulation of materials. *arXiv preprint arXiv:2205.15256* (2022).

[45] A. Kitaev, W. A. Webb, Wavefunction preparation and resampling using a quantum computer. *arXiv preprint arXiv:0801.0342* (2008).

[46] H. M. Wiseman, R. B. Killip, Adaptive single-shot phase measurements: A semiclassical approach. *Physical Review A* **56**, 944 (1997).

[47] B. Sanders, G. Milburn, Optimal quantum measurements for phase estimation. *Physical review letters* **75**, 2944 (1995).

[48] L. Lin, Y. Tong, Heisenberg-limited ground-state energy estimation for early fault-tolerant quantum computers. *PRX Quantum* **3**, 010318 (2022).

[49] N. M. Linke, D. Maslov, M. Roetteler, S. Debnath, C. Figgatt, K. A. Landsman, K. Wright, C. Monroe, Experimental comparison of two quantum computing architectures. *Proceedings of the National Academy of Sciences* **114**, 3305–3310 (2017).

[50] J. G. Bohnet, B. C. Sawyer, J. W. Britton, M. L. Wall, A. M. Rey, M. Foss-Feig, J. J. Bollinger, Quantum spin dynamics and entanglement generation with hundreds of trapped ions. *Science* **352**, 1297–1301 (2016).

[51] A. Friedenauer, H. Schmitz, J. T. Glueckert, D. Porras, T. Schätz, Simulating a quantum magnet with trapped ions. *Nature Physics* **4**, 757 (2008).

[52] P. Jurcevic, H. Shen, P. Hauke, C. Maier, T. Brydges, C. Hempel, B. Lanyon, M. Heyl, R. Blatt, C. Roos, Direct observation of dynamical quantum phase transitions in an interacting many-body system. *Physical review letters* **119**, 080501 (2017).

[53] R. A. Hart, P. M. Duarte, T.-L. Yang, X. Liu, T. Paiva, E. Khatami, R. T. Scalettar, N. Trivedi, D. A. Huse, R. G. Hulet, Observation of antiferromagnetic correlations in the hubbard model with ultracold atoms. *Nature* **519**, 211 (2015).

[54] H. Bernien, S. Schwartz, A. Keesling, H. Levine, A. Omran, H. Pichler, S. Choi, A. S. Zibrov, M. Endres, M. Greiner, *et al.*, Probing many-body dynamics on a 51-atom quantum simulator. *Nature* **551**, 579 (2017).

[55] J. Li, R. Fan, H. Wang, B. Ye, B. Zeng, H. Zhai, X. Peng, J. Du, Measuring out-of-time-order correlators on a nuclear magnetic resonance quantum simulator. *Physical Review X* **7**, 031011 (2017).

[56] Z. Li, H. Zhou, C. Ju, H. Chen, W. Zheng, D. Lu, X. Rong, C. Duan, X. Peng, J. Du, Experimental realization of a compressed quantum simulation of a 32-spin ising chain. *Physical review letters* **112**, 220501 (2014).

[57] M. Gärttner, J. G. Bohnet, A. Safavi-Naini, M. L. Wall, J. J. Bollinger, A. M. Rey, Measuring out-of-time-order correlations and multiple quantum spectra in a trapped-ion quantum magnet. *Nature Physics* **13**, 781 (2017).

[58] J. Struck, C. Ölschläger, R. Le Targat, P. Soltan-Panahi, A. Eckardt, M. Lewenstein, P. Windpassinger, K. Sengstock, Quantum simulation of frustrated classical magnetism in triangular optical lattices. *Science* **333**, 996–999 (2011).

[59] P. Jurcevic, B. P. Lanyon, P. Hauke, C. Hempel, P. Zoller, R. Blatt, C. F. Roos, Quasiparticle engineering and entanglement propagation in a quantum many-body system. *Nature* **511**, 202–205 (2014).

[60] N. Wiebe, C. Granade, C. Ferrie, D. G. Cory, Hamiltonian learning and certification using quantum resources. *Physical review letters* **112**, 190501 (2014).

[61] J. Wang, S. Paesani, R. Santagati, S. Knauer, A. A. Gentile, N. Wiebe, M. Petruzzella, J. L. OBrien, J. G. Rarity, A. Laing, *et al.*, Experimental quantum hamiltonian learning. *Nature Physics* **13**, 551–555 (2017).

[62] N. Wiebe, C. Granade, C. Ferrie, D. Cory, Quantum hamiltonian learning using imperfect quantum resources. *Physical Review A* **89**, 042314 (2014).

[63] M. Mohseni, A. Rezakhani, D. Lidar, Quantum-process tomography: Resource analysis of different strategies. *Physical Review A* **77**, 032322 (2008).

[64] W. Yu, J. Sun, Z. Han, X. Yuan, Practical and efficient hamiltonian learning. *arXiv preprint arXiv:2201.00190* (2022).

[65] S. McArdle, Learning from physics experiments with quantum computers: Applications in muon spectroscopy. *PRX Quantum* **2**, 020349 (2021).

[66] T. E. O'Brien, L. B. Ioffe, Y. Su, D. Fushman, H. Neven, R. Babbush, V. Smelyanskiy, Quantum computation of molecular structure using data from challenging-to-classically-simulate nuclear magnetic resonance experiments. *arXiv preprint arXiv:2109.02163* (2021).

[67] D. Sels, H. Dashti, S. Mora, O. Demler, E. Demler, Quantum approximate bayesian computation for nmr model inference. *Nature machine intelligence* **2**, 396–402 (2020).

[68] E. T. Campbell, B. M. Terhal, C. Vuillot, Roads towards fault-tolerant universal quantum computation. *Nature* **549**, 172–179 (2017).

[69] S. Guo, J. Sun, H. Qian, M. Gong, Y. Zhang, F. Chen, Y. Ye, Y. Wu, S. Cao, K. Liu, *et al.*, Scalable quantum computational chemistry with superconducting qubits. *arXiv preprint arXiv:2212.08006* (2022).

[70] C. Neill, T. McCourt, X. Mi, Z. Jiang, M. Niu, W. Mruczkiewicz, I. Aleiner, F. Arute, K. Arya, J. Atalaya, *et al.*, Accurately computing the electronic properties of a quantum ring. *Nature* **594**, 508–512 (2021).

[71] J. Preskill, Quantum computing in the nisq era and beyond. *Quantum* **2**, 79 (2018).

[72] J. R. McClean, J. Romero, R. Babbush, A. Aspuru-Guzik, The theory of variational hybrid quantum-classical algorithms. *New Journal of Physics* **18**, 023023 (2016).

[73] A. Peruzzo, J. McClean, P. Shadbolt, M.-H. Yung, X.-Q. Zhou, P. J. Love, A. Aspuru-Guzik, J. L. Obrien, A variational eigenvalue solver on a photonic quantum processor. *Nature communications* **5**, 1–7 (2014).

[74] A. Kandala, A. Mezzacapo, K. Temme, M. Takita, M. Brink, J. M. Chow, J. M. Gambetta, Hardware-efficient variational quantum eigensolver for small molecules and quantum magnets. *Nature* **549**, 242–246 (2017).

[75] Y. Wu, *et al.*, Strong quantum computational advantage using a superconducting quantum processor. *Phys. Rev. Lett.* **127**, 180501 (2021).

[76] S. Ebadi, A. Keesling, M. Cain, T. T. Wang, H. Levine, D. Bluvstein, G. Semeghini, A. Omran, J.-G. Liu, R. Samajdar, *et al.*, Quantum optimization of maximum independent set using rydberg atom arrays. *Science* p. eabo6587 (2022).

[77] C. Huang, F. Zhang, M. Newman, J. Cai, X. Gao, Z. Tian, J. Wu, H. Xu, H. Yu, B. Yuan, *et al.*, Classical simulation of quantum supremacy circuits. *arXiv preprint arXiv:2005.06787* (2020).

[78] F. Pan, P. Zhang, Simulation of quantum circuits using the big-batch tensor network method. *Physical Review Letters* **128**, 030501 (2022).

[79] H.-S. Zhong, H. Wang, Y.-H. Deng, M.-C. Chen, L.-C. Peng, Y.-H. Luo, J. Qin, D. Wu, X. Ding, Y. Hu, *et al.*, Quantum computational advantage using photons. *Science* **370**, 1460–1463 (2020).

[80] S. Endo, J. Sun, Y. Li, S. C. Benjamin, X. Yuan, Variational quantum simulation of general processes. *Physical Review Letters* **125**, 010501 (2020).

[81] S. Endo, Z. Cai, S. C. Benjamin, X. Yuan, Hybrid quantum-classical algorithms and quantum error mitigation. *arXiv preprint arXiv:2011.01382* (2020).

[82] M. Cerezo, A. Arrasmith, R. Babbush, S. C. Benjamin, S. Endo, K. Fujii, J. R. McClean, K. Mitarai, X. Yuan, L. Cincio, P. J. Coles, Variational quantum algorithms. *arXiv preprint arXiv:2012.09265* (2020).

[83] J. R. McClean, M. E. Kimchi-Schwartz, J. Carter, W. A. de Jong, Hybrid quantum-classical hierarchy for mitigation of decoherence and determination of excited states. *Physical Review A* **95**, 042308 (2017).

[84] A. Callison, N. Chancellor, Hybrid quantum-classical algorithms in the noisy intermediate-scale quantum era and beyond. *Physical Review A* **106**, 010101 (2022).

[85] M. Cerezo, A. Arrasmith, R. Babbush, S. C. Benjamin, S. Endo, K. Fujii, J. R. McClean, K. Mitarai, X. Yuan, L. Cincio, P. J. Coles, Variational quantum algorithms (2020).

[86] Y. Du, Z. Tu, X. Yuan, D. Tao, An efficient measure for the expressivity of variational quantum algorithms (2021).

[87] B. Wu, J. Sun, Q. Huang, X. Yuan, Overlapped grouping measurement: A unified framework for measuring quantum states. *Quantum* **7**, 896 (2023).

[88] L. Bittel, M. Kliesch, Training variational quantum algorithms is np-hard. *Physical Review Letters* **127**, 120502 (2021).

[89] D. Stilck França, R. Garcia-Patron, Limitations of optimization algorithms on noisy quantum devices. *Nature Physics* **17**, 1221–1227 (2021).

[90] X. Yuan, J. Sun, J. Liu, Q. Zhao, Y. Zhou, Quantum simulation with hybrid tensor networks. *Phys. Rev. Lett.* **127**, 040501 (2021).

[91] J. Sun, X. Yuan, V. Vedral, Quantum dynamics simulation with quantum tensor networks, to appear.

[92] J. Sun, P. Zeng, V. Vedral, High precision eigenenergy estimation and eigenstate property estimation: a framework and resource estimation, to appear.

[93] P. Zeng, J. Sun, X. Yuan, Universal quantum algorithmic cooling on a quantum computer (2021).

[94] Z.-J. Zhang, J. Sun, X. Yuan, M.-H. Yung, Low-depth hamiltonian simulation by an adaptive product formula. *Physical Review Letters* **130**, 040601 (2023).

[95] J. Sun, S. Endo, H. Lin, P. Hayden, V. Vedral, X. Yuan, Perturbative quantum simulation (2021).

[96] J. Sun, Lucia Vilchez Estevez, V. Vedral, A. T. Boothroyd, Probing spectroscopic features of quantum many-body systems on a quantum simulator, to appear.

[97] J. Sun, X. Yuan, T. Tsunoda, V. Vedral, S. C. Benjamin, S. Endo, Mitigating realistic noise in practical noisy intermediate-scale quantum devices. *Physical Review Applied* **15**, 034026 (2021).

[98] T. Zhang, J. Sun, X.-X. Fang, X.-M. Zhang, X. Yuan, H. Lu, Experimental quantum state measurement with classical shadows. *Phys. Rev. Lett.* **127**, 200501 (2021).

[99] D. W. Berry, C. Gidney, M. Motta, J. R. McClean, R. Babbush, Qubitization of arbitrary basis quantum chemistry leveraging sparsity and low rank factorization. *Quantum* **3**, 208 (2019).

[100] M. Motta, E. Ye, J. R. McClean, Z. Li, A. J. Minnich, R. Babbush, G. K. Chan, Low rank representations for quantum simulation of electronic structure. *npj Quantum Information* **7**, 1–7 (2021).

[101] J. Lee, D. W. Berry, C. Gidney, W. J. Huggins, J. R. McClean, N. Wiebe, R. Babbush, Even more efficient quantum computations of chemistry through tensor hypercontraction. *PRX Quantum* **2**, 030305 (2021).

[102] F. Verstraete, J. I. Cirac, Matrix product states represent ground states faithfully. *Physical review b* **73**, 094423 (2006).

[103] M. B. Hastings, An area law for one-dimensional quantum systems. *Journal of statistical mechanics: theory and experiment* **2007**, P08024 (2007).

[104] M. B. Hastings, Solving gapped hamiltonians locally. *Physical review b* **73**, 085115 (2006).

[105] R. Orús, Tensor networks for complex quantum systems. *Nature Reviews Physics* **1**, 538–550 (2019).

[106] R. Balian, M. Vénéroni, Static and dynamic variational principles for expectation values of observables. *Annals of Physics* **187**, 29 - 78 (1988).

[107] F. Dalfovo, S. Giorgini, L. P. Pitaevskii, S. Stringari, Theory of bose-einstein condensation in trapped gases. *Rev. Mod. Phys.* **71**, 463–512 (1999).

[108] J. Haegeman, J. I. Cirac, T. J. Osborne, I. Pižorn, H. Verschelde, F. Verstraete, Time-dependent variational principle for quantum lattices. *Phys. Rev. Lett.* **107**, 070601 (2011).

[109] T. Shi, E. Demler, J. I. Cirac, Variational study of fermionic and bosonic systems with non-gaussian states: Theory and applications. *Annals of Physics* **390**, 245 - 302 (2018).

[110] L. Vanderstraeten, J. Haegeman, F. Verstraete, Tangent-space methods for uniform matrix product states. *SciPost Phys. Lect. Notes* **7** (2019).

[111] E. Farhi, J. Goldstone, S. Gutmann, A quantum approximate optimization algorithm (2014).

[112] Y. Wang, F. Dolde, J. Biamonte, R. Babbush, V. Bergholm, S. Yang, I. Jakobi, P. Neumann, A. Aspuru-Guzik, J. D. Whitfield, *et al.*, Quantum simulation of helium hydride cation in a solid-state spin register. *ACS nano* **9**, 7769–7774 (2015).

[113] P. J. J. O'Malley, R. Babbush, I. D. Kivlichan, J. Romero, J. R. McClean, R. Barends, J. Kelly, P. Roushan, A. Tranter, N. Ding, B. Campbell, Y. Chen, Z. Chen, B. Chiaro, A. Dunsworth, A. G. Fowler, E. Jeffrey, E. Lucero, A. Megrant, J. Y. Mutus, M. Neeley, C. Neill, C. Quintana, D. Sank, A. Vainsencher, J. Wenner, T. C. White, P. V. Coveney, P. J. Love, H. Neven, A. Aspuru-Guzik, J. M. Martinis, Scalable quantum simulation of molecular energies. *Phys. Rev. X* **6**, 031007 (2016).

[114] Y. Shen, X. Zhang, S. Zhang, J.-N. Zhang, M.-H. Yung, K. Kim, Quantum implementation of the unitary coupled cluster for simulating molecular electronic structure. *Phys. Rev. A* **95**, 020501 (2017).

[115] S. Paesani, A. A. Gentile, R. Santagati, J. Wang, N. Wiebe, D. P. Tew, J. L. O'Brien, M. G. Thompson, Experimental bayesian quantum phase estimation on a silicon photonic chip. *Phys. Rev. Lett.* **118**, 100503 (2017).

[116] Y. Li, S. C. Benjamin, Efficient variational quantum simulator incorporating active error minimization. *Physical Review X* **7**, 021050 (2017).

[117] J. I. Colless, V. V. Ramasesh, D. Dahlen, M. S. Blok, M. E. Kimchi-Schwartz, J. R. McClean, J. Carter, W. A. de Jong, I. Siddiqi, Computation of molecular spectra on a quantum processor with an error-resilient algorithm. *Phys. Rev. X* **8**, 011021 (2018).

[118] C. Hempel, C. Maier, J. Romero, J. McClean, T. Monz, H. Shen, P. Jurcevic, B. P. Lanyon, P. Love, R. Babbush, A. Aspuru-Guzik, R. Blatt, C. F. Roos, Quantum chemistry calculations on a trapped-ion quantum simulator. *Phys. Rev. X* **8**, 031022 (2018).

[119] J. Romero, R. Babbush, J. R. McClean, C. Hempel, P. J. Love, A. Aspuru-Guzik, Strategies for quantum computing molecular energies using the unitary coupled cluster ansatz. *Quantum Science and Technology* **4**, 014008 (2018).

[120] S. McArdle, T. Jones, S. Endo, Y. Li, S. C. Benjamin, X. Yuan, Variational ansatz-based quantum simulation of imaginary time evolution. *npj Quantum Information* **5**, 1–6 (2019).

[121] T. Jones, S. Endo, S. McArdle, X. Yuan, S. C. Benjamin, Variational quantum algorithms for discovering hamiltonian spectra. *Physical Review A* **99**, 062304 (2019).

[122] O. Higgott, D. Wang, S. Brierley, Variational quantum computation of excited states. *Quantum* **3**, 156 (2019).

[123] J. I. Colless, V. V. Ramasesh, D. Dahlen, M. S. Blok, M. Kimchi-Schwartz, J. McClean, J. Carter, W. De Jong, I. Siddiqi, Computation of molecular spectra on a quantum processor with an error-resilient algorithm. *Physical Review X* **8**, 011021 (2018).

[124] C. Kokail, C. Maier, R. van Bijnen, T. Brydges, M. K. Joshi, P. Jurcevic, C. A. Muschik, P. Silvi, R. Blatt, C. F. Roos, *et al.*, Self-verifying variational quantum simulation of lattice models. *Nature* **569**, 355–360 (2019).

[125] K. Heya, K. M. Nakanishi, K. Mitarai, K. Fujii, Subspace variational quantum simulator. *arXiv preprint arXiv:1904.08566* (2019).

[126] K. M. Nakanishi, K. Mitarai, K. Fujii, Subspace-search variational quantum eigensolver for excited states. *Physical Review Research* **1**, 033062 (2019).

[127] I. Kassal, J. D. Whitfield, A. Perdomo-Ortiz, M.-H. Yung, A. Aspuru-Guzik, Simulating chemistry using quantum computers. *Annual review of physical chemistry* **62**, 185–207 (2011).

[128] D. Lu, B. Xu, N. Xu, Z. Li, H. Chen, X. Peng, R. Xu, J. Du, Quantum chemistry simulation on quantum computers: theories and experiments. *Phys. Chem. Chem. Phys.* **14**, 9411-9420 (2012).

[129] K. B. Whaley, A. R. Dinner, S. A. Rice, *Quantum information and computation for chemistry* (John Wiley & Sons, 2014).

[130] S. McArdle, S. Endo, A. Aspuru-Guzik, S. C. Benjamin, X. Yuan, Quantum computational chemistry. *Reviews of Modern Physics* **92**, 015003 (2020).

[131] X. Yuan, S. Endo, Q. Zhao, Y. Li, S. C. Benjamin, Theory of variational quantum simulation. *Quantum* **3**, 191 (2019).

[132] J. Gacon, C. Zoufal, G. Carleo, S. Woerner, Simultaneous perturbation stochastic approximation of the quantum fisher information. *Quantum* **5**, 567 (2021).

[133] B. Koczor, S. C. Benjamin, Quantum natural gradient generalised to noisy and non-unitary circuits. *arXiv preprint arXiv:1912.08660* (2019).

[134] B. van Straaten, B. Koczor, Measurement cost of metric-aware variational quantum algorithms. *PRX Quantum* **2**, 030324 (2021).

[135] J. Stokes, J. Izaac, N. Killoran, G. Carleo, Quantum natural gradient. *Quantum* **4**, 269 (2020).

[136] S. McArdle, S. Endo, T. Jones, Y. Li, S. Benjamin, X. Yuan, Variational quantum simulation of imaginary time evolution with applications in chemistry and beyond. *arXiv preprint arXiv:1804.03023* (2018).

[137] A. McLachlan, A variational solution of the time-dependent schrodinger equation. *Molecular Physics* **8**, 39-44 (1964).

[138] K. Mitarai, K. Fujii, Methodology for replacing indirect measurements with direct measurements. *Physical Review Research* **1**, 013006 (2019).

[139] M. Suzuki, General theory of fractal path integrals with applications to many-body theories and statistical physics. *Journal of Mathematical Physics* **32**, 400–407 (1991).

[140] A. M. Childs, A. Ostrander, Y. Su, Faster quantum simulation by randomization. *Quantum* **3**, 182 (2019).

[141] T. Albash, D. A. Lidar, Adiabatic quantum computation. *Rev. Mod. Phys.* **90**, 015002 (2018).

[142] S. Lu, M. C. Bañuls, J. I. Cirac, Algorithms for quantum simulation at finite energies. *PRX Quantum* **2**, 020321 (2021).

[143] Y. R. Sanders, D. W. Berry, P. C. Costa, L. W. Tessler, N. Wiebe, C. Gidney, H. Neven, R. Babbush, Compilation of fault-tolerant quantum heuristics for combinatorial optimization. *PRX Quantum* **1**, 020312 (2020).

[144] A. M. Childs, Y. Su, M. C. Tran, N. Wiebe, S. Zhu, Theory of trotter error with commutator scaling. *Phys. Rev. X* **11**, 011020 (2021).

[145] E. Campbell, Random compiler for fast hamiltonian simulation. *Physical review letters* **123**, 070503 (2019).

[146] Y. Su, H.-Y. Huang, E. T. Campbell, Nearly tight trotterization of interacting electrons. *Quantum* **5**, 495 (2021).

[147] M. Heyl, P. Hauke, P. Zoller, Quantum localization bounds trotter errors in digital quantum simulation. *Science advances* **5**, eaau8342 (2019).

[148] C.-F. Chen, R. Kueng, J. A. Tropp, *et al.*, Quantum simulation via randomized product formulas: Low gate complexity with accuracy guarantees. *arXiv preprint arXiv:2008.11751* (2020).

[149] S. Endo, S. C. Benjamin, Y. Li, Practical quantum error mitigation for near-future applications. *Physical Review X* **8**, 031027 (2018).

[150] K. Temme, S. Bravyi, J. M. Gambetta, Error mitigation for short-depth quantum circuits. *Physical review letters* **119**, 180509 (2017).

[151] S. R. White, Density matrix formulation for quantum renormalization groups. *Phys. Rev. Lett.* **69**, 2863–2866 (1992).

[152] S. Rommer, S. Östlund, Class of ansatz wave functions for one-dimensional spin systems and their relation to the density matrix renormalization group. *Phys. Rev. B* **55**, 2164–2181 (1997).

[153] U. Schollwöck, The density-matrix renormalization group in the age of matrix product states. *Annals of Physics* **326**, 96–192 (2011).

[154] E. Garlatti, T. Guidi, S. Ansbro, P. Santini, G. Amoretti, J. Ollivier, H. Mutka, G. Timco, I. Vitorica-Yrezabal, G. Whitehead, *et al.*, Portraying entanglement between molecular qubits with four-dimensional inelastic neutron scattering. *Nature communications* **8**, 1–7 (2017).

[155] G. A. Timco, S. Carretta, F. Troiani, F. Tuna, R. J. Pritchard, C. A. Muryn, E. J. McInnes, A. Ghirri, A. Candini, P. Santini, *et al.*, Engineering the coupling between molecular spin qubits by coordination chemistry. *Nature Nanotechnology* **4**, 173–178 (2009).

[156] T. Takeshita, N. C. Rubin, Z. Jiang, E. Lee, R. Babbush, J. R. McClean, Increasing the representation accuracy of quantum simulations of chemistry without extra quantum resources. *Phys. Rev. X* **10**, 011004 (2020).

[157] Y. Gu, X.-L. Qi, D. Stanford, Local criticality, diffusion and chaos in generalized Sachdev-Ye-Kitaev models. *JHEP* **05**, 125 (2017).

[158] M. Lubasch, J. I. Cirac, M.-C. Banuls, Unifying projected entangled pair state contractions. *New Journal of Physics* **16**, 033014 (2014).

[159] A. A. Abrikosov, L. P. Gorkov, I. E. Dzyaloshinski, *Methods of quantum field theory in statistical physics* (Courier Corporation, 2012).

[160] P.-O. Löwdin, Studies in perturbation theory. x. lower bounds to energy eigenvalues in perturbation-theory ground state. *Phys. Rev.* **139**, A357–A372 (1965).

[161] A. Szabo, N. S. Ostlund, *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory* (Courier Corporation, 2012).

[162] R. Orús, Tensor networks for complex quantum systems. *Nature Reviews Physics* **1**, 538-550 (2019).

[163] D. Sholl, J. A. Steckel, *Density Functional Theory: a Practical Introduction* (John Wiley & Sons, 2011).

[164] G. Knizia, G. K.-L. Chan, Density matrix embedding: A simple alternative to dynamical mean-field theory. *Phys. Rev. Lett.* **109**, 186404 (2012).

[165] G. Carleo, M. Troyer, Solving the quantum many-body problem with artificial neural networks. *Science* **355**, 602-606 (2017).

[166] J. Hermann, Z. Schätzle, F. Noé, Deep-neural-network solution of the electronic schrödinger equation. *Nature Chemistry* **12**, 891-897 (2020).

[167] D. Pfau, J. S. Spencer, A. G. D. G. Matthews, W. M. C. Foulkes, Ab initio solution of the many-electron schrödinger equation with deep neural networks. *Phys. Rev. Research* **2**, 033429 (2020).

[168] N. Metropolis, S. Ulam, The monte carlo method. *Journal of the American Statistical Association* **44**, 335-341 (1949). PMID: 18139350.

[169] M. Troyer, U.-J. Wiese, Computational complexity and fundamental limitations to fermionic quantum monte carlo simulations. *Phys. Rev. Lett.* **94**, 170201 (2005).

[170] M. Marvian, D. A. Lidar, I. Hen, On the computational complexity of curing non-stoquastic hamiltonians. *Nature Communications* **10**, 1571 (2019).

[171] J. Kolorenč, L. Mitas, Applications of quantum monte carlo methods in condensed systems. *Reports on Progress in Physics* **74**, 026502 (2011).

[172] B. M. Austin, D. Y. Zubarev, W. A. Lester, Quantum monte carlo and related approaches. *Chemical Reviews* **112**, 263-288 (2012).

[173] S. Oh, Quantum computational method of finding the ground-state energy and expectation values. *Phys. Rev. A* **77**, 012326 (2008).

[174] T. Koma, B. Nachtergaele, The spectral gap of the ferromagnetic xxz-chain. *Letters in Mathematical Physics* **40**, 1–16 (1997).

[175] I. Buluta, F. Nori, Quantum simulators. *Science* **326**, 108–111 (2009).

[176] R. Lewis-Swan, A. Safavi-Naini, A. Kaufman, A. Rey, Dynamics of quantum information. *Nature Reviews Physics* **1**, 627–634 (2019).

[177] M. Motta, C. Sun, A. T. Tan, M. J. ORourke, E. Ye, A. J. Minnich, F. G. Brandão, G. K.-L. Chan, Determining eigenstates and thermal states on a quantum computer using quantum imaginary time evolution. *Nature Physics* **16**, 205–210 (2020).

[178] S.-N. Sun, M. Motta, R. N. Tazhigulov, A. T. Tan, G. K.-L. Chan, A. J. Minnich, Quantum computation of finite-temperature static and dynamical properties of spin systems using quantum imaginary time evolution. *PRX Quantum* **2**, 010317 (2021).

[179] A. Aspuru-Guzik, A. D. Dutoi, P. J. Love, M. Head-Gordon, Simulated quantum computation of molecular energies. *Science* **309**, 1704–1707 (2005).

[180] N. H. Stair, R. Huang, F. A. Evangelista, A multireference quantum krylov algorithm for strongly correlated electrons. *Journal of Chemical Theory and Computation* **16**, 2236–2245 (2020).

[181] R. M. Parrish, P. L. McMahon, Quantum filter diagonalization: Quantum eigendecomposition without full quantum phase estimation. *arXiv preprint arXiv:1909.08925* (2019).

[182] R. D. Somma, Quantum eigenvalue estimation via time series analysis. *New Journal of Physics* **21**, 123025 (2019).

[183] O. Kyriienko, Quantum inverse iteration algorithm for programmable quantum simulators. *npj Quantum Information* **6**, 1–8 (2020).

[184] P. Zeng, J. Sun, X. Yuan, Universal quantum algorithmic cooling on a quantum computer. *arXiv e-prints* p. arXiv:2109.15304 (2021).

[185] M. Suzuki, Decomposition formulas of exponential operators and lie exponentials with some applications to quantum mechanics and statistical physics. *Journal of mathematical physics* **26**, 601–612 (1985).

[186] D. W. Berry, A. M. Childs, Y. Su, X. Wang, N. Wiebe, Time-dependent hamiltonian simulation with $l-1$-norm scaling. *Quantum* **4**, 254 (2020).

[187] I. D. Kivlichan, C. Gidney, D. W. Berry, N. Wiebe, J. McClean, W. Sun, Z. Jiang, N. Rubin, A. Fowler, A. Aspuru-Guzik, *et al.*, Improved fault-tolerant quantum simulation of condensed-phase correlated electrons via trotterization. *Quantum* **4**, 296 (2020).

[188] E. T. Campbell, Early fault-tolerant simulations of the hubbard model. *Quantum Science and Technology* **7**, 015007 (2021).

[189] A. M. Childs, Y. Su, Nearly optimal lattice simulation by product formulas. *Physical review letters* **123**, 050503 (2019).

[190] Q. Zhao, Y. Zhou, A. F. Shaw, T. Li, A. M. Childs, Hamiltonian simulation with random inputs. *arXiv e-prints* p. arXiv:2111.04773 (2021).

[191] X. Gu, A. F. Kockum, A. Miranowicz, Y.-x. Liu, F. Nori, Microwave photonics with superconducting quantum circuits. *Physics Reports* **718**, 1–102 (2017).

[192] P. Krantz, M. Kjaergaard, F. Yan, T. P. Orlando, S. Gustavsson, W. D. Oliver, A quantum engineer's guide to superconducting qubits. *Applied Physics Reviews* **6**, 021318 (2019).

[193] A. Kandala, K. Temme, A. D. Córcoles, A. Mezzacapo, J. M. Chow, J. M. Gambetta, Error mitigation extends the computational reach of a noisy quantum processor. *Nature* **567**, 491 (2019).

[194] H. R. Grimsley, S. E. Economou, E. Barnes, N. J. Mayhall, An adaptive variational algorithm for exact molecular simulations on a quantum computer. *Nature comm.* **10**, 1–9 (2019).

[195] H. L. Tang, V. Shkolnikov, G. S. Barron, H. R. Grimsley, N. J. Mayhall, E. Barnes, S. E. Economou, qubit-adapt-vqe: An adaptive algorithm for constructing hardware-efficient ansätze on a quantum processor. *PRX Quantum* **2**, 020310 (2021).

[196] Y.-X. Yao, N. Gomes, F. Zhang, C.-Z. Wang, K.-M. Ho, T. Iadecola, P. P. Orth, Adaptive variational quantum dynamics simulations. *PRX Quantum* **2**, 030307 (2021).

[197] Z. Zhang, J. Sun, under preparation.

[198] V. Verteletskyi, T.-C. Yen, A. F. Izmaylov, Measurement optimization in the variational quantum eigensolver using a minimum clique cover. *The Journal of Chemical Physics* **152**, 124114 (2020).

[199] H.-Y. Huang, R. Kueng, J. Preskill, Predicting many properties of a quantum system from very few measurements. *arXiv preprint arXiv:2002.08953* (2020).

[200] C. Hadfield, S. Bravyi, R. Raymond, A. Mezzacapo, Measurements of quantum hamiltonians with locally-biased classical shadows. *arXiv preprint arXiv:2006.15788* (2020).

[201] X. Bonet-Monroig, R. Babbush, T. E. OBrien, Nearly optimal measurement scheduling for partial tomography of quantum states. *Physical Review X* **10**, 031064 (2020).

[202] M. Rigol, T. Bryant, R. R. Singh, Numerical linked-cluster algorithms. i. spin systems on square, triangular, and kagomé lattices. *Phys. Rev. E* **75**, 061118 (2007).

[203] B. Bauer, S. Bravyi, M. Motta, G. K.-L. Chan, Quantum algorithms for quantum chemistry and quantum materials science. *Chemical Reviews* **120**, 12685–12717 (2020).

[204] E. Fradkin, *Field theories of condensed matter physics* (Cambridge University Press, 2013).

[205] M. Rigol, T. Bryant, R. R. Singh, Numerical linked-cluster approach to quantum lattice models. *Phys. Rev. Lett.* **97**, 187202 (2006).

[206] S. Wilson, Diagrammatic many-body perturbation theory of atomic and molecular electronic structure. *Computer Physics Reports* **2**, 391–480 (1985).

[207] G. Rohringer, H. Hafermann, A. Toschi, A. Katanin, A. Antipov, M. Katsnelson, A. Lichtenstein, A. Rubtsov, K. Held, Diagrammatic routes to nonlocal correlations beyond dynamical mean field theory. *Reviews of Modern Physics* **90**, 025003 (2018).

[208] I. H. Kim, Holographic quantum simulation. *arXiv preprint arXiv:1702.02093* (2017).

[209] T. Peng, A. W. Harrow, M. Ozols, X. Wu, Simulating large quantum circuits on a small quantum computer. *Physical Review Letters* **125**, 150504 (2020).

[210] Q. Sun, G. K.-L. Chan, Quantum embedding theories. *Accounts of chemical research* **49**, 2705–2712 (2016).

[211] C. Vorwerk, N. Sheng, M. Govoni, B. Huang, G. Galli, Quantum embedding theories to simulate condensed systems on quantum computers. *Nature Computational Science* **2**, 424–432 (2022).

[212] B. Bauer, D. Wecker, A. J. Millis, M. B. Hastings, M. Troyer, Hybrid quantum-classical approach to correlated materials. *Physical Review X* **6**, 031045 (2016).

[213] H. Ma, M. Govoni, G. Galli, Quantum simulations of materials on near-term quantum computers. *npj Computational Materials* **6**, 1–8 (2020).

[214] G. Kotliar, S. Y. Savrasov, K. Haule, V. S. Oudovenko, O. Parcollet, C. A. Marianetti, Electronic structure calculations with dynamical mean-field theory. *Rev. Mod. Phys.* **78**, 865–951 (2006).

[215] I. Rungger, N. Fitzpatrick, H. Chen, C. Alderete, H. Apel, A. Cowtan, A. Patterson, D. M. Ramo, Y. Zhu, N. H. Nguyen, *et al.*, Dynamical mean field theory algorithm and experiment on quantum computers (2019).

[216] N. C. Rubin, A hybrid classical/quantum approach for large-scale studies of quantum systems with density matrix embedding theory (2016).

[217] G. Knizia, G. K.-L. Chan, Density matrix embedding: A simple alternative to dynamical mean-field theory. *Physical review letters* **109**, 186404 (2012).

[218] Y. Kawashima, M. P. Coons, Y. Nam, E. Lloyd, S. Matsuura, A. J. Garza, S. Johri, L. Huntington, V. Senicourt, A. O. Maksymov, *et al.*, Efficient and accurate electronic structure simulation demonstrated on a trapped-ion quantum computer (2021).

[219] J. I. Cirac, P. Maraner, J. K. Pachos, Cold atom simulation of interacting relativistic quantum field theories. *Phys. Rev. Lett.* **105**, 190403 (2010).

[220] S. Choi, C. J. Turner, H. Pichler, W. W. Ho, A. A. Michailidis, Z. Papić, M. Serbyn, M. D. Lukin, D. A. Abanin, Emergent su(2) dynamics and perfect quantum many-body scars. *Phys. Rev. Lett.* **122**, 220603 (2019).

[221] L. Villa, S. Thomson, L. Sanchez-Palencia, Finding the phase diagram of strongly correlated disordered bosons using quantum quenches. *Physical Review A* **104**, 023323 (2021).

[222] G. Wang, D. Stilck-França, R. Zhang, S. Zhu, P. D. Johnson, Quantum algorithm for ground state energy estimation using circuit depth with exponentially improved dependence on precision. *arXiv preprint arXiv:2209.06811* (2022).

[223] P. Dai, Antiferromagnetic order and spin dynamics in iron-based superconductors. *Reviews of Modern Physics* **87**, 855 (2015).

[224] M. Korshunov, I. Eremin, Theory of magnetic excitations in iron-based layered superconductors. *Physical Review B* **78**, 140509 (2008).

[225] M. Cazalilla, R. Citro, T. Giamarchi, E. Orignac, M. Rigol, One dimensional bosons: From condensed matter systems to ultracold gases. *Rev. Mod. Phys.* **83**, 1405 (2011).

[226] Z. Yan, Y.-R. Zhang, M. Gong, Y. Wu, Y. Zheng, S. Li, C. Wang, F. Liang, J. Lin, Y. Xu, *et al.*, Strongly correlated quantum walks with a 12-qubit superconducting processor. *Science* **364**, 753–756 (2019).

[227] S. Bravyi, M. B. Hastings, F. Verstraete, Lieb-robinson bounds and the generation of correlations and topological quantum order. *Phys. Rev. Lett.* **97**, 050401 (2006).

[228] Y. Lahini, M. Verbin, S. D. Huber, Y. Bromberg, R. Pugatch, Y. Silberberg, Quantum walk of two interacting bosons. *Phys. Rev. A* **86**, 011603 (2012).

[229] J. Šmakov, A. Chernyshev, S. R. White, Binding of holons and spinons in the one-dimensional anisotropic t- j model. *Phys. Rev. Lett.* **98**, 266401 (2007).

[230] F. Arute, K. Arya, R. Babbush, D. Bacon, J. C. Bardin, R. Barends, A. Bengtsson, S. Boixo, M. Broughton, B. B. Buckley, *et al.*, Observation of separated dynamics of charge and spin in the fermi-hubbard model. *arXiv preprint arXiv:2010.07965* (2020).

[231] Z. Jiang, K. J. Sung, K. Kechedzhi, V. N. Smelyanskiy, S. Boixo, Quantum algorithms to simulate many-body physics of correlated fermions. *Phys. Rev. Appl.* **9**, 044036 (2018).

[232] I. D. Kivlichan, J. McClean, N. Wiebe, C. Gidney, A. Aspuru-Guzik, G. K.-L. Chan, R. Babbush, Quantum Simulation of Electronic Structure with Linear Depth and Connectivity. *Phys. Rev. Lett.* **120**, 110501 (2018).

[233] M. Heyl, A. Polkovnikov, S. Kehrein, Dynamical quantum phase transitions in the transverse-field ising model. *Phys. Rev. Lett.* **110**, 135704 (2013).

[234] J. Zhang, G. Pagano, P. W. Hess, A. Kyprianidis, P. Becker, H. Kaplan, A. V. Gorshkov, Z.-X. Gong, C. Monroe, Observation of a many-body dynamical phase transition with a 53-qubit quantum simulator. *Nature* **551**, 601 (2017).

[235] K. Xu, Z.-H. Sun, W. Liu, Y.-R. Zhang, H. Li, H. Dong, W. Ren, P. Zhang, F. Nori, D. Zheng, *et al.*, Probing dynamical phase transitions with a superconducting quantum simulator. *Science advances* **6**, eaba4935 (2020).

[236] C. Monroe, W. Campbell, L.-M. Duan, Z.-X. Gong, A. Gorshkov, P. Hess, R. Islam, K. Kim, N. Linke, G. Pagano, *et al.*, Programmable quantum simulations of spin systems with trapped ions. *Rev. Mod. Phys.* **93**, 025001 (2021).

[237] G. Aleksandrowicz, T. Alexander, P. Barkoutsos, L. Bello, Y. Ben-Haim, *et al.*, Qiskit: An Open-source Framework for Quantum Computing (2019).

[238] S. Bravyi, S. Sheldon, A. Kandala, D. C. Mckay, J. M. Gambetta, Mitigating measurement errors in multi-qubit experiments. *arXiv preprint arXiv:2006.14044* (2020).

[239] R. Babbush, C. Gidney, D. W. Berry, N. Wiebe, J. McClean, A. Paler, A. Fowler, H. Neven, Encoding electronic spectra in quantum circuits with linear t complexity. *Physical Review X* **8**, 041015 (2018).

[240] Y. Su, D. W. Berry, N. Wiebe, N. Rubin, R. Babbush, Fault-tolerant quantum simulations of chemistry in first quantization. *PRX Quantum* **2**, 040332 (2021).

[241] Y. Dong, L. Lin, Y. Tong, Ground state preparation and energy estimation on early fault-tolerant quantum computers via quantum eigenvalue transformation of unitary matrices. *arXiv preprint arXiv:2204.05955* (2022).

[242] Y. Yang, B.-N. Lu, Y. Li, Accelerated quantum monte carlo with mitigated error on noisy quantum computer. *PRX Quantum* **2**, 040361 (2021).

[243] K. Wan, M. Berta, E. T. Campbell, A randomized quantum algorithm for statistical phase estimation (2021).

[244] N. J. Ross, P. Selinger, Optimal ancilla-free clifford+ t approximation of z-rotations. *Quantum Inf. Comput.* **16**, 901–953 (2016).

[245] A. Bocharov, M. Roetteler, K. M. Svore, Efficient synthesis of universal repeat-until-success quantum circuits. *Physical review letters* **114**, 080502 (2015).

[246] M. Mosca, P. Mukhopadhyay, A polynomial time and space heuristic algorithm for t-count. *Quantum Science and Technology* **7**, 015003 (2021).

[247] D. Maslov, Advantages of using relative-phase toffoli gates with an application to multiple control toffoli optimization. *Physical Review A* **93**, 022311 (2016).

[248] J. R. McClean, M. E. Kimchi-Schwartz, J. Carter, W. A. de Jong, Hybrid quantum-classical hierarchy for mitigation of decoherence and determination of excited states. *Phys. Rev. A* **95**, 042308 (2017).

[249] M. Huo, Y. Li, Temporally correlated error tomography and mitigation in quantum computer. *arXiv preprint arXiv:1811.02734* (2018).

[250] X. Bonet-Monroig, R. Sagastizabal, M. Singh, T. O'Brien, Low-cost error mitigation by symmetry verification. *Physical Review A* **98**, 062339 (2018).

[251] E. F. Dumitrescu, A. J. McCaskey, G. Hagen, G. R. Jansen, T. D. Morris, T. Papenbrock, R. C. Pooser, D. J. Dean, P. Lougovski, Cloud quantum computing of an atomic nucleus. *Physical review letters* **120**, 210501 (2018).

[252] M. Otten, S. K. Gray, Recovering noise-free quantum observables. *Physical Review A* **99**, 012338 (2019).

[253] S. McArdle, X. Yuan, S. Benjamin, Error-mitigated digital quantum simulation. *Physical review letters* **122**, 180501 (2019).

[254] R. Sagastizabal, X. Bonet-Monroig, M. Singh, M. A. Rol, C. Bultink, X. Fu, C. Price, V. Ostroukh, N. Muthusubramanian, A. Bruno, *et al.*, Experimental error mitigation via symmetry verification in a variational quantum eigensolver. *Physical Review A* **100**, 010302 (2019).

[255] W. J. Huggins, J. McClean, N. Rubin, Z. Jiang, N. Wiebe, K. B. Whaley, R. Babbush, Efficient and noise resilient measurements for quantum chemistry on near-term quantum computers. *arXiv preprint arXiv:1907.13117* (2019).

[256] M. Otten, S. K. Gray, Accounting for errors in quantum algorithms via individual error reduction. *Npj Quantum Inf.* **5**, 11 (2019).

[257] J. R. McClean, Z. Jiang, N. C. Rubin, R. Babbush, H. Neven, Decoding quantum errors with subspace expansions. *Nature Communications* **11**, 1–9 (2020).

[258] T. Giurgica-Tiron, Y. Hindy, R. LaRose, A. Mari, W. J. Zeng, Digital zero noise extrapolation for quantum error mitigation. *arXiv preprint arXiv:2005.10921* (2020).

[259] T. Keen, T. Maier, S. Johnston, P. Lougovski, Quantum-classical simulation of two-site dynamical mean-field theory on noisy quantum hardware. *Quantum Science and Technology* **5**, 035001 (2020).

[260] Z. Cai, Multi-exponential error extrapolation and combining error mitigation techniques for nisq applications. *arXiv preprint arXiv:2007.01265* (2020).

[261] A. He, B. Nachman, W. A. de Jong, C. W. Bauer, Resource efficient zero noise extrapolation with identity insertions. *arXiv preprint arXiv:2003.04941* (2020).

[262] F. B. Maciejewski, Z. Zimborás, M. Oszmaniec, Mitigation of readout noise in near-term quantum devices by classical post-processing based on detector tomography. *Quantum* **4**, 257 (2020).

[263] Y. Chen, M. Farahzad, S. Yoo, T.-C. Wei, Detector tomography on ibm quantum computers and mitigation of an imperfect measurement. *Physical Review A* **100**, 052315 (2019).

[264] H. Kwon, J. Bae, A hybrid quantum-classical approach to mitigating measurement errors. *arXiv preprint arXiv:2003.12314* (2020).

[265] A. Strikis, D. Qin, Y. Chen, S. C. Benjamin, Y. Li, Learning-based quantum error mitigation. *arXiv preprint arXiv:2005.07601* (2020).

[266] P. Czarnik, A. Arrasmith, P. J. Coles, L. Cincio, Error mitigation with clifford quantum-circuit data. *arXiv preprint arXiv:2005.10189* (2020).

[267] A. Zlokapa, A. Gheorghiu, A deep learning model for noise prediction on near-term quantum devices. *arXiv preprint arXiv:2005.10811* (2020).

[268] S. Sheldon, E. Magesan, J. M. Chow, J. M. Gambetta, Procedure for systematically tuning up cross-talk in the cross-resonance gate. *Physical Review A* **93**, 060302 (2016).

[269] J. Plantenberg, P. De Groot, C. Harmans, J. Mooij, Demonstration of controlled-not quantum gates on a pair of superconducting quantum bits. *Nature* **447**, 836–839 (2007).

[270] C. Neill, P. Roushan, K. Kechedzhi, S. Boixo, S. V. Isakov, V. Smelyanskiy, A. Megrant, B. Chiaro, A. Dunsworth, K. Arya, *et al.*, A blueprint for demonstrating quantum supremacy with superconducting qubits. *Science* **360**, 195–199 (2018).

[271] A. D. Córcoles, J. M. Gambetta, J. M. Chow, J. A. Smolin, M. Ware, J. Strand, B. L. Plourde, M. Steffen, Process verification of two-qubit quantum gates by randomized benchmarking. *Physical Review A* **87**, 030301 (2013).

[272] J. M. Chow, J. M. Gambetta, A. D. Corcoles, S. T. Merkel, J. A. Smolin, C. Rigetti, S. Poletto, G. A. Keefe, M. B. Rothwell, J. R. Rozen, *et al.*, Universal quantum gate set approaching fault-tolerant thresholds with superconducting qubits. *Physical review letters* **109**, 060501 (2012).

[273] C. Rigetti, M. Devoret, Fully microwave-tunable universal gates in superconducting qubits with linear couplings and fixed transition frequencies. *Physical Review B* **81**, 134507 (2010).

[274] M. Takita, A. W. Cross, A. Córcoles, J. M. Chow, J. M. Gambetta, Experimental demonstration of fault-tolerant state preparation with superconducting qubits. *Physical review letters* **119**, 180501 (2017).

[275] M. Reagor, C. B. Osborn, N. Tezak, A. Staley, G. Prawiroatmodjo, M. Scheer, N. Alidoust, E. A. Sete, N. Didier, M. P. da Silva, *et al.*, Demonstration of universal parametric entangling gates on a multi-qubit lattice. *Science advances* **4**, eaao3603 (2018).

[276] J. Eisert, M. Friesdorf, C. Gogolin, Quantum many-body systems out of equilibrium. *Nature Physics* **11**, 124 (2015).

[277] S. An, J.-N. Zhang, M. Um, D. Lv, Y. Lu, J. Zhang, Z.-Q. Yin, H. Quan, K. Kim, Experimental test of the quantum jarzynski equality with a trapped-ion system. *Nature Physics* **11**, 193 (2015).

[278] P. Schindler, M. Müller, D. Nigg, J. T. Barreiro, E. A. Martinez, M. Hennrich, T. Monz, S. Diehl, P. Zoller, R. Blatt, Quantum simulation of dynamical maps with trapped ions. *Nature Physics* **9**, 361 (2013).

[279] K. A. Landsman, C. Figgatt, T. Schuster, N. M. Linke, B. Yoshida, N. Y. Yao, C. Monroe, Verified quantum information scrambling. *Nature* **567**, 61 (2019).

[280] A. A. Houck, H. E. Türeci, J. Koch, On-chip quantum simulation with superconducting circuits. *Nature Physics* **8**, 292 (2012).

[281] R. Harris, Y. Sato, A. Berkley, M. Reis, F. Altomare, M. Amin, K. Boothby, P. Bunyk, C. Deng, C. Enderud, *et al.*, Phase transitions in a programmable quantum spin glass simulator. *Science* **361**, 162–165 (2018).

[282] Z.-X. Gong, L.-M. Duan, Prethermalization and dynamic phase transition in an isolated trapped ion spin chain. *New Journal of Physics* **15**, 113051 (2013).

[283] Y.-L. Wu, S. D. Sarma, Understanding analog quantum simulation dynamics in coupled ion-trap qubits. *Physical Review A* **93**, 022332 (2016).

[284] E. Altman, K. R. Brown, G. Carleo, L. D. Carr, E. Demler, C. Chin, B. DeMarco, S. E. Economou, M. Eriksson, K.-M. C. Fu, *et al.*, Quantum simulators: Architectures and opportunities. *arXiv preprint arXiv:1912.06938* (2019).

[285] P. M. Poggi, N. K. Lysne, K. W. Kuper, I. H. Deutsch, P. S. Jessen, Quantifying the sensitivity to errors in analog quantum simulation. *arXiv preprint arXiv:2007.01901* (2020).

[286] P. Hauke, F. M. Cucchietti, L. Tagliacozzo, I. Deutsch, M. Lewenstein, Can one trust quantum simulators? *Reports on Progress in Physics* **75**, 082401 (2012).

[287] J. Bylander, S. Gustavsson, F. Yan, F. Yoshihara, K. Harrabi, G. Fitch, D. G. Cory, Y. Nakamura, J.-S. Tsai, W. D. Oliver, Noise spectroscopy through dynamical decoupling with a superconducting flux qubit. *Nature Physics* **7**, 565 (2011).

[288] P. Bertet, I. Chiorescu, G. Burkard, K. Semba, C. Harmans, D. P. DiVincenzo, J. Mooij, Dephasing of a superconducting qubit induced by photon noise. *Physical review letters* **95**, 257002 (2005).

[289] J. M. Martinis, S. Nam, J. Aumentado, K. Lang, C. Urbina, Decoherence of a superconducting qubit due to bias noise. *Physical Review B* **67**, 094510 (2003).

[290] A. Walther, F. Ziesel, T. Ruster, S. T. Dawkins, K. Ott, M. Hettrich, K. Singer, F. Schmidt-Kaler, U. Poschinger, Controlling fast transport of cold trapped ions. *Physical review letters* **109**, 080501 (2012).

[291] B. Pokharel, N. Anand, B. Fortman, D. A. Lidar, Demonstration of fidelity improvement using dynamical decoupling with superconducting qubits. *Physical review letters* **121**, 220502 (2018).

[292] T. P. Harty, D. T. C. Allcock, C. J. Ballance, L. Guidoni, H. A. Janacek, N. M. Linke, D. N. Stacey, D. M. Lucas, High-fidelity preparation, gates, memory, and readout of a trapped-ion quantum bit. *Phys. Rev. Lett.* **113**, 220501 (2014).

[293] C. J. Ballance, T. P. Harty, N. M. Linke, M. A. Sepiol, D. M. Lucas, High-fidelity quantum logic gates using trapped-ion hyperfine qubits. *Phys. Rev. Lett.* **117**, 060504 (2016).

[294] J. P. Gaebler, T. R. Tan, Y. Lin, Y. Wan, R. Bowler, A. C. Keith, S. Glancy, K. Coakley, E. Knill, D. Leibfried, D. J. Wineland, High-fidelity universal gate set for $^{9}Be^{+}$ ion qubits. *Phys. Rev. Lett.* **117**, 060505 (2016).

[295] M. Gong, S. Wang, C. Zha, M.-C. Chen, H.-L. Huang, Y. Wu, Q. Zhu, Y. Zhao, S. Li, S. Guo, *et al.*, Quantum walks on a programmable two-dimensional 62-qubit superconducting processor. *Science* **372**, 948952 (2021).

[296] A. Mezzacapo, U. Las Heras, J. Pedernales, L. DiCarlo, E. Solano, L. Lamata, Digital quantum rabi and dicke models in superconducting circuits. *Scientific reports* **4**, 1–4 (2014).

[297] L. García-Álvarez, J. Casanova, A. Mezzacapo, I. Egusquiza, L. Lamata, G. Romero, E. Solano, Fermion-fermion scattering in quantum field theory with superconducting circuits. *Physical review letters* **114**, 070502 (2015).

[298] S. Asaad, C. Dickel, N. K. Langford, S. Poletto, A. Bruno, M. A. Rol, D. Deurloo, L. DiCarlo, Independent, extensible control of same-frequency superconducting qubits by selective broadcasting. *npj Quantum Information* **2**, 1–7 (2016).

[299] S. J. Weber, G. O. Samach, D. Hover, S. Gustavsson, D. K. Kim, A. Melville, D. Rosenberg, A. P. Sears, F. Yan, J. L. Yoder, *et al.*, Coherent coupled qubits for quantum annealing. *Physical Review Applied* **8**, 014004 (2017).

[300] H. Häffner, W. Hänsel, C. Roos, J. Benhelm, M. Chwalla, T. Körber, U. Rapol, M. Riebe, P. Schmidt, C. Becher, *et al.*, Scalable multiparticle entanglement of trapped ions. *Nature* **438**, 643–646 (2005).

[301] M. Saffman, Quantum computing with atomic qubits and rydberg interactions: progress and challenges. *Journal of Physics B: Atomic, Molecular and Optical Physics* **49**, 202001 (2016).

[302] A. Parra-Rodriguez, P. Lougovski, L. Lamata, E. Solano, M. Sanz, Digital-analog quantum computation. *Physical Review A* **101**, 022305 (2020).

[303] A. Martin, L. Lamata, E. Solano, M. Sanz, Digital-analog quantum algorithm for the quantum fourier transform. *Physical Review Research* **2**, 013012 (2020).

[304] E. Bairey, C. Guo, D. Poletti, N. H. Lindner, I. Arad, Learning the dynamics of open quantum systems from their steady states. *New Journal of Physics* **22**, 032001 (2020).

[305] M. P. da Silva, O. Landon-Cardinal, D. Poulin, Practical characterization of quantum devices without tomography. *Physical Review Letters* **107**, 210404 (2011).

[306] M. H. Lee, I. Kim, R. Dekeyser, Time-dependent behavior of the spin-1/2 anisotropic heisenberg model in infinite lattice dimensions. *Physical review letters* **52**, 1579 (1984).

[307] L. Siurakshina, D. Ihle, R. Hayn, Theory of magnetic order in the three-dimensional spatially anisotropic heisenberg model. *Physical Review B* **61**, 14601 (2000).

[308] A. Hucht, A. Moschel, K. Usadel, Monte-carlo study of the reorientation transition in heisenberg models with dipole interactions. *Journal of magnetism and magnetic materials* **148**, 32–33 (1995).

[309] D. Torelli, T. Olsen, Calculating critical temperatures for ferromagnetic order in two-dimensional materials. *2D Materials* **6**, 015028 (2018).

[310] C. Soukoulis, S. Datta, Y. H. Lee, Spin-wave theory for anisotropic heisenberg antiferromagnets. *Physical Review B* **44**, 446 (1991).

[311] L. Lamata, A. Parra-Rodriguez, M. Sanz, E. Solano, Digital-analog quantum simulations with superconducting circuits. *Advances in Physics: X* **3**, 1457981 (2018).

[312] W. D. Oliver, P. B. Welander, Materials in superconducting quantum bits. *MRS bulletin* **38**, 816–825 (2013).

[313] Z. Zhou, S.-I. Chu, S. Han, Relaxation and decoherence in a resonantly driven qubit. *Journal of Physics B: Atomic, Molecular and Optical Physics* **41**, 045506 (2008).

[314] C. Müller, J. Lisenfeld, A. Shnirman, S. Poletto, Interacting two-level defects as sources of fluctuating high-frequency noise in superconducting circuits. *Physical Review B* **92**, 035442 (2015).

[315] S. M. Meißner, A. Seiler, J. Lisenfeld, A. V. Ustinov, G. Weiss, Probing individual tunneling fluctuators with coherently controlled tunneling systems. *Physical Review B* **97**, 180505 (2018).

[316] C. Neill, A. Megrant, R. Barends, Y. Chen, B. Chiaro, J. Kelly, J. Mutus, P. O'Malley, D. Sank, J. Wenner, *et al.*, Fluctuations from edge defects in superconducting resonators. *Applied Physics Letters* **103**, 072601 (2013).

[317] H.-Y. Huang, R. Kueng, J. Preskill, Information-theoretic bounds on quantum advantage in machine learning. *Phys. Rev. Lett.* **126**, 190505 (2021).

[318] C. Hadfield, Adaptive pauli shadows for energy estimation. *arXiv preprint arXiv:2105.12207* (2021).

[319] S. Hillmich, C. Hadfield, R. Raymond, A. Mezzacapo, R. Wille, *2021 IEEE International Conference on Quantum Computing and Engineering (QCE)* (IEEE, 2021), pp. 24–34.

[320] T. Maier, M. Jarrell, T. Pruschke, M. H. Hettler, Quantum cluster theories. *Reviews of Modern Physics* **77**, 1027 (2005).

[321] W. Li, Z. Huang, C. Cao, Y. Huang, Z. Shuai, X. Sun, J. Sun, X. Yuan, D. Lv, Toward practical quantum embedding simulation of realistic chemical systems on near-term quantum computers. *Chemical science* **13**, 8953–8962 (2022).

[322] I. W. Bulik, G. E. Scuseria, J. Dukelsky, Density matrix embedding from broken symmetry lattice mean fields. *Physical Review B* **89**, 035140 (2014).

[323] J. Tilly, P. V. Sriluckshmy, A. Patel, E. Fontana, I. Rungger, E. Grant, R. Anderson, J. Tennyson, G. H. Booth, Reduced density matrix sampling: Self-consistent embedding and multiscale electronic structure on current generation quantum computers. *arXiv preprint arXiv:2104.05531* (2021).

[324] E. Fertitta, G. H. Booth, Energy-weighted density matrix embedding of open correlated chemical fragments. *J. Chem. Phys.* **151**, 014115 (2019).

[325] Z.-H. Cui, T. Zhu, G. K.-L. Chan, Efficient implementation of ab initio quantum embedding in periodic systems: Density matrix embedding theory. *J. Chem. Theory Comput.* **16**, 119-129 (2020).

[326] H. Q. Pham, M. R. Hermes, L. Gagliardi, Periodic electronic structure calculations with the density matrix embedding theory. *J. Chem. Theory Comput.* **16**, 130-140 (2020).

[327] H. Chen, M. Nusspickel, J. Tilly, G. H. Booth, Variational quantum eigensolver for dynamic correlation functions. *Phys. Rev. A* **104** (2021).

[328] C. J. Scott, G. H. Booth, Extending density matrix embedding: A static two-particle theory. *arXiv preprint arXiv:2107.04335* (2021).

[329] T. Zhu, G. K.-L. Chan, Ab initio full cell $GW + \mathrm{DMFT}$ for correlated materials. *Phys. Rev. X* **11**, 021006 (2021).

[330] R. G. Parr, W. Yang, *Density-Functional Theory of Atoms and Molecules*, International Series of Monographs on Chemistry (Oxford University Press, New York, 1995).

[331] M. Rossmannek, P. K. Barkoutsos, P. J. Ollitrault, I. Tavernelli, Quantum hf/dft-embedding algorithms for electronic structure calculations: Scaling up to complex molecular systems. *J. Chem. Phys.* **154**, 114105 (2021).

[332] Y. Du, M.-H. Hsieh, T. Liu, D. Tao, Expressive power of parametrized quantum circuits. *Physical Review Research* **2**, 033125 (2020).

[333] Z. Holmes, K. Sharma, M. Cerezo, P. J. Coles, Connecting ansatz expressibility to gradient magnitudes and barren plateaus. *PRX Quantum* **3**, 010313 (2022).

[334] Y. Du, Z. Tu, X. Yuan, D. Tao, Efficient measure for the expressivity of variational quantum algorithms. *Physical Review Letters* **128**, 080506 (2022).

[335] J. Eisert, Entangling power and quantum circuit complexity. *Physical Review Letters* **127**, 020501 (2021).

[336] J. Haferkamp, P. Faist, N. B. Kothakonda, J. Eisert, N. Yunger Halpern, Linear growth of quantum circuit complexity. *Nature Physics* **18**, 528–532 (2022).

[337] W. Kutzelnigg, Quantum chemistry in Fock space. I. the universal wave and energy operators. *J. Chem. Phys.* **77**, 3081-3097 (1982).

[338] R. J. Bartlett, S. A. Kucharski, J. Noga, Alternative coupled-cluster anstze II. the unitary coupled-cluster method. *Chem. Phys. Lett.* **155**, 133-140 (1989).

[339] A. G. Taube, R. J. Bartlett, New perspectives on unitary coupled-cluster theory. *Int. J. Quantum Chem.* **106**, 3393-3401 (2006).

[340] R. J. Bartlett, M. Musiał, Coupled-cluster theory in quantum chemistry. *Rev. Mod. Phys.* **79**, 291 (2007).

[341] B. P. Lanyon, J. D. Whitfield, G. G. Gillett, M. E. Goggin, M. P. Almeida, I. Kassal, J. D. Biamonte, M. Mohseni, B. J. Powell, M. Barbieri, *et al.*, Towards quantum chemistry on a quantum computer. *Nature chemistry* **2**, 106 (2010).

[342] C. Cao, J. Hu, W. Zhang, X. Xu, D. Chen, F. Yu, J. Li, H.-S. Hu, D. Lv, M.-H. Yung, Progress toward larger molecular simulation on a quantum computer: Simulating a system with up to 28 qubits accelerated by point-group symmetry. *Phys. Rev. A* **105**, 062452 (2022).

[343] K. Setia, R. Chen, J. E. Rice, A. Mezzacapo, M. Pistoia, J. D. Whitfield, Reducing qubit requirements for quantum simulations using molecular point group symmetries. *J. Chem. Theory Comput.* **16**, 6091-6097 (2020).

[344] J. F. Stanton, J. Gauss, J. D. Watts, R. J. Bartlett, A direct product decomposition approach for symmetry exploitation in many-body methods. i. energy calculations. *The Journal of Chemical Physics* **94**, 4334–4345 (1991).

[345] C. Cao, J. Sun, X. Yuan, H.-S. Hu, H. Q. Pham, D. Lv, Ab initio quantum simulation of strongly correlated materials with quantum embedding. *arXiv preprint arXiv:2209.03202* (2022).

[346] G. Vidal, Class of quantum many-body states that can be efficiently simulated. *Phys. Rev. Lett.* **101**, 110501 (2008).

[347] R. Orús, A practical introduction to tensor networks: Matrix product states and projected entangled pair states. *Annals of Physics* **349**, 117–158 (2014).

[348] Y.-Y. Shi, L.-M. Duan, G. Vidal, Classical simulation of quantum many-body systems with a tree tensor network. *Phys. Rev. A* **74**, 022320 (2006).

[349] G. Evenbly, G. Vidal, Algorithms for entanglement renormalization. *Phys. Rev. B* **79**, 144108 (2009).

[350] G. Evenbly, G. Vidal, Entanglement renormalization in two spatial dimensions. *Phys. Rev. Lett.* **102**, 180406 (2009).

[351] M. Kieferová, A. Scherer, D. W. Berry, Simulating the dynamics of time-dependent hamiltonians with a truncated dyson series. *Phys. Rev. A* **99**, 042314 (2019).

[352] Y.-H. Chen, A. Kalev, I. Hen, A quantum algorithm for time-dependent hamiltonian simulation by permutation expansion. *arXiv preprint arXiv:2103.15334* (2021).

[353] Y. Wan, N. Armitage, Resolving continua of fractional excitations by spinon echo in thz 2d coherent spectroscopy. *Physical Review Letters* **122**, 257401 (2019).

[354] T. V. Zache, T. Schweigler, S. Erne, J. Schmiedmayer, J. Berges, Extracting the field theory description of a quantum many-body system from experimental data. *Physical Review X* **10**, 011020 (2020).

[355] G. Xu, Z. Xu, J. Tranquada, Absolute cross-section normalization of magnetic neutron scattering data. *Review of Scientific Instruments* **84**, 083906 (2013).

[356] E. Stoudenmire, S. R. White, Minimally entangled typical thermal state algorithms. *New Journal of Physics* **12**, 055026 (2010).

[357] Y. Huang, Y. Shao, W. Ren, J. Sun, D. Lv, Efficient quantum imaginary time evolution by drifting real time evolution: an approach with low gate and measurement complexity. *arXiv preprint arXiv:2203.11112* (2022).

[358] C. L. Cortes, S. K. Gray, Quantum krylov subspace algorithms for ground-and excited-state energy estimation. *Physical Review A* **105**, 022417 (2022).

[359] G. H. Low, I. L. Chuang, Hamiltonian simulation by uniform spectral amplification. *arXiv preprint arXiv:1707.05391* (2017).