

# Theory of variational quantum simulation

Xiao Yuan<sup>1</sup>, Suguru Endo<sup>1</sup>, Qi Zhao<sup>2</sup>, Ying Li<sup>3</sup>, and Simon C. Benjamin<sup>1</sup>

<sup>1</sup>Department of Materials, University of Oxford, Parks Road, Oxford OX1 3PH, United Kingdom

<sup>2</sup>Center for Quantum Information, Institute for Interdisciplinary Information Sciences, Tsinghua University, Beijing 100084, China

<sup>3</sup>Graduate School of China Academy of Engineering Physics, Beijing 100193, China

The variational method is a versatile tool for classical simulation of a variety of quantum systems. Great efforts have recently been devoted to its extension to quantum computing for efficiently solving static many-body problems and simulating real and imaginary time dynamics. In this work, we first review the conventional variational principles, including the Rayleigh-Ritz method for solving static problems, and the Dirac and Frenkel variational principle, the McLachlan's variational principle, and the time-dependent variational principle, for simulating real time dynamics. We focus on the simulation of dynamics and discuss the connections of the three variational principles. Previous works mainly focus on the unitary evolution of pure states. In this work, we introduce variational quantum simulation of mixed states under general stochastic evolution. We show how the results can be reduced to the pure state case with a correction term that takes accounts of global phase alignment. For variational simulation of imaginary time evolution, we also extend it to the mixed state scenario and discuss variational Gibbs state preparation. We further elaborate on the design of ansatz that is compatible with post-selection measurement and the implementation of the generalised variational algorithms with quantum circuits. Our work completes the theory of variational quantum simulation of general real and imaginary time evolution and it is applicable to near-term quantum hardware.

## Contents

<b>1</b>	<b>Introduction</b>	<b>2</b>
<b>2</b>	<b>Preliminary</b>	<b>4</b>
2.1	Static problem: Rayleigh-Ritz method . . . . .	4
2.2	Dynamics of closed systems: three variational principles . . . . .	5
2.2.1	The Dirac and Frenkel variational principle . . . . .	6
2.2.2	McLachlan's variational principle . . . . .	6
2.2.3	The time-dependent variational principle . . . . .	7
2.3	Implementation . . . . .	8
<b>3</b>	<b>Equivalence of the three variational principles</b>	<b>9</b>
<b>4</b>	<b>Real time evolution: open quantum systems</b>	<b>10</b>
4.1	General evolution . . . . .	11
4.1.1	The Dirac and Frenkel variational principle . . . . .	11
4.1.2	McLachlan's variational principle . . . . .	12

4.1.3	The time-dependent variational principle . . . . .	12
4.2	Reduction to unitary evolution . . . . .	12
<b>5</b>	<b>Imaginary time evolution</b>	<b>14</b>
5.1	Pure states . . . . .	14
5.1.1	The Dirac and Frenkel variational principle . . . . .	14
5.1.2	McLachlan’s variational principle . . . . .	14
5.1.3	Time-dependent variational principles . . . . .	15
5.2	Mixed state . . . . .	15
5.2.1	The Dirac and Frenkel variational principle . . . . .	15
5.2.2	MacLachlan’s variational principle . . . . .	16
5.2.3	The time-dependent variational principle . . . . .	16
5.3	Variational Gibbs state preparation . . . . .	16
<b>6</b>	<b>Implementation</b>	<b>17</b>
6.1	Trial state implementation . . . . .	17
6.1.1	Pure states . . . . .	17
6.1.2	Mixed states and unitary evolution . . . . .	17
6.1.3	Stochastic and imaginary time evolution . . . . .	18
6.2	Coefficients evaluation . . . . .	18
6.3	Numerical simulation . . . . .	20
<b>7</b>	<b>Discussion</b>	<b>21</b>
<b>A</b>	<b>Real time evolution</b>	<b>22</b>
A.1	Real time evolution: pure state and unitary evolution . . . . .	22
A.1.1	The Dirac and Frenkel variational principle . . . . .	22
A.1.2	McLachlan’s variational principle . . . . .	23
A.1.3	Time-dependent variational principles . . . . .	24
A.2	Mixed states and general evolution . . . . .	25
A.2.1	General evolution . . . . .	25
A.2.2	Reduction to the pure state case . . . . .	27
<b>B</b>	<b>Imaginary time evolution</b>	<b>29</b>
B.1	Imaginary time evolution: pure states . . . . .	29
B.1.1	The Dirac and Frenkel variational principle . . . . .	29
B.1.2	McLachlan’s variational principle . . . . .	30
B.1.3	Time-dependent variational principles . . . . .	31
B.2	Imaginary time evolution of mixed state . . . . .	32
B.2.1	Evolution of mixed states . . . . .	32
B.2.2	Reduction to the pure state case . . . . .	35

# 1 Introduction

Variational simulation is a widely used technique in many-body physics [1–7] and chemistry [8–10]. As the full Hilbert space of a quantum system grows exponentially with respect to the system size, direct classical simulation of a large many-body state is generally impossible. The variational method avoids the exponential space problem by considering trial states from a physically motivated small subset of the exponentially large Hilbert space [11, 12]. Variational classical simulation (VCS) works

for both static problems in finding the ground state and ground state energy, and dynamical problems in simulating the real and imaginary time evolution of quantum states [4–6, 12–16]. In modern science, the variational method has become a versatile tool for simulating various problems when the target system state can be well modelled classically.

Nevertheless, there also exist many problems that may not be solved classically even with the classical variational method [17–19]. This is because there exist highly entangled many-body states that may not be efficiently represented via any classical method. This problem motivates the development of quantum simulation with universal quantum computers [20], because quantum systems can be efficiently encoded or represented by another quantum system and real time evolution of the Schrödinger equation can be efficiently realised via a unitary quantum circuit [21, 22]. However, implementing a universal quantum computer requires the coherent and accurate control of millions of qubits [23–25]; While the state-of-the-art quantum hardware can only accurately control tens of qubits [19, 26]. Although this number may be extended by one or two orders in the near future, realising a universal fault tolerant quantum computer remains a challenging task.

With noisy intermediate-scale quantum hardware [27], quantum advantages may still be achieved in many tasks with recently proposed hybrid or, more specifically, variational quantum simulation (VQS) methods [28–42]. By dividing the problem into two levels, VQS methods only use the quantum processor to solve the core and classically intractable problem and leave the relatively easy task to a classical computer. For solving static problems, the Rayleigh-Ritz variational method can be naturally generalised to the quantum regime, such as the variational quantum eigensolver (VQE) [29–40] and the quantum approximate optimisation algorithm (QAOA) [28]. Considering parameterised trial states prepared by quantum circuits [43–45], the target static problem is mapped to a cost function of the measurement outcomes of the quantum state, which is further optimised via a classical algorithm. As the state preparation and measurement procedures may be hard to simulate by a classical computer [17–19], hybrid or variational quantum simulation can have an intrinsic quantum advantage over VCS and it has been proposed as the standard tool for studying quantum computational chemistry [46, 47].

A broader and harder challenge is the simulation of dynamics. Classically, there are three variational principles—the Dirac and Frenkel variational principle [48, 49], the McLachlan variational principle [50], and the time-dependent variational method principle [51, 52]. These variational principles have been extensively studied in VCS for example on Bose-Einstein condensation [2] and with matrix-product states in the so called tangent-bundle formalism [53]. Recently, the variational methods have been generalised to simulating the real [41, 54] and imaginary [55, 56] time evolution of pure quantum systems, and been experimentally implemented with superconducting qubits [57]. By encoding quantum states with a parameterised circuit, the evolution of the state can be mapped to the evolution of the parameters controlling the circuit. However, different variational principles can lead to different evolution equations of parameters. It has not been clear which variational principle or which evolution equation should be used in practice. Furthermore, previous works only focus on the evolution of pure states of closed systems, rather than the more general case of the stochastic evolution of mixed states of open systems. While practical quantum systems always suffer from unavoidable interaction with the environment, it makes almost every quantum system open. Therefore, simulating the evolution of open quantum systems can be useful for studying realistic quantum systems.

In this work, we complete the VQS theory by studying the equivalence and difference of the variational principles and extending the principles to the general stochastic evolution of mixed states. We first review the theory of variational simulation in Sec. 2. Then we study the equivalence and difference of the variational principles and the derived evolution equations in Sec. 3. We summarise the results in Table 1 under various conditions. In particular, we consider general real time stochastic evolution of mixed states in Sec. 4. As an example, we re-derive the evolution equation of parameters for unitary evolution of pure states and find a correction term which applies to previous results and

handles the global phase alignment. After applying all the three variational principles to the general mixed state case, we find that McLachlan's variational principle is the most appropriate variational principle that leads to a consistent theory of VQS. In Sec. 5, we study variational simulation of imaginary time evolution of mixed states and discuss variational Gibbs state preparation. We further elaborate on the implementation with quantum circuits in Sec. 6. We conclude our work and discuss future directions in Sec. 7.

## 2 Preliminary

In this section, we introduce the background of variational classical and quantum simulation. We first review the Rayleigh-Ritz principle for solving static problems and the Dirac and Frenkel variational principle [51, 52], the McLachlan variational principle [50], and the time-dependent variational principle [51, 52] for simulating dynamics. We also review the implementation of the simulation with quantum circuits. It is worth noting that the variational principles work similarly for both variational classical and quantum simulation. The main differences between variational classical and quantum simulation are about how the trial state or the ansatz is implemented and how the evolution equation of the parameters is determined, as we shortly discuss. Variational classical simulation has been extensively studied for different tasks [4–6, 12–16]. This section will not review those results. Instead, we aim to have a self-consistent introduction of different variational principles with a focus on variational quantum simulation.

### 2.1 Static problem: Rayleigh-Ritz method

The Rayleigh-Ritz method is designed to solve static problems of finding the ground state energy  $E_0$  of a given Hamiltonian  $H = \sum_j h_j \sigma_j$ . Here, we assume that  $H$  is a linear combination of tensor products of local operators  $\sigma_j$  with coefficients  $h_j$ . The ground state energy is the solution of the problem,

$$E_0 = \min_{|\psi\rangle} \frac{\langle \psi | H | \psi \rangle}{\langle \psi | \psi \rangle}, \quad (1)$$

where the minimisation is over states from the whole Hilbert space. Because the size of the Hilbert space grows exponentially to the system size, it is in general computationally hard to brutal force search the whole Hilbert space. Instead, the Rayleigh-Ritz method only searches states from a subset of the whole Hilbert space to find an approximate solution.

For example, classically, one can consider trial states that are linear combinations of a small number of basis states  $\{|\phi_i\rangle\}$ ,

$$|\phi(\vec{a})\rangle = \sum_i a_i |\phi_i\rangle, \quad (2)$$

with  $\vec{a} = (a_1, a_2, \dots)$ . Suppose the ground state can be approximated by the trial state  $|\phi(\vec{a})\rangle$  with a proper vector of coefficients  $\vec{a}$ , then the ground state energy can be approximated or upper bounded by

$$E_0 \leq E_0^{est} = \min_{\vec{a}} \frac{\langle \phi(\vec{a}) | H | \phi(\vec{a}) \rangle}{\langle \phi(\vec{a}) | \phi(\vec{a}) \rangle} = \min_{\vec{a}} \frac{\sum_{i,j} a_i^* a_j \langle \phi_i | H | \phi_j \rangle}{\sum_{i,j} a_i^* a_j \langle \phi_i | \phi_j \rangle}. \quad (3)$$

Suppose each term  $\langle \phi_i | H | \phi_j \rangle$  can be efficiently obtained, the minimisation problem can be solved by calculating the partial derivative of  $E_0^{est}$  over  $a_i^*$ , and is equivalent to the minimal solution to

$$\det(H - \lambda S) = 0, \quad (4)$$

where  $H$  and  $S$  are the matrix of  $\langle \phi_i | H | \phi_j \rangle$  and  $\langle \phi_i | \phi_j \rangle$ , respectively. We refer to the trial state also as the ansatz. In practice, different ansätze have been invented in condensed matter physics and

computational chemistry [6, 11, 44–46]. When the minimisation cannot be solved analytically, one can also numerically solve it with classical optimisation algorithms. Such a variational method with classical ansatz is called VCS.

With quantum computers or in VQS, the trial state  $|\phi(\vec{\theta})\rangle$  can be prepared by applying a sequence of parameterised gates  $R_k(\theta_k)$  to an initial state  $|\bar{0}\rangle$  as

$$|\phi(\vec{\theta})\rangle = R_N(\theta_N) \dots R_k(\theta_k) \dots R_1(\theta_1) |\bar{0}\rangle.$$

Here,  $R_k(\theta_k)$  is the  $k^{\text{th}}$  gate controlled by the real parameter  $\theta_k$  and  $\vec{\theta} = (\theta_1, \theta_2, \dots, \theta_N)$ . The ansatz is automatically normalised and the average energy  $\langle\phi(\vec{\theta})|H|\phi(\vec{\theta})\rangle$  can be obtained by measuring each term  $\langle\phi(\vec{\theta})|h_i|\phi(\vec{\theta})\rangle$  and linearly combining the measurement outcomes. In Sec. 6, we also introduce other possible ways of realising the trial states. To approximate the ground state, we can minimise  $\langle\phi(\vec{\theta})|H|\phi(\vec{\theta})\rangle$  over the parameter space with a classical optimisation algorithm. For instance, by calculating the gradient of  $\langle\phi(\vec{\theta})|H|\phi(\vec{\theta})\rangle$ , a local minimum of  $\langle\phi(\vec{\theta})|H|\phi(\vec{\theta})\rangle$  can be found via the gradient descent algorithm. Such a hybrid algorithm for solving the eigenvalue of the Hamiltonian is called variational quantum eigensolver (VQE) [29–40]. Note that parameters can be complex for classical simulation but it suffices for them to be real for quantum simulation. This is because without loss of generality the trial state can be prepared by a quantum circuit by applying unitary gates  $R_k(\theta_k)$  that is in the form of  $e^{i\theta_k\sigma}$  with Hermitian operator  $\sigma$  and real parameter  $\theta_k$ .

## 2.2 Dynamics of closed systems: three variational principles

Now we introduce three variational principles for simulating real time dynamics of closed systems with pure quantum states [4, 41]. We only review the results here and leave the detailed derivation in Appendix A. To simulate Schrödinger’s equation

$$\frac{d|\psi(t)\rangle}{dt} = -iH|\psi(t)\rangle, \quad (5)$$

we consider a parametrised trial state  $|\phi(\vec{\theta}(t))\rangle$  with time dependent parameters. Again, the parameters can be complex for VCS and are assumed to be real for VQS. Suppose the state at time  $t$  is represented by the trial state  $|\psi(t)\rangle = |\phi(\vec{\theta}(t))\rangle$  with parameters  $\vec{\theta}(t)$ , then we want to approximate the state  $|\psi(t + \delta t)\rangle$  at time  $t + \delta t$  by  $|\phi(\vec{\theta}(t + \delta t))\rangle$ . Note that the state is evolved from  $|\psi(t)\rangle$  to  $|\psi(t + \delta t)\rangle$  according to Schrödinger’s equation as

$$|\psi(t + \delta t)\rangle \approx |\phi(\vec{\theta}(t))\rangle - i\delta t H |\phi(\vec{\theta}(t))\rangle.$$

Such a state may be impossible to be represented by the trial state with any parameters. A variational method is used to project the state  $|\psi(t + \delta t)\rangle$  back to the trial state manifold or find the best solution  $\vec{\theta}(t + \delta t)$  that approximates

$$|\phi(\vec{\theta}(t + \delta t))\rangle \approx |\phi(\vec{\theta}(t))\rangle - i\delta t H |\phi(\vec{\theta}(t))\rangle. \quad (6)$$

Note that

$$|\phi(\vec{\theta}(t + \delta t))\rangle \approx |\phi(\vec{\theta}(t))\rangle + \sum_j \frac{\partial |\phi(\vec{\theta}(t))\rangle}{\partial \theta_j} \delta \theta_j,$$

the target problem reduces to approximate

$$\sum_j \frac{\partial |\phi(\vec{\theta}(t))\rangle}{\partial \theta_j} \delta \theta_j \approx -i\delta t H |\phi(\vec{\theta}(t))\rangle. \quad (7)$$

Here, the L.H.S. and R.H.S. denote the change of the trial state by varying the parameters and the evolution according to the Schrödinger equation, respectively. Under different variational principles, we aim to minimise the difference so that the evolution of the state  $|\psi\rangle$  under the Schrödinger equation can be mapped onto the trial state manifold as the evolution of parameters.

### 2.2.1 The Dirac and Frenkel variational principle

In Eq. (7), the L.H.S is in the tangent subspace  $\{\frac{\partial|\phi(\vec{\theta}(t))\rangle}{\partial\theta_j}\}$ , but the R.H.S.  $-iH|\phi(\vec{\theta}(t))\rangle$  may not be in this subspace. So the Dirac and Frenkel variational principle [51, 52] directly projects the equation onto the subspace, which can be equivalently expressed by

$$\left\langle \delta\phi(\vec{\theta}(t)) \left| \left( \frac{d}{dt} + iH \right) \right| \phi(\vec{\theta}(t)) \right\rangle = 0, \quad (8)$$

with  $\langle \delta\phi(\vec{\theta}(t)) | = \sum_i \frac{\partial \langle \phi(\vec{\theta}(t)) |}{\partial \theta_i} \delta\theta_i$ . The evolution of parameters can be solved to be

$$\sum_j A_{i,j} \dot{\theta}_j = -iC_i, \quad (9)$$

where  $A$  and  $C$  are

$$\begin{aligned} A_{i,j} &= \frac{\partial \langle \phi(\vec{\theta}(t)) |}{\partial \theta_i} \frac{\partial |\phi(\vec{\theta}(t))\rangle}{\partial \theta_j}, \\ C_i &= \frac{\partial \langle \phi(\vec{\theta}(t)) |}{\partial \theta_i} H |\phi(\vec{\theta}(t))\rangle. \end{aligned} \quad (10)$$

In general,  $A$  and  $C$  are complex even if the parameters are real, which leads to a complex solution  $\dot{\theta}_j$ . A complex solution is not problematic in classical simulation as parameters are complex. However, parameters in quantum simulation are generally real, a complex solution thus leads to a state outside of the ansatz space. Naively, one can take the real or imaginary part of Eq. (9) to make the solution real. Although this seems artificial, we show that the corresponding equations can be obtained from the McLachlan's variational principle and the time-dependent variational principle, respectively.

### 2.2.2 McLachlan's variational principle

The McLachlan's variational principle [50] is to minimise the distance between the L.H.S and R.H.S of Eq. (7), which can be equivalently expressed as

$$\delta \|(d/dt + iH) |\phi(\vec{\theta}(t))\rangle\| = 0. \quad (11)$$

Here  $\| |\psi\rangle \| = \sqrt{\langle \psi | \psi \rangle}$  is the norm of quantum states  $|\psi\rangle$ . When the parameters  $\theta$  are complex, the solution gives the same evolution of Eq. (9) under the Dirac and Frenkel variational principle, indicating the equivalence of these two variational principles. However, assuming  $\theta$  to be real, the equation becomes

$$\sum_j A_{i,j}^R \dot{\theta}_j = C_i^I, \quad (12)$$

which can be obtained by taking the real part of Eq. (9). Here  $A_{i,j}^R$  is the real part of  $A_{i,j}$  and  $C_i^I$  is the imaginary part of  $C_i$ . As shown in Sec. 3, this equation is equivalent to Eq. (9) when the trial state  $|\phi(\vec{\theta}(t))\rangle$  is powerful enough to represent the target state  $|\psi(t)\rangle$ . The advantage of this equation is that it always gives a real solution of  $\dot{\theta}$ , which makes it consistent with VQS with real parameters. Furthermore, we will show shortly that McLachlan's variational principle is the most consistent variational principle that works similarly to general stochastic evolution.

However, it is worth noting that the evolution according to Eq. (12) is not the ultimate one owing to a time-dependent global phase mismatch between the target state  $|\psi(t)\rangle$  and the trial state  $|\phi(\vec{\theta}(t))\rangle$ . That is, even if  $|\phi(\vec{\theta}(t))\rangle$  can represent  $|\psi(t)\rangle$  up to a global phase, Eq. (12) may still be incorrect to be the evolution equation of the parameters. The main reason for this seemingly counter-intuitive phenomenon stems from the fact that even if  $|\psi(t)\rangle$  and  $|\phi(\vec{\theta}(t))\rangle$  are equivalent up to a time-dependent global phase, their derivatives  $d|\psi(t)\rangle/dt$  and  $d|\phi(\vec{\theta}(t))\rangle/dt$  can be very different. This problem can be either addressed by introducing an actual time-dependent phase gate with an additional parameter as  $e^{i\theta_0(t)}|\phi(\vec{\theta}(t))\rangle$ . Alternatively, we show that this is not necessary as the problem can be resolved with a modification of Eq. (12). Suppose the state at time  $t$  is represented by the trial state,  $|\psi(t)\rangle = e^{i\theta_0(t)}|\phi(\vec{\theta}(t))\rangle$ , up to a global phase  $\theta_0(t)$ . Although global phase is physically irrelevant, it plays an important role in the variational method. Considering Eq. (7), the L.H.S. becomes

$$e^{i\theta_0(t)} \sum_j \frac{\partial \langle \phi(\vec{\theta}(t)) |}{\partial \theta_j} \delta \theta_j + i e^{i\theta_0(t)} \delta \theta_0 |\phi(\vec{\theta}(t))\rangle$$

and the R.H.S. is  $-i\delta t H e^{i\theta_0(t)} |\phi(\vec{\theta}(t))\rangle$ . According to McLachlan's variational principle, the minimisation between the difference of them gives a similar evolution equation of the parameters  $\vec{\theta}$ ,

$$\sum_j M_{i,j} \dot{\theta}_j = V_i, \quad (13)$$

with

$$\begin{aligned} M_{i,j} &= A_{i,j}^R + \frac{\partial \langle \phi(\vec{\theta}(t)) |}{\partial \theta_i} |\phi(\vec{\theta}(t))\rangle \frac{\partial \langle \phi(\vec{\theta}(t)) |}{\partial \theta_j} |\phi(\vec{\theta}(t))\rangle, \\ V_i &= C_i^I + i \frac{\partial \langle \phi(\vec{\theta}(t)) |}{\partial \theta_i} |\phi(\vec{\theta}(t))\rangle \langle \phi(\vec{\theta}(t)) | H | \phi(\vec{\theta}(t)) \rangle. \end{aligned} \quad (14)$$

Therefore, as long as we measure the correction terms, we can automatically resolve the global phase problem without actually introducing the global phase. As shown in Sec. 4, those correction terms can be also obtained by considering the McLachlan's variational principle for mixed states, which automatically handles the global phase problem.

A further point of interest concerns our ability to track the accuracy of the simulation. In practice, after solving the derivatives of  $\dot{\theta}$  in Eq. (12), we can also get the distance between the true evolution and the evolution of the trial state,

$$\|(d/dt + iH) |\phi(\vec{\theta}(t))\rangle\|^2 = \sum_{i,j} A_{i,j}^R \dot{\theta}_i \dot{\theta}_j - 2 \sum_i C_i^I \dot{\theta}_i + \langle H^2 \rangle. \quad (15)$$

Here  $\langle H^2 \rangle = \langle \phi(\vec{\theta}(t)) | H^2 | \phi(\vec{\theta}(t)) \rangle$ . Therefore, by measuring  $A^R$ ,  $C^I$ , and  $\langle H^2 \rangle$ , we can calculate the distance to verify the variational algorithm. This works similarly for the modified equation that takes account of the global phase.

### 2.2.3 The time-dependent variational principle

The Schrödinger equation can be obtained from the Lagrangian  $L = \langle \psi(t) | (d/dt + iH) | \psi(t) \rangle$  or more precisely  $iL$  so that the Lagrangian is real. As the imaginary term  $i$  does not affect the evolution equation, we omit it for simplicity. Replacing  $|\psi(t)\rangle$  with  $|\phi(\vec{\theta}(t))\rangle$ , the Lagrangian becomes

$$L = \left\langle \phi(\vec{\theta}(t)) \left| \left( \frac{d}{dt} + iH \right) \right| \phi(\vec{\theta}(t)) \right\rangle. \quad (16)$$



For complex  $\theta_i$ , parameters are  $\theta_i$ ,  $\theta_i^*$ ,  $\dot{\theta}_i$ , and  $\dot{\theta}_i^*$ . After applying the Euler-Lagrange equation, the evolution of parameters can be derived and shown to coincide with Eq. (9). For real  $\theta_i$ , parameters of the Lagrangian are  $\theta_i$  and  $\dot{\theta}_i$ , and the evolution of parameters are

$$\sum_j A_{i,j}^I \dot{\theta}_j = -C_i^R, \quad (17)$$

which can be obtained by taking the imaginary part of Eq. (9). Here  $A_{i,j}^I$  is the imaginary part of  $A_{i,j}$  and  $C_i^R$  is the real part of  $C_i$ . As shown in Sec. 3, this evolution equation is equivalent to Eq. (9) when the trial state  $|\phi(\vec{\theta}(t))\rangle$  is powerful enough to represent the target state  $|\psi(t)\rangle$ . However, as we take the imaginary part of the  $A$  matrix, the  $A^I$  matrix is more likely to be singular, making the evolution unstable. Note that there also exists other ways of choosing the Lagrangian, such as the one considered in Ref. [53],

$$L = \frac{i}{2} \left( \langle \psi | \frac{d|\psi\rangle}{dt} - \frac{d\langle \psi |}{dt} |\psi\rangle \right) - \langle \psi | H | \psi \rangle.$$

Nevertheless, it is not hard to verify that it leads to the same evolution equation as Eq. (17).

### 2.3 Implementation

In variational classical simulation, the  $A$  matrix and the  $C$  vector are calculated classically. Here we show how to evaluate  $A$  and  $C$  with a quantum circuit introduced by [41]. Suppose the parameterised trial state is prepared by  $|\phi(\vec{\theta})\rangle = R|\bar{0}\rangle$  with  $R = R_N(\theta_N) \dots R_k(\theta_k) \dots R_1(\theta_1)$ . In conventional quantum computing, each unitary gate involves a small subset of qubits, typically one or two. Therefore, the derivative of the unitary can be efficiently decomposed via

$$\frac{\partial R_k}{\partial \theta_k} = \sum_i f_{k,i} R_k \sigma_{k,i}, \quad (18)$$

where  $\sigma_{k,i}$  are also unitary operators and  $f_{k,i}$  are complex coefficients. For most single- and two-qubit-gates  $R_k$ , we find that there are only one or two terms in this expression, and  $\sigma_{k,i}$  is also a one-qubit or two-qubit gate. For example, when  $R_k(\theta_k) = e^{-i\theta_k \sigma/2}$  with a one- or two-qubit Hermitian operator  $\sigma$ , we have  $\partial R_k(\theta_k)/\partial \theta_k = -i/2 \cdot \sigma \cdot e^{-i\theta_k \sigma/2}$  with  $f_k = -i/2$  and  $\sigma_k = \sigma$  in Eq. (18).

Using the expression (18), we rewrite the derivative of the state as

$$\frac{\partial |\phi(\vec{\theta}(t))\rangle}{\partial \theta_k} = \sum_i f_{k,i} R_{k,i} |\bar{0}\rangle, \quad (19)$$

where

$$R_{k,i} = R_N R_{N-1} \dots R_{k+1} R_k \sigma_{k,i} \dots R_2 R_1. \quad (20)$$

Then, based on the definition of the Hamiltonian, the elements of  $A$  and  $C$  can be expressed as

$$A_{k,q} = \sum_{i,j} \left( f_{k,i}^* f_{q,j} \langle \bar{0} | R_{k,i}^\dagger R_{q,j} | \bar{0} \rangle \right), \quad (21)$$

and

$$C_k = \sum_{i,j} \left( f_{k,i}^* h_j \langle \bar{0} | R_{k,i}^\dagger \sigma_j R | \bar{0} \rangle \right). \quad (22)$$



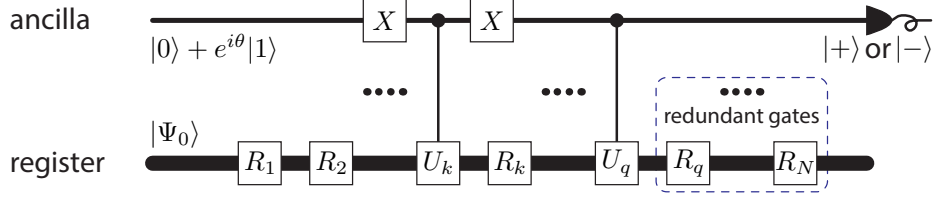


Figure 1: Quantum circuit for the evaluation of coefficients in the variational pure-state simulator (from Ref. [41]). The ancillary qubit is initialised in the state  $(|0\rangle + e^{i\theta}|1\rangle)/\sqrt{2}$  and measured in the  $|\pm\rangle = (|0\rangle \pm |1\rangle)/\sqrt{2}$  basis. The probability  $P_+$  that the qubit is in the state  $|+\rangle$  is  $(\Re(e^{i\theta}\langle\bar{0}|U|\bar{0}\rangle) + 1)/2$ , where  $U = R_1^\dagger \cdots U_k^\dagger R_k^\dagger \cdots R_N^\dagger R_N \cdots R_q U_q \cdots R_1$ . Here,  $U_k$  is one of  $\sigma_{k,i}$ , and  $U_q$  is one of  $\sigma_{q,j}$  or  $\sigma_j$  (By taking  $q = N + 1$ ,  $\sigma_j$  is put on the left side of  $R_N$  in the product), assuming that  $k < q$ . We would like to remark that this circuit is a variant of the circuit proposed in Ref. [59]. This circuit contains  $N$  gates on the register, two flip gates ( $X$ ) on the ancillary qubit, and two controlled unitary gates between the ancillary qubit and the register. Note that gates on the register after the second controlled unitary gate can be omitted. Usually,  $U_k$  and  $U_q$  are unitary operators on only one or two register qubits rather than the entire register. This circuit can be further reduced to a direct measurement on the register qubits without the ancilla [58].

The real and imaginary part of expressions (21) and (22) are in the form

$$a\Re\left(e^{i\theta}\langle\bar{0}|U|\bar{0}\rangle\right),$$

where the amplitude  $a$  and phase  $\theta$  are determined by the coefficients, and  $U$  is a unitary operator either  $R_{k,i}^\dagger R_{q,j}$  or  $R_{k,i}^\dagger \sigma_j R$ . Such a term can be evaluated with the quantum circuit shown in Fig. 1. This circuit needs an ancillary qubit initialised in the state  $(|0\rangle + e^{i\theta}|1\rangle)/\sqrt{2}$  and a register initialised in the state  $|\Psi_0\rangle$ . The ancillary qubit is measured after a sequence of unitary operations on the register and two controlled unitary operations, where the ancillary qubit is the control qubit. The value is given by  $\Re(e^{i\theta}\langle\bar{0}|U|\bar{0}\rangle) = 2P_+ - 1$ , where  $P_+$  is the probability that the qubit is in the state  $|+\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$ . Recently, [58] also showed a way to measure  $A$  and  $C$  without the ancillary qubit, making the realisation more friendly to near-term quantum hardware.

### 3 Equivalence of the three variational principles

With complex parameters of the trial state, the three variational principles lead to the same evolution equation of parameters. We will see that this is also true for general stochastic evolution and imaginary time evolution of mixed states. However, when considering real parameters as in quantum simulation, we can get three different equations (9), (12), and (17) for the evolution of parameters. With classical ansatz, a necessary condition for the equivalence of the three equations is discussed by [52]. That is these three equations are equivalent when for any parameter  $\theta_i$ , there always exists a parameter  $\theta_j$ , such that

$$\frac{\partial |\phi(\vec{\theta}(t))\rangle}{\partial \theta_i} = i \frac{\partial |\phi(\vec{\theta}(t))\rangle}{\partial \theta_j}. \quad (23)$$

However, such a condition is hard to be satisfied in quantum simulation with real parameters. Generally, when  $|\phi(\vec{\theta}(t))\rangle$  can uniquely represent any state  $|\psi(t)\rangle$  at time  $t$ , Eq. (9) and Eq. (12) both give real solutions and are also equivalent. The necessary and sufficient condition for the equivalence of Eq. (9) and Eq. (12) only requires the solution of Eq. (9) to be real.

Broadly speaking we may expect that Eq. (12) and Eq. (17) each suffice to define the evolution of our parameters. Practically, there are differences however. In particular, we can consider a single qubit system with Hamiltonian  $\sigma_x$  and initial state  $|0\rangle$ . Suppose the trial state is chosen to be  $e^{-i\theta\sigma_x}|0\rangle$ ,

Table 1: Summary of evolution equation of parameters under different conditions. We consider real/imaginary time dynamics for pure/mixed states with complex/real parameters and three different variational principles. TDVP: time-dependent variational principle. The evolution equation is the same for complex parameters of the three variational principles. As pure state unitary dynamics is a special case of the mixed state open system dynamics, the evolution equation for mixed states also works for pure states, which further corrects the global phase problem for pure states as discussed in Sec. 4.2. The definition of the matrices are:  $A_{i,j} = \frac{\partial \langle \phi(\vec{\theta}(t)) |}{\partial \theta_i} \frac{\partial \langle \phi(\vec{\theta}(t)) |}{\partial \theta_j}$  and  $C_i = \frac{\partial \langle \phi(\vec{\theta}(t)) |}{\partial \theta_i} H | \phi(\vec{\theta}(t)) \rangle$  are defined in Eq. (10);  $M_{i,j} = \text{Tr} \left[ \left( \frac{\partial \rho(\vec{\theta}(t))}{\partial \theta_i} \right)^\dagger \frac{\partial \rho(\vec{\theta}(t))}{\partial \theta_j} \right]$ ,  $V_i = \text{Tr} \left[ \left( \frac{\partial \rho(\vec{\theta}(t))}{\partial \theta_i} \right)^\dagger \mathcal{L}(\rho) \right]$  are defined in Eq. (29).  $C'_i = \frac{\partial \langle \phi(\vec{\theta}(\tau)) |}{\partial \theta_i} (H - E_\tau) | \phi(\vec{\theta}(\tau)) \rangle$  is defined in Eq. (40);  $Y_i = -\text{Tr} \left[ \frac{\partial \rho(\vec{\theta}(t))}{\partial \theta_i} \{H, \rho(\vec{\theta}(t))\} \right]$  is defined in Eq. (51).

Dynamics	Pure/mixed states	Parameters	Variational principles	Evolution equations
Real	Pure $\frac{d \psi(t)\rangle}{dt} = -iH \psi(t)\rangle$ (5)	Complex	-	$\sum_j A_{i,j} \dot{\theta}_j = -iC_i$ (9)
		Real	Dirac and Frenkel MacLachlan	$\sum_j A_{i,j} \dot{\theta}_j = -iC_i$ (9) $\sum_j A_{i,j}^R \dot{\theta}_j = C_i^I$ (12)
			TDVP	$\sum_j A_{i,j}^I \dot{\theta}_j = -C_i^R$ (17)
	Mixed $\frac{d\rho}{dt} = \mathcal{L}(\rho)$ (24)	Complex	-	$\sum_j M_{i,j} \dot{\theta}_j = V_i$ (28)
		Real	Dirac and Frenkel MacLachlan	$\sum_j M_{i,j} \dot{\theta}_j = V_i$ (28) $\sum_j M_{i,j} \dot{\theta}_j = V_i$ (28)
			TDVP	-
Imag	Pure $\frac{d \psi(\tau)\rangle}{d\tau} = -(H - E_\tau) \psi(\tau)\rangle$ (36)	Complex	-	$\sum_j A_{i,j} \dot{\theta}_j = -C'_i$ (39)
		Real	Dirac and Frenkel MacLachlan	$\sum_j A_{i,j} \dot{\theta}_j = -C'_i$ (39) $\sum_j A_{i,j}^R \dot{\theta}_j = -C_i^R$ (42)
			TDVP	-
	Mixed $\frac{d\rho(\tau)}{d\tau} = -(\{H, \rho(\tau)\} - 2\langle H \rangle \rho(\tau))$ (46)	Complex	-	$\sum_j M_{i,j} \dot{\theta}_j = Y_i$ (49)
		Real	Dirac and Frenkel MacLachlan	$\sum_j M_{i,j} \dot{\theta}_j = Y_i$ (49) $\sum_j M_{i,j} \dot{\theta}_j = Y_i$ (49)
			TDVP	-

which can exactly represent the evolution of the state. For such a simple case, the evolution equation (17) based on the time-dependent variational principle cannot work as  $A^I$  and  $C^R$  are always zero; nevertheless, Eq. (9) and Eq. (12) can both simulate the evolution. In fact, as the diagonal elements of  $A$  are in general real constants, the diagonal elements of  $A^R$  and  $A^I$  are real constants and zero, respectively. As the diagonal elements of  $A^I$  are all zero, it makes the inverse of  $A^I$  unstable when the off-diagonal elements are also close to 0. This suggests that the evolution equation (17) may not be the ideal choice for VQS.

In the following, we will re-derive the evolution equations from the mixed state perspective. Surprisingly, we will show that Eq. (17) from the time-dependent variational principle cannot be obtained in the mixed state case, whereas a variant of Eq. (12) can be consistently derived. Therefore, our work suggests that McLachlan's principle is the most consistent variational principle in variational quantum simulation. We also summarise the evolution equations of parameters under different conditions in Table 1.

## 4 Real time evolution: open quantum systems

In this section, we extend VQS to mixed states under general stochastic evolutions, which can describe open quantum systems and noisy quantum hardware. Note that this approach differs from that in Ref. [60] where we can variationally simulate the stochastic master equation.

## 4.1 General evolution

The general stochastic evolution of an open quantum system can be modelled by

$$\frac{d\rho}{dt} = \mathcal{L}(\rho), \quad (24)$$

where the superoperator  $\mathcal{L}(\rho)$  can be decomposed as

$$\mathcal{L}(\rho) = \sum_i g_i S_i \rho T_i^\dagger, \quad (25)$$

with unitary operators  $S_i$  and  $T_i$ , and coefficients  $g_i$ . Suppose  $\rho = \rho(\vec{\theta}(t))$  is a parameterised state, we show how to transform the evolution of  $\rho$  into the evolution of parameters  $\vec{\theta}$ . Here, we mainly focus on the evolution of parameters and refer to Sec. 6 for the design of the mixed states ansatz with unitary circuits.

### 4.1.1 The Dirac and Frenkel variational principle

Let  $\rho = \rho(\vec{\theta}(t))$ , the general stochastic evolution becomes

$$\sum_i \frac{\partial \rho(\vec{\theta}(t))}{\partial \theta_i} \dot{\theta}_i = \mathcal{L}(\rho). \quad (26)$$

Because  $\frac{d\rho(\vec{\theta}(t))}{dt} = \sum_i \frac{\partial \rho(\vec{\theta}(t))}{\partial \theta_i} \dot{\theta}_i$  is in the tangent subspace of  $\{\frac{\partial \rho(\vec{\theta}(t))}{\partial \theta_i}\}$ , we can project the equation onto this tangent subspace by following the Dirac and Frenkel variational principle,

$$\text{Tr} \left[ (\delta \rho(\vec{\theta}(t)))^\dagger \left( \frac{d\rho}{dt} - \mathcal{L}(\rho) \right) \right] = 0. \quad (27)$$

The evolution of parameters is

$$\sum_j M_{i,j} \dot{\theta}_j = V_i, \quad (28)$$

with elements of  $M$  and  $V$  defined by

$$\begin{aligned} M_{i,j} &= \text{Tr} \left[ \left( \frac{\partial \rho(\vec{\theta}(t))}{\partial \theta_i} \right)^\dagger \frac{\partial \rho(\vec{\theta}(t))}{\partial \theta_j} \right], \\ V_i &= \text{Tr} \left[ \left( \frac{\partial \rho(\vec{\theta}(t))}{\partial \theta_i} \right)^\dagger \mathcal{L}(\rho) \right]. \end{aligned} \quad (29)$$

In classical simulation,  $\vec{\theta}$  can be complex and  $\left( \frac{\partial \rho(\vec{\theta}(t))}{\partial \theta_i} \right)^\dagger \neq \frac{\partial \rho(\vec{\theta}(t))}{\partial \theta_j}$ . In this case,  $M$ ,  $V$ , and the solution  $\dot{\theta}$  can all be complex. However, considering quantum simulation with real parameters  $\theta$ , we have  $\left( \frac{\partial \rho(\vec{\theta}(t))}{\partial \theta_i} \right)^\dagger = \frac{\partial \rho(\vec{\theta}(t))}{\partial \theta_j}$ . Therefore,  $M$  and  $V$  are real and the solution  $\dot{\theta}$  is also real. **This is different from the pure state case where even if parameters are real, the solution  $\dot{\theta}$  can still be complex.** Each term of  $M$  and  $V$  can be calculated with a quantum circuit that is discussed in Sec. 6.

#### 4.1.2 McLachlan's variational principle

We can also minimise the distance between the evolution of the trial state and the true evolution,

$$\delta \|d\rho/dt - \mathcal{L}(\rho)\| = 0, \quad (30)$$

where  $\|\rho\| = \sqrt{\text{Tr}[\rho^2]}$  is the  $l_2$  norm of matrices. With both complex and real parameters, we can get the same result with the Dirac and Frenkel variational principle. Meanwhile, after solving the derivatives of  $\dot{\theta}$ , we can also get the distance between the true evolution and the evolution of the trial state by calculating

$$\|d\rho/dt - \mathcal{L}(\rho)\|^2 = \sum_{i,j} M_{i,j} \dot{\theta}_i \dot{\theta}_j - 2 \sum_i V_i \dot{\theta}_i + \text{Tr}[\mathcal{L}(\rho)^2]. \quad (31)$$

Therefore, by measuring  $M$ ,  $V$ , and  $\text{Tr}[\mathcal{L}(\rho)^2]$ , we can calculate the distance to verify the variational algorithm.

#### 4.1.3 The time-dependent variational principle

In general, **Hamiltonians for open systems may not be well defined due to the dissipation.** Here, we construct the Lagrangian as follows,

$$L = \text{Tr}[\rho(\vec{\theta}(t))^\dagger (d\rho(\vec{\theta}(t))/dt - \mathcal{L}(\rho))], \quad (32)$$

so that after the Euler-Lagrange equation, we can recover the evolution equation defined in Eq. (24). Here, we assume that  $\theta$  are complex parameters so that  $\rho^\dagger$  and  $\rho$  can be also regarded to be independent. Applying the Euler-Lagrange equation to  $\theta^*$  gives the same evolution Eq. (28). However, when parameters are real with  $\theta = \theta^*$ , we also have  $\rho^\dagger = \rho$  and the Euler-Lagrange equation cannot lead to Eq. (24) or any evolution equation of the parameters. We refer to Appendix A.2 for details.

From the above result, we find that the three variational principles are equivalent for complex parameters. However, when parameters are real, the Dirac and Frenkel and the McLachlan variational principles are also equivalent, while the time-dependent variational principle cannot lead to a nontrivial evolution of parameters. **Therefore, the McLachlan variational principle always gives a consistent real evolution of parameters in VQS.**

## 4.2 Reduction to unitary evolution

Note that the **unitary evolution of Schrödinger's equation is a special case of the stochastic evolution** defined in Eq. (24). Therefore, it would be interesting to verify whether the evolution equations of parameters for Schrödinger's equation can be consistently obtained from the more general scenario. First, we consider unitary evolution of mixed input states. Suppose the initial state is  $\rho(0)$ , the evolution under Hamiltonian  $H$  is governed by the von Neumann equation with  $\mathcal{L}(\rho) = -i[H, \rho(t)]$ ,

$$\frac{d\rho(t)}{dt} = -i[H, \rho(t)]. \quad (33)$$

In this case,  $M$  is defined in Eq. (29) and  $V$  is explicitly given by

$$V_i = -i \text{Tr} \left[ \left( \frac{\partial \rho(\vec{\theta}(t))}{\partial \theta_i} \right)^\dagger [H, \rho(\vec{\theta}(t))] \right]. \quad (34)$$

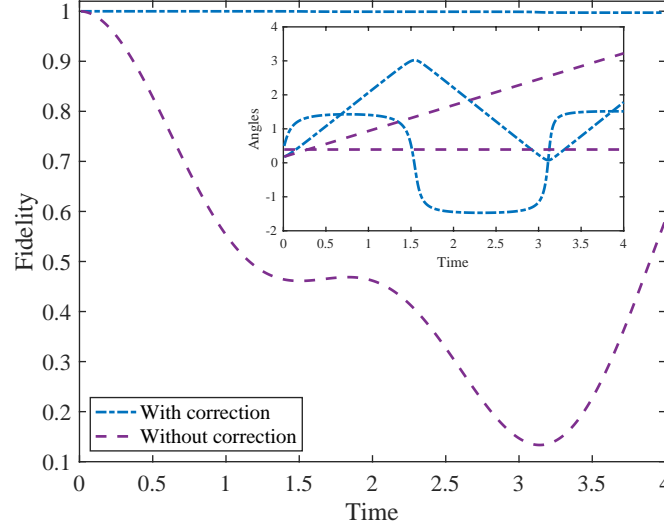


Figure 2: Comparison between the evolution equations Eq. (35) and Eq. (12) for real time evolution with Hamiltonian  $\sigma_y$  and trial state  $|\phi(\vec{\theta}(t))\rangle = R_z(\theta_z)R_x(\theta_x)|0\rangle$ . Here, we consider randomly initialised state ( $\theta_1 = 0.1734$  and  $\theta_2 = 0.3909$  for the shown results) and find that the evolution under Eq. (12) generally fails to simulate the true evolution even if the ansatz can represent all pure qubit state. The Y-axis is the fidelity of the state under variational simulation to the state under the exact evolution. Dash-dot line corresponds to Eq. (35); Dashed line corresponds to Eq. (12). The evolution of parameters are shown in the inset.

Suppose  $\rho(\vec{\theta}(t)) = |\phi(\vec{\theta}(t))\rangle \langle \phi(\vec{\theta}(t))|$  is a pure state, it becomes the Schrödinger equation and we have

$$M_{i,j} = A_{i,j}^R + \frac{\partial \langle \phi(\vec{\theta}(t)) |}{\partial \theta_i} |\phi(\vec{\theta}(t))\rangle \frac{\partial \langle \phi(\vec{\theta}(t)) |}{\partial \theta_j} |\phi(\vec{\theta}(t))\rangle, \quad (35)$$

$$V_i = C_i^I + i \frac{\partial \langle \phi(\vec{\theta}(t)) |}{\partial \theta_i} |\phi(\vec{\theta}(t))\rangle \langle \phi(\vec{\theta}(t)) | H | \phi(\vec{\theta}(t)) \rangle.$$

The evolution equation obtained from directly applying the McLachlan's principle to pure states with unitary evolution is  $\sum_j A_{i,j}^R \dot{\theta}_j = C_i^I$ , as given in Eq. (12). Compared to the  $A^R$  matrix and  $C^I$  vector, there are additional terms that estimate the overlap between the trial state and the derivative of the trial state in  $M$  and  $V$ . The terms  $\frac{\partial \langle \phi(\vec{\theta}(t)) |}{\partial \theta_i} |\phi(\vec{\theta}(t))\rangle$  can be measured with the circuit in Fig. 1 and the average energy  $\langle \phi(\vec{\theta}(t)) | H | \phi(\vec{\theta}(t)) \rangle$  can be directly measured. As we have discussed in Sec. 2, the additional terms take account of the global phase difference between the trial state and the target evolution state. Consider a trial state  $e^{i\theta_0} |\phi(\vec{\theta}(t))\rangle$  with  $N$  parameters in  $|\phi(\vec{\theta}(t))\rangle$  and an extra parameter  $\theta_0$  for the global phase. Using Eq. (12), the evolution of the first  $N$  parameters of  $e^{i\theta_0} |\phi(\vec{\theta}(t))\rangle$  is exactly given by Eq. (35). Because the global phase  $\theta_0$  is irrelevant to any measurement, we can thus only evolve the first  $N$  parameters with Eq. (35). However, the extra parameter  $\theta_0$  is important for the global phase alignment between the trial state and the exact state as we confirm with the following example.

Here, we consider a single-qubit state evolution to illustrate the difference between Eq. (35) and Eq. (12). Suppose the trial state is prepared by  $|\phi(\vec{\theta}(t))\rangle = R_z(\theta_z)R_x(\theta_x)|0\rangle$  with two parameters  $\theta_z$  and  $\theta_x$ . Here,  $R_x(\theta) = e^{-i\theta\sigma_x/2}$  and  $R_z(\theta) = e^{-i\theta\sigma_z/2}$ , with Pauli matrices  $\sigma_x$  and  $\sigma_z$ . Suppose the system Hamiltonian is the Pauli matrix  $\sigma_y$ , we compare the simulation of the time evolution  $e^{-i\sigma_y t}$  with randomly initialised state, as shown in Fig. 2. We find that the evolution under Eq. (12) generally fails to simulate the exact real time dynamics. Nevertheless, as the trial state can represent all pure

qubit state up to a global phase, the evolution under Eq. (35) can indeed simulate the true time evolution  $e^{-i\sigma_y t}$ . Therefore, Eq. (35) is recommended for variational simulation of real time unitary dynamics of pure states.

## 5 Imaginary time evolution

In this section, we show that the variational principles can be also applied to the simulation of imaginary time evolution.

### 5.1 Pure states

We first focus on the pure state case studied in Ref. [55]. Replacing real time with imaginary time  $\tau = it$ , the state at time  $\tau$  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,

$$\frac{d|\psi(\tau)\rangle}{d\tau} = -(H - E_\tau)|\psi(\tau)\rangle, \quad (36)$$

where  $E_\tau = \langle\psi(\tau)|H|\psi(\tau)\rangle$  is the expected energy at time  $\tau$ . Consider a normalised parametrised trial state  $|\phi(\vec{\theta}(\tau))\rangle$ , the imaginary time evolution of the Schrödinger equation on the trial state space is

$$\sum_i \frac{\partial |\phi(\vec{\theta}(\tau))\rangle}{\partial \theta_i} \dot{\theta}_i = -(H - E_\tau) |\phi(\vec{\theta}(\tau))\rangle. \quad (37)$$

Therefore, we can evolve parameters to simulate the unphysical imaginary time evolution.

#### 5.1.1 The Dirac and Frenkel variational principle

Applying the Dirac and Frenkel variational principle, we have

$$\left\langle \delta\phi(\vec{\theta}(\tau)) \left| \left( \frac{d}{d\tau} + H - E_\tau \right) \right| \phi(\vec{\theta}(\tau)) \right\rangle = 0, \quad (38)$$

and the evolution of parameters is simplified to

$$\sum_j A_{i,j} \dot{\theta}_j = -C'_i, \quad (39)$$

with

$$C'_i = \frac{\partial \langle \phi(\vec{\theta}(\tau)) |}{\partial \theta_i} (H - E_\tau) |\phi(\vec{\theta}(\tau))\rangle. \quad (40)$$

Similar to the real time evolution, even if the parameter  $\theta_j$  is real, the solution of its derivative  $\dot{\theta}_j$  may not be real as  $A$  and  $C'$  are complex.

#### 5.1.2 McLachlan's variational principle

Applying the McLachlan's variational principle [50] to imaginary time evolution, we have

$$\delta \|(d/d\tau + H - E_\tau) |\phi(\vec{\theta}(\tau))\rangle\| = 0. \quad (41)$$

When  $\dot{\theta}_i$  is complex, the evolution of parameters is the same as Eq. (39). When  $\dot{\theta}_i$  can only take real values, the evolution becomes

$$\sum_j A_{i,j}^R \dot{\theta}_j = -C_i^R. \quad (42)$$

Due to the normalisation requirement for the trial state  $|\phi(\vec{\theta}(\tau))\rangle$ , i.e.,  $\langle\phi(\vec{\theta}(\tau))|\phi(\vec{\theta}(\tau))\rangle = 1$ , we have  $\Re\left(\frac{\partial\langle\phi(\vec{\theta}(t))|}{\partial\theta_i} E_\tau |\phi(\vec{\theta}(t))\rangle\right) = 0$  and therefore  $C_i^R = C_i'^R$ .

### 5.1.3 Time-dependent variational principles

The Lagrangian for the Schrödinger equation is

$$L = \left\langle \psi \left| \frac{d}{d\tau} + H \right| \psi \right\rangle + \langle \psi | H | \psi \rangle (1 - \langle \psi | \psi \rangle), \quad (43)$$

which reproduces the imaginary time evolution equation defined in Eq. (36). When  $|\psi\rangle = \phi(\vec{\theta}(t))$  has complex  $\theta$ , we get the evolution of Eq. (39). Suppose  $\theta$  is real, the evolution becomes

$$i \sum_j A_{i,j}^I \dot{\theta}_j = C_i^R, \quad (44)$$

where,  $A_{i,j}^I$  and  $C_i^R$  are the imaginary and real parts of  $A_{i,j}$  and  $C_i$ , respectively. The solution is imaginary and hence not physical.

To summarise, when parameters are complex, the three variational principles are still equivalent. However, considering real parameters, only the McLachlan's variational principle can guarantee a real solution of the evolution equation.

## 5.2 Mixed state

Now, we consider the case that the input state is a mixed state. Under the imaginary time evolution of Hamiltonian  $H$ , the state  $\rho(\tau)$  at imaginary time  $\tau$  is

$$\rho(\tau) = \frac{e^{-H\tau} \rho(0) e^{-H\tau}}{\text{Tr}[e^{-2H\tau} \rho(0)]}, \quad (45)$$

where  $\rho(0)$  is the initial state. Equivalently,  $\rho(\tau)$  follows the time derivative equation

$$\frac{d\rho(\tau)}{d\tau} = -\left(\{H, \rho(\tau)\} - 2\langle H \rangle \rho(\tau)\right), \quad (46)$$

where  $\{H, \rho(\tau)\} = H\rho(\tau) + \rho(\tau)H$ . Consider a parametrised state  $\rho(\vec{\theta}(t))$ , the evolution becomes

$$\sum_i \frac{\partial\rho(\vec{\theta})}{\partial\theta_i} \dot{\theta}_i = -\left(\{H, \rho(\vec{\theta}(t))\} - 2\langle H \rangle \rho(\vec{\theta}(t))\right). \quad (47)$$

### 5.2.1 The Dirac and Frenkel variational principle

To get the evolution of parameters, we project the equation onto  $\delta\rho(\vec{\theta})$  according to the Dirac and Frenkel variational principle. Therefore, we have

$$\text{Tr}\left[\delta\rho(\vec{\theta})\left(\sum_i \frac{\partial\rho(\vec{\theta})}{\partial\theta_i} \dot{\theta}_i + \{H, \rho(\vec{\theta}(t))\} - 2\langle H \rangle \rho(\vec{\theta}(t))\right)\right] \quad (48)$$



equal zero and the solution is

$$\sum_j M_{i,j} \dot{\theta}_j = Y_i, \quad (49)$$

with  $M$  defined in Eq. (29) and  $Y$  given by

$$Y_i = -\text{Tr} \left[ \left( \frac{\partial \rho(\vec{\theta})}{\partial \theta_i} \right)^\dagger \left( \{H, \rho(\vec{\theta}(t))\} - 2 \langle H \rangle \rho(\vec{\theta}(t)) \right) \right]. \quad (50)$$

Note that when parameters  $\vec{\theta}$  are real,  $M$  and  $Y$  are also real, and we have

$$Y_i = -\text{Tr} \left[ \frac{\partial \rho(\vec{\theta}(t))}{\partial \theta_i} \{H, \rho(\vec{\theta}(t))\} \right]. \quad (51)$$

In this case, the solution  $\vec{\theta}$  is also real note that this is opposite to the pure state case.

### 5.2.2 MacLachlan's variational principle

We can also minimise the distance between the evolution via parameters and the evolution,

$$\delta \|d\rho/d\tau + (\{H, \rho\} - 2 \langle H \rangle \rho)\| = 0. \quad (52)$$

The evolution to  $\vec{\theta}$  is the same as Eq. (49). Similarly, when  $\vec{\theta}$  is real,  $Y_i$  is reduced to Eq. (51) and we always have real solutions.

### 5.2.3 The time-dependent variational principle

The Lagrangian for the imaginary time evolution is

$$L = \text{Tr} \left[ \rho(\vec{\theta}(t))^\dagger \left( \frac{d\rho(\vec{\theta}(t))}{dt} + \{H, \rho(\vec{\theta}(t))\} - 2 \langle H \rangle \rho(\vec{\theta}(t)) \right) \right]. \quad (53)$$

When  $\vec{\theta}$  is complex, we can regard  $\rho(\vec{\theta}(t))^\dagger$  and  $\rho(\vec{\theta}(t))$  as independent parameters and hence recover the equation defined in Eq. (46). Applying the Euler-Lagrangian equation to  $\theta^*$ , we can obtain the evolution the same as in Eq. (49). Instead, when  $\vec{\theta}$  is real, we cannot regard  $\rho(\vec{\theta}(t))^\dagger$  and  $\rho(\vec{\theta}(t))$  as independent parameters and thus we cannot recover Eq. (46) or obtain the evolution equation of the parameters  $\vec{\theta}$ .

As a special case, suppose  $\rho(\vec{\theta}(t))$  is a pure state  $|\phi(\vec{\theta}(t))\rangle$ , we have the same  $M_{i,j}$  as in real time evolution given in Eq. (35) and  $Y_i = -C_i^R$ .

## 5.3 Variational Gibbs state preparation

Evolving the maximally mixed state  $I_d/d$  with imaginary time  $\tau$  can be used to prepare the thermal state  $\rho(T) = e^{-H/T}/\text{Tr}[e^{-H/T}]$  with temperature  $T = 1/2\tau$ . Here,  $d$  is the dimension of the system. Variational classical simulation of finite-temperature Gibbs state with matrix product states has been introduced in Ref. [5]. For variational quantum simulation, we cannot simply apply a variational unitary circuit to the maximally mixed state and change parameters in the unitary circuit realise the simulation. This is because  $I_d/d$  is invariant under unitary,  $U(\theta) \cdot I_d/d \cdot U(\theta)^\dagger = I_d/d$ . Instead, we can input a maximally entangled state  $|\Phi\rangle_d = 1/\sqrt{d} \sum_i |ii\rangle_{AE}$  of system  $AE$  and evolve the whole system with Hamiltonian  $H \otimes I_d$  under imaginary time  $\tau$ . Then, the state of system  $A$  at time  $\tau$  will be the thermal state with temperature  $T = 1/2\tau$ .

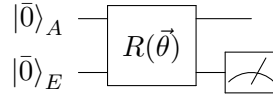
## 6 Implementation

In this section, we first show how to realise pure and mixed trial states with quantum circuits possibly assisted with post-selection on measurement outcomes. Then we show how to measure each term of  $M$ ,  $V$ , and  $Y$  for simulating general evolutions of mixed states.

### 6.1 Trial state implementation

#### 6.1.1 Pure states

We first consider the pure state case. The conventional way of realising the trial state is to apply a sequence of parameterised gates to an initial state as  $|\phi(\vec{\theta})\rangle = R_N(\theta_N) \dots R_k(\theta_k) \dots R_1(\theta_1) |\bar{0}\rangle$ . In practice, not only the gate, but also the measurement performed to the state can be also parametrised, which however can always be effectively realised by applying a parameterised gate before a fixed measurement. Therefore, we also regard parametrised measurements equivalently as parametrised gates. Next, we consider state preparation via post-selection of measurement outcomes. Specifically, after initialising the target system together with ancillae, we can apply a joint parametrised circuit, measure the ancillae, and post-select the trial state conditioned on the measurement outcome. The quantum circuit of such a procedure is as follows.



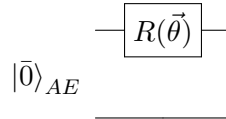
As we only consider a pure trial state, the measurement should be a rank one projector, such as  $|0\rangle\langle 0|_E$  without loss of generality. Therefore, the trial state is

$$|\phi_0(\vec{\theta})\rangle_A = \langle 0|_E |\phi(\vec{\theta})\rangle_{AE} / \sqrt{p(\vec{\theta})}, \quad (54)$$

with the joint state before the measurement  $|\phi(\vec{\theta})\rangle = R(\vec{\theta}) |\bar{0}\rangle_A |\bar{0}\rangle_E$ , and  $p(\vec{\theta}) = |\langle \phi(\vec{\theta})|_{AE} |0\rangle_E|^2$  being the post-selection probability of measuring 0 of the ancillae.

#### 6.1.2 Mixed states and unitary evolution

For mixed states under unitary evolution, we can design the trial states via two different ways. We can either input a mixed state  $\rho(0)$  and apply a unitary circuit  $R(\vec{\theta})$  to prepare the ansatz. Equivalently, we can input a state  $|\bar{0}\rangle_{AE}$  of system  $AE$  that is a purification of system  $A$  satisfying  $\text{Tr}_E[|\bar{0}\rangle_{AE}\langle\bar{0}|_{AE}] = \rho_A(0)$  with partial trace  $\text{Tr}_E$  of system  $E$ . Note that there is not a unique representation of the purified state and different choices may lead to greater or lower complexity of the simulation. With the whole purified state, we only need to apply the circuit to system  $A$  such as the following one.



In this case, we use one trial state  $\rho(\vec{\theta}(t)) = \text{Tr}_E[R(\vec{\theta}) |\bar{0}\rangle_{AE} \langle\bar{0}|_{AE} R^\dagger(\vec{\theta})]$  with one parameter setting  $\vec{\theta}$  to directly represent the state  $\rho(t)$  at time  $t$ . Here, we can also consider more complicated ansatz with measurements such as Eq. (54). Alternatively, we can decompose  $\rho$  into pure states and simulate the unitary evolution of each pure state separately. Suppose the initial state  $\rho(0)$  has a decomposition  $\rho(0) = \sum_i p_i |\psi_i(0)\rangle \langle\psi_i(0)|$ . In practice, one can randomly input state  $|\psi_i(0)\rangle$  with probability  $p_i$  and evolve  $|\psi_i(t)\rangle$  to time  $t$  under the Hamiltonian  $H$ . Then the state  $\rho(t)$  at time  $t$  would be  $\rho(t) =$

$\sum_i p_i |\psi_i(t)\rangle \langle \psi_i(t)|$ . With VQS,  $|\psi_i(t)\rangle$  is represented by the trial state  $|\phi(\vec{\theta}_i(t))\rangle$  with parameters  $\vec{\theta}_i(t)$ . As the ansatz circuit  $R(\vec{\theta})$  and parameters  $\vec{\theta}$  can be different for the evolution of each pure state, evolving the pure states separately could lead to more accurate results than evolving the whole mixed state with one parameter settings. However, as the evolution equation for one trial state is obtained by minimising the distance of the whole state, evolving the pure states separately may not always lead more accurate simulation. Furthermore, it can also be practically hard to decompose the initial state into a mixture of pure states.

### 6.1.3 Stochastic and imaginary time evolution

For stochastic and imaginary time evolution evolutions, as they are generally not reversible, we have to introduce ancillae and jointly apply the parameterised circuit. We can also measure a subset of the ancillae to postselect the trial state. However, we can only use one trial state and it cannot be naively simulated by decomposing  $\rho(0)$  into pure states as  $\rho(0) = \sum_i p_i |\psi_i(0)\rangle \langle \psi_i(0)|$  and evolving each pure state separately. This is because stochastic and imaginary time evolution is not a linear process acting on the initial state so that  $\rho(T) \neq \sum_i p_i \rho_i(T)$  for real time  $T = t$  and imaginary time  $T = \tau$ . Here  $\rho(T)$  and  $\rho_i(T)$  are the evolved state of  $\rho(0)$  and  $|\psi_i(0)\rangle$  at time  $T$ , respectively.

In Ref. [60], we do present an alternative method to simulate the stochastic master equation with different pure states evolving with different parameters. While the simulation involves several new techniques including variational simulation of general unphysical processes, we refer the readers to Ref. [60] for more information.

## 6.2 Coefficients evaluation

Now, we consider how to measure the coefficients  $M$ ,  $V$ , and  $Y$  in the evolution equations. For unitary evolution of mixed states, we can effective regard the input state as a pure state and only evolve the subsystem of  $\rho_0$ . Suppose the trial state is prepared by directly applying a unitary circuit to the initial state, the  $M$ ,  $V$ , and  $Y$  can all be measured with the circuit in Fig. 1. Instead, when the trial state is prepared conditioned on postselecting a measurement outcome as defined in Eq. (54), we need to reconsider how to measure the coefficients. The derivative of the state can be written as

$$\frac{\partial |\phi_0(\vec{\theta})\rangle_A}{\partial \theta_k} = \frac{\langle 0|_E \frac{\partial |\phi(\vec{\theta}(t))\rangle_{AE}}{\partial \theta_k}}{\sqrt{p(\vec{\theta})}} - \frac{\langle 0|_E |\phi(\vec{\theta}(t))\rangle_{AE} \frac{\partial p(\vec{\theta})}{\partial \theta_k}}{2p^{3/2}(\vec{\theta})} \quad (55)$$

The probability  $p(\vec{\theta})$  can be directly measured and the derivative  $\frac{\partial p(\vec{\theta})}{\partial \theta_k}$  can be measured by the circuit in Fig. 1 To evaluate  $M$ ,  $V$ , and  $Y$ , we need to measure the following terms for all  $i$  and  $j$ ,

$$\begin{aligned} & \Re \left( \frac{\partial \langle \phi(\vec{\theta}(t)) |}{\partial \theta_i} \frac{\partial |\phi(\vec{\theta}(t))\rangle}{\partial \theta_j} \right), \Im \left( \frac{\partial \langle \phi(\vec{\theta}(t)) |}{\partial \theta_i} H |\phi(\vec{\theta}(t))\rangle \right), \\ & \Im \left( \frac{\partial \langle \phi(\vec{\theta}(t)) |}{\partial \theta_i} |\phi(\vec{\theta}(t))\rangle \right), \Re \left( \frac{\partial \langle \phi(\vec{\theta}(t)) |}{\partial \theta_i} H |\phi(\vec{\theta}(t))\rangle \right). \end{aligned} \quad (56)$$

All these terms can be efficiently measured with the circuit shown in Fig. 1 as they are in the form of elements of either  $A$  or  $C$  as defined in Eq. (10). Note that the precision of the estimated derivatives depend on the size of the system or the number of measurements, which can affect the accuracy of the simulation. We refer to Ref. [41] for a detailed analysis of the scaling in terms of the system size and the simulation accuracy.

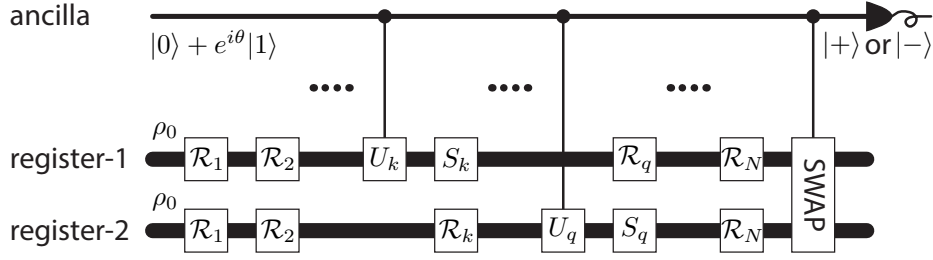


Figure 3: Circuit for the evaluation of differential-equation coefficients of the mixed-state simulator. To evaluate  $\Re[e^{i\theta}\text{Tr}(\rho_1\rho_2)]$ , where  $\rho_1 = \mathcal{R}_N\mathcal{R}_{N-1}\cdots\mathcal{R}_{k+1}[S_k(\mathcal{R}_{k-1}\cdots\mathcal{R}_2\mathcal{R}_1\rho_0)T_k^\dagger]$ ,  $\rho_2 = \mathcal{R}_N\mathcal{R}_{N-1}\cdots\mathcal{R}_{q+1}[S_q(\mathcal{R}_{q-1}\cdots\mathcal{R}_2\mathcal{R}_1\rho_0)T_q^\dagger]$ , the ancillary qubit is initialised in the state  $(|0\rangle + e^{i\theta}|1\rangle)/\sqrt{2}$  and measured in the  $|\pm\rangle = (|0\rangle \pm |1\rangle)/\sqrt{2}$  basis. Here,  $U_k = T_k S_k^\dagger$  and  $U_q = T_q S_q^\dagger$ . The last gate is a controlled swap gate on two registers. In the figure, we have assumed that  $k < q$ , and  $q = N + 1$  when the circuit is used to evaluate  $\tilde{V}_k$  coefficients.

Now, we consider the case of general evolutions. Suppose the trial state is prepared by  $\rho = \mathcal{R}_N\mathcal{R}_{N-1}\cdots\mathcal{R}_1(\rho_0)$ , where  $\rho_0$  is a joint state of the system and ancillae and  $\mathcal{R}_k$  are quantum gates that applied on the whole system. We express the generator as

$$\mathcal{L}(\rho) = \sum_i g_i S_i \rho T_i^\dagger, \quad (57)$$

where  $S_i$  and  $T_i$  are unitary operators, and  $g_i$  are coefficients. Similarly, we write

$$\frac{\partial \mathcal{R}_k(\rho)}{\partial \lambda_k} = \sum_i r_{k,i} S_{k,i} \rho T_{k,i}^\dagger, \quad (58)$$

where  $S_{k,i}$  and  $T_{k,i}$  are unitary operators, and  $r_{k,i}$  are coefficients. For example, consider gate  $\mathcal{R}_k(\theta_k) = e^{-i\theta_k\sigma/2}\rho e^{i\theta_k\sigma/2}$  with a one or two qubit Hermitian operator  $\sigma$ , the derivative is  $\partial \mathcal{R}_k(\rho)/\partial \lambda_k = -i\sigma/2 \cdot e^{-i\theta_k\sigma/2}\rho e^{i\theta_k\sigma/2} + e^{-i\theta_k\sigma/2}\rho e^{i\theta_k\sigma/2} \cdot i\sigma/2$ .

Using the expression (58), we rewrite the derivative of the mixed state as

$$\frac{\partial \rho}{\partial \lambda_k} = \sum_i r_{k,i} \mathcal{R}_{k,i} \rho_0, \quad (59)$$

where

$$\mathcal{R}_{k,i} \rho_0 = \mathcal{R}_N\mathcal{R}_{N-1}\cdots\mathcal{R}_{k+1}[S_{k,i}(\mathcal{R}_{k-1}\cdots\mathcal{R}_2\mathcal{R}_1\rho_0)T_{k,i}^\dagger]. \quad (60)$$

Then, using the expression (57), coefficients can be expressed as

$$\begin{aligned} M_{k,q} &= \sum_{i,j} \frac{1}{2} \{r_{k,i} r_{q,j} \text{Tr}[(\mathcal{R}_{k,i}\rho_0)(\mathcal{R}_{q,j}\rho_0)] + \text{h.c.}\}, \\ V_k &= \sum_{i,j} \frac{1}{2} \{r_{k,i} g_j \text{Tr}[(\mathcal{R}_{k,i}\rho_0)(S_j \rho T_j^\dagger)] + \text{h.c.}\}, \\ Y_k &= - \sum_i \frac{1}{2} \{r_{k,i} \text{Tr}[(\mathcal{R}_{k,i}\rho_0)\rho H] + \text{h.c.}\} \end{aligned} \quad (61)$$

In Eq. (61), each term is in the form

$$a \Re[e^{i\theta} \text{Tr}(\rho_1 \rho_2)],$$

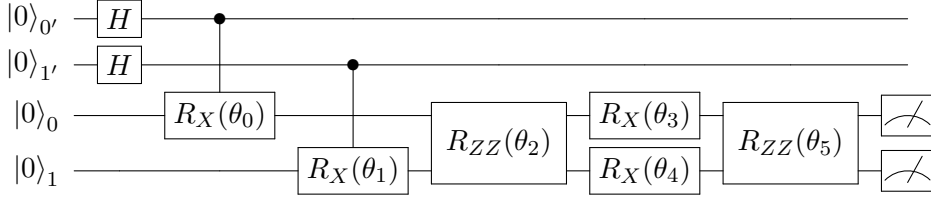


Figure 4: The ansatz used in our numerical simulation. We introduced two ancillary qubits  $0'1'$  as a purification of the two system qubits  $01$ . We define  $R_X(\theta_i) = e^{-i\theta_i X}$  and  $R_{ZZ}(\theta) = e^{-i\theta_i Z_1 \otimes Z_2}$ , which are the evolution of the system Hamiltonian terms. We used six parameters in total.

where the amplitude  $a$  and phase  $\theta$  are determined by either  $r_{k,i}r_{q,j}$  or  $r_{k,i}g_j$ . We would like to remark that, in general,  $\rho_1$  and  $\rho_2$  are not reduced density matrices as  $S$  and  $T$  are separately applied to the left and right sides of the state. Nevertheless, such a term can be evaluated with the quantum circuit shown in Fig. 3. This circuit needs an ancillary qubit initialised in the state  $(|0\rangle + e^{i\theta}|1\rangle)/\sqrt{2}$  and two registers initialised in the state  $\rho_0$ . The ancillary qubit is measured after a sequence of operations on registers and some controlled unitary operations, where the ancillary qubit is the control qubit. The value is given by  $\Re[e^{i\theta}\text{Tr}(\rho_1\rho_2)] = 2P_+ - 1$ , where  $P_+$  is the probability that the qubit is in the state  $|+\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$ . We can also consider trial state prepared by post-selection and the matrix elements can be similarly evaluated.

### 6.3 Numerical simulation

Here, we show a numerical example of the variational algorithm for simulating the real dynamics of open quantum systems. We consider a two-qubit Ising model under independent amplitude damping noise with the Lindblad master equation

$$\frac{d}{dt}\rho = -i[H, \rho] + \sum_{j=0}^1 \mathcal{L}_j(\rho). \quad (62)$$

Here the system Hamiltonian is  $H = X_0 + X_1 + \frac{1}{4}Z_0Z_1$ , the Lindblad terms are  $\mathcal{L}_j(\rho) = \frac{1}{2}(2\sigma_j^- \rho \sigma_j^+ - \sigma_j^+ \sigma_j^- \rho - \rho \sigma_j^+ \sigma_j^-)$ , and  $\sigma_j^- = |0\rangle\langle 1|_j$  and  $\sigma_j^+ = |1\rangle\langle 0|_j$  are lowering and raising operators acting on the  $j^{\text{th}}$  qubit, respectively. We start with a product initial state  $\rho_0 = |00\rangle\langle 00|$  and simulate the dynamics from  $t = 0$  to  $t = 10$ .

For the variational algorithm, we consider the Hamiltonian ansatz as shown in Fig. 4, where we introduced two ancillary qubits to purify the state. For the ansatz, we first apply parametrised gates to entangle the ancillae and the system qubits; then we apply the Hamiltonian ansatz, which consists of parameterised gates determined by the Hamiltonian of the system, to the system qubits. In general, the ansatz can also depend on the Lindblad terms and we leave the design of ansatz for general open system evolutions in a future work.

To simulate the evolution of Eq. (62), we consider discretised timestep  $\delta t = 0.01$ . For the initial step, we set all the parameters to 0; for the later steps, we evaluate all terms of  $M$  and  $V$  defined in Eq. (29) and solve Eq. (28) to update the parameters.

The simulation result is shown in Fig. 5. We also numerically solve the exact evolution of Eq. (62) and compare it with the result obtained from our variational algorithm. We measure the expectation value of  $Z_0$  and we found excellent agreement between the results from the exact solution and our variational algorithm with a deviation less than  $10^{-2}$ .

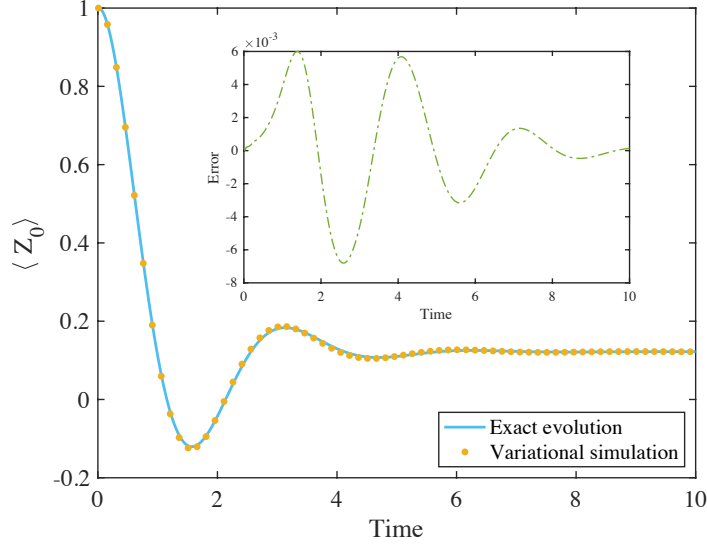


Figure 5: The comparison between the exact result and the result obtained from variational algorithm. We set the time step  $\delta t = 0.01$  and simulated from  $t = 0$  to  $t = 10$ . We used a software package QuTIP for the numerical simulation [61, 62].

## 7 Discussion

In this work, we focus on the theory of variational quantum simulation. We first study the equivalence and difference of the three variational principles. We find McLachlan’s variational principle is the most consistent principle for variational quantum simulation with quantum gates controlled by real parameters. Then, we extend variational quantum simulation of pure states to the general evolution of mixed states under both real and imaginary time. We discuss possible realisation of the trial states and show how to efficiently implement the simulation with quantum circuits.

In future works, one can study the design of trials states for specific problems and test our theory for simulating open quantum systems and preparing Gibbs states. It is also interesting to experimentally realise our variational simulation algorithms with current and near-term noisy quantum hardware. As the variational method only uses shallow quantum circuits, error mitigation techniques can be applied to suppress errors [41, 63–70].

## Acknowledgements.

This work is supported by the EPSRC National Quantum Technology Hub in Networked Quantum Information Technology (EP/M013243/1). SE is supported by Japan Student Services Organization (JASSO) Student Exchange Support Program (Graduate Scholarship for Degree Seeking Students). QZ acknowledges support by the National Natural Science Foundation of China Grant No. 11674193. YL is supported by NSAF (Grant No. U1730449).

## Appendix

### A Real time evolution

#### A.1 Real time evolution: pure state and unitary evolution

##### A.1.1 The Dirac and Frenkel variational principle

The Schrödinger equation is,

$$\frac{d|\psi(t)\rangle}{dt} = -iH|\psi(t)\rangle. \quad (63)$$

Consider a parametrised trial state  $|\phi(\vec{\theta}(t))\rangle$ , with  $\vec{\theta}(t) = (\theta_1(t), \theta_2(t), \dots, \theta_N(t))$ , the real time evolution of the Schrödinger equation on the trial state space is

$$\frac{d|\phi(\vec{\theta}(t))\rangle}{dt} = \sum_i \frac{\partial |\phi(\vec{\theta}(t))\rangle}{\partial \theta_i} \dot{\theta}_i = -iH|\phi(\vec{\theta}(t))\rangle. \quad (64)$$

Such an equation has a state vector on both sides. While, state  $\frac{\partial |\phi(\vec{\theta}(t))\rangle}{\partial \theta_i}$  can be regarded as the tangent vector of state  $|\phi(\vec{\theta}(t))\rangle$  at  $\vec{\theta}(t)$ , we can thus apply a projector

$$P = \sum_i \frac{\partial |\phi(\vec{\theta}(t))\rangle}{\partial \theta_i} \frac{\partial \langle \phi(\vec{\theta}(t))|}{\partial \theta_i} \quad (65)$$

to project the right hand side onto the tangent space. Therefore, the equation becomes,

$$\begin{aligned} P \sum_j \frac{\partial |\phi(\vec{\theta}(t))\rangle}{\partial \theta_j} \dot{\theta}_j &= -iPH|\phi(\vec{\theta}(t))\rangle, \\ i \sum_i \frac{\partial |\phi(\vec{\theta}(t))\rangle}{\partial \theta_i} \frac{\partial \langle \phi(\vec{\theta}(t))|}{\partial \theta_i} \sum_j \frac{\partial |\phi(\vec{\theta}(t))\rangle}{\partial \theta_j} \dot{\theta}_j &= -i \sum_i \frac{\partial |\phi(\vec{\theta}(t))\rangle}{\partial \theta_i} \frac{\partial \langle \phi(\vec{\theta}(t))|}{\partial \theta_i} H|\phi(\vec{\theta}(t))\rangle. \end{aligned} \quad (66)$$

For each term of  $\frac{\partial |\phi(\vec{\theta}(t))\rangle}{\partial \theta_i}$ , the coefficient should satisfy

$$\sum_j \frac{\partial \langle \phi(\vec{\theta}(t))|}{\partial \theta_i} \frac{\partial |\phi(\vec{\theta}(t))\rangle}{\partial \theta_j} \dot{\theta}_j = -i \frac{\partial \langle \phi(\vec{\theta}(t))|}{\partial \theta_i} H|\phi(\vec{\theta}(t))\rangle. \quad (67)$$

Define the matrix elements of  $M$  and  $V$  as

$$\begin{aligned} A_{i,j} &= \frac{\partial \langle \phi(\vec{\theta}(t))|}{\partial \theta_i} \frac{\partial |\phi(\vec{\theta}(t))\rangle}{\partial \theta_j}, \\ C_i &= \frac{\partial \langle \phi(\vec{\theta}(t))|}{\partial \theta_i} H|\phi(\vec{\theta}(t))\rangle, \end{aligned} \quad (68)$$

the evolution of parameters is simplified to

$$\sum_j A_{i,j} \dot{\theta}_j = -iC_i. \quad (69)$$

The same equation can be obtained by applying the Dirac and Frenkel variational principle

$$\left\langle \delta\phi(\vec{\theta}(t)) \left| \left( \frac{d}{dt} + iH \right) \right| \phi(\vec{\theta}(t)) \right\rangle = 0. \quad (70)$$



This is can be verified with

$$\langle \delta \phi(\vec{\theta}(t)) | = \sum_i \frac{\partial \langle \phi(\vec{\theta}(t)) |}{\partial \theta_i} \delta \theta_i, \quad (71)$$

and Eq. (64).

### A.1.2 McLachlan's variational principle

The McLachlan's variational principle [50], applied to real time evolution, is given by

$$\delta \| (d/dt + iH) |\phi(\vec{\theta}(t))\rangle \| = 0 \quad (72)$$

where

$$\begin{aligned} \| (d/dt + iH) |\phi(\vec{\theta}(t))\rangle \|^2 &= (d/dt + iH) |\psi(t)\rangle^\dagger (d/dt + iH) |\phi(\vec{\theta}(t))\rangle, \\ &= \sum_{i,j} \frac{\partial \langle \phi(\vec{\theta}(t)) |}{\partial \theta_i} \frac{\partial |\phi(\vec{\theta}(t))\rangle}{\partial \theta_j} \dot{\theta}_i^* \dot{\theta}_j + i \sum_i \frac{\partial \langle \phi(\vec{\theta}(t)) |}{\partial \theta_i} H |\phi(\vec{\theta}(t))\rangle \dot{\theta}_i^* \\ &\quad - i \sum_i \langle \phi(\vec{\theta}(t)) | H \frac{\partial |\phi(\vec{\theta}(t))\rangle}{\partial \theta_i} \dot{\theta}_i + \langle \phi(\vec{\theta}(t)) | H^2 |\phi(\vec{\theta}(t))\rangle. \end{aligned} \quad (73)$$

Note that the variation respect to  $\| (d/dt + iH) |\phi(\vec{\theta}(t))\rangle \|$  is equivalent to the variation respect to  $\| (d/dt + iH) |\phi(\vec{\theta}(t))\rangle \|^2$ , so we focus on  $\| (d/dt + iH) |\phi(\vec{\theta}(t))\rangle \|^2$  instead.

Suppose  $\dot{\theta}_i$  can be complex, then we have

$$\begin{aligned} \delta \| (d/dt + iH) |\phi(\vec{\theta}(t))\rangle \|^2 &= \left( \sum_j \frac{\partial \langle \phi(\vec{\theta}(t)) |}{\partial \theta_i} \frac{\partial |\phi(\vec{\theta}(t))\rangle}{\partial \theta_j} \dot{\theta}_j + i \frac{\partial \langle \phi(\vec{\theta}(t)) |}{\partial \theta_i} H |\phi(\vec{\theta}(t))\rangle \right) \delta \dot{\theta}_i^*, \\ &\quad + \left( \sum_j \frac{\partial \langle \phi(\vec{\theta}(t)) |}{\partial \theta_j} \frac{\partial |\phi(\vec{\theta}(t))\rangle}{\partial \theta_i} \dot{\theta}_j^* - i \frac{\partial \langle \phi(\vec{\theta}(t)) |}{\partial \theta_i} H |\phi(\vec{\theta}(t))\rangle \right) \delta \dot{\theta}_i. \end{aligned} \quad (74)$$

Then the evolution is

$$\sum_j \frac{\partial \langle \phi(\vec{\theta}(t)) |}{\partial \theta_i} \frac{\partial |\phi(\vec{\theta}(t))\rangle}{\partial \theta_j} \dot{\theta}_j = -i \frac{\partial \langle \phi(\vec{\theta}(t)) |}{\partial \theta_i} H |\phi(\vec{\theta}(t))\rangle. \quad (75)$$

On the other hand, suppose  $\dot{\theta}_i$  is real, then

$$\begin{aligned} \delta \| (d/dt + iH) |\phi(\vec{\theta}(t))\rangle \|^2 &= \sum_j \left( \frac{\partial \langle \phi(\vec{\theta}(t)) |}{\partial \theta_i} \frac{\partial |\phi(\vec{\theta}(t))\rangle}{\partial \theta_j} + \frac{\partial \langle \phi(\vec{\theta}(t)) |}{\partial \theta_j} \frac{\partial |\phi(\vec{\theta}(t))\rangle}{\partial \theta_i} \right) \dot{\theta}_j \delta \theta_i \\ &\quad + i \left( \frac{\partial \langle \phi(\vec{\theta}(t)) |}{\partial \theta_i} H |\phi(\vec{\theta}(t))\rangle - \langle \phi(\vec{\theta}(t)) | H \frac{\partial |\phi(\vec{\theta}(t))\rangle}{\partial \theta_i} \right) \delta \theta_i. \end{aligned} \quad (76)$$

The corresponding evolution equation for parameters is

$$\begin{aligned} &\sum_j \left( \frac{\partial \langle \phi(\vec{\theta}(t)) |}{\partial \theta_i} \frac{\partial |\phi(\vec{\theta}(t))\rangle}{\partial \theta_j} + \frac{\partial \langle \phi(\vec{\theta}(t)) |}{\partial \theta_j} \frac{\partial |\phi(\vec{\theta}(t))\rangle}{\partial \theta_i} \right) \dot{\theta}_j \\ &= -i \left( \frac{\partial \langle \phi(\vec{\theta}(t)) |}{\partial \theta_i} H |\phi(\vec{\theta}(t))\rangle - \langle \phi(\vec{\theta}(t)) | H \frac{\partial |\phi(\vec{\theta}(t))\rangle}{\partial \theta_i} \right) \end{aligned} \quad (77)$$

or equivalently

$$\sum_j A_{i,j}^R \dot{\theta}_j = C_i^I, \quad (78)$$

where  $A_{i,j}^R$  and  $C_i^I$  are the real and imaginary parts of  $A_{i,j}$  and  $C_i$ , respectively.

### A.1.3 Time-dependent variational principles

The Lagrangian for the Schrödinger equation is

$$L = \left\langle \psi \left| \left( \frac{d}{dt} + iH \right) \right| \psi \right\rangle, \quad (79)$$

and the Schrödinger equation is obtained by

$$\frac{\partial L}{\partial \langle \psi |} - \frac{d}{dt} \frac{\partial L}{\partial \frac{\partial \langle \psi |}{\partial t}} = \left( \frac{d}{dt} + iH \right) |\psi\rangle = 0. \quad (80)$$

Suppose  $|\psi\rangle = |\phi(\vec{\theta}(t))\rangle$ , then the Lagrangian is

$$\begin{aligned} L &= \left\langle \phi(\vec{\theta}(t)) \left| \left( \frac{d}{dt} + iH \right) \right| \phi(\vec{\theta}(t)) \right\rangle, \\ &= \sum_i \dot{\theta}_i \left\langle \phi(\vec{\theta}(t)) \left| \frac{\partial}{\partial \theta_i} \right| \phi(\vec{\theta}(t)) \right\rangle + i \left\langle \phi(\vec{\theta}(t)) \left| H \right| \phi(\vec{\theta}(t)) \right\rangle \end{aligned} \quad (81)$$

Suppose  $\theta$  is complex, then

$$\frac{\partial L}{\partial \theta_i^*} - \frac{d}{dt} \frac{\partial L}{\partial \dot{\theta}_i^*} = \sum_j \dot{\theta}_j \frac{\partial \langle \phi(\vec{\theta}(t)) |}{\partial \theta_i} \frac{\partial |\phi(\vec{\theta}(t))\rangle}{\partial \theta_j} + i \frac{\partial \langle \phi(\vec{\theta}(t)) |}{\partial \theta_i} H |\phi(\vec{\theta}(t))\rangle, \quad (82)$$

and the evolution is

$$\sum_j \frac{\partial \langle \phi(\vec{\theta}(t)) |}{\partial \theta_i} \frac{\partial |\phi(\vec{\theta}(t))\rangle}{\partial \theta_j} \dot{\theta}_j = -i \frac{\partial \langle \phi(\vec{\theta}(t)) |}{\partial \theta_i} H |\phi(\vec{\theta}(t))\rangle. \quad (83)$$

Instead, suppose  $\theta$  is real, then we have

$$\begin{aligned} \frac{\partial L}{\partial \theta_i} - \frac{d}{dt} \frac{\partial L}{\partial \dot{\theta}_i} &= \sum_j \dot{\theta}_j \frac{\partial \left\langle \phi(\vec{\theta}(t)) \left| \frac{\partial}{\partial \theta_j} \right| \phi(\vec{\theta}(t)) \right\rangle}{\partial \theta_i} + i \frac{\partial \left\langle \phi(\vec{\theta}(t)) \left| H \right| \phi(\vec{\theta}(t)) \right\rangle}{\partial \theta_i} - \frac{d}{dt} \left\langle \phi(\vec{\theta}(t)) \left| \frac{\partial}{\partial \theta_i} \right| \phi(\vec{\theta}(t)) \right\rangle, \\ &= \sum_j \dot{\theta}_j \left( \frac{\partial \langle \phi(\vec{\theta}(t)) |}{\partial \theta_i} \frac{\partial |\phi(\vec{\theta}(t))\rangle}{\partial \theta_j} + \left\langle \phi(\vec{\theta}(t)) \left| \frac{\partial^2}{\partial \theta_i \partial \theta_j} \right| \phi(\vec{\theta}(t)) \right\rangle \right) + i \frac{\partial \langle \phi(\vec{\theta}(t)) |}{\partial \theta_i} H |\phi(\vec{\theta}(t))\rangle \\ &\quad + i \langle \phi(\vec{\theta}(t)) | H \frac{\partial |\phi(\vec{\theta}(t))\rangle}{\partial \theta_i} - \sum_j \dot{\theta}_j \left( \frac{\partial \langle \phi(\vec{\theta}(t)) |}{\partial \theta_j} \frac{\partial |\phi(\vec{\theta}(t))\rangle}{\partial \theta_i} + \left\langle \phi(\vec{\theta}(t)) \left| \frac{\partial^2}{\partial \theta_i \partial \theta_j} \right| \phi(\vec{\theta}(t)) \right\rangle \right), \\ &= \sum_j \dot{\theta}_j \left( \frac{\partial \langle \phi(\vec{\theta}(t)) |}{\partial \theta_i} \frac{\partial |\phi(\vec{\theta}(t))\rangle}{\partial \theta_j} - \frac{\partial \langle \phi(\vec{\theta}(t)) |}{\partial \theta_j} \frac{\partial |\phi(\vec{\theta}(t))\rangle}{\partial \theta_i} \right) \\ &\quad + i \frac{\partial \langle \phi(\vec{\theta}(t)) |}{\partial \theta_i} H |\phi(\vec{\theta}(t))\rangle + i \langle \phi(\vec{\theta}(t)) | H \frac{\partial |\phi(\vec{\theta}(t))\rangle}{\partial \theta_i}, \end{aligned} \quad (84)$$

and

$$\begin{aligned} &\sum_j \left( \frac{\partial \langle \phi(\vec{\theta}(t)) |}{\partial \theta_i} \frac{\partial |\phi(\vec{\theta}(t))\rangle}{\partial \theta_j} - \frac{\partial \langle \phi(\vec{\theta}(t)) |}{\partial \theta_j} \frac{\partial |\phi(\vec{\theta}(t))\rangle}{\partial \theta_i} \right) \dot{\theta}_j \\ &= -i \left( \frac{\partial \langle \phi(\vec{\theta}(t)) |}{\partial \theta_i} H |\phi(\vec{\theta}(t))\rangle + \langle \phi(\vec{\theta}(t)) | H \frac{\partial |\phi(\vec{\theta}(t))\rangle}{\partial \theta_i} \right), \end{aligned} \quad (85)$$

Equivalently, the evolution to parameters is

$$\sum_j A_{i,j}^I \dot{\theta}_j = -C_i^R, \quad (86)$$

where,  $A_{i,j}^I$  and  $C_i^R$  are the imaginary and real parts of  $A_{i,j}$  and  $C_i$ , respectively.

## A.2 Mixed states and general evolution

### A.2.1 General evolution

Suppose the initial state is a mixed state  $\rho$ , the stochastic evolution is defined by

$$\frac{d\rho}{dt} = \mathcal{L}(\rho). \quad (87)$$

Consider a parameterised trial state  $\rho(\vec{\theta}(t))$ , we can get the evolution of  $\theta$  that simulates the stochastic evolution.

**The time-dependent variational principle.** Let  $\rho = \rho(\vec{\theta}(t))$ , the stochastic evolution equation is

$$\sum_i \frac{\partial \rho(\vec{\theta}(t))}{\partial \theta_i} \dot{\theta}_i = \mathcal{L}(\rho). \quad (88)$$

Because  $\frac{d\rho(\vec{\theta}(t))}{dt} = \sum_i \frac{\partial \rho(\vec{\theta}(t))}{\partial \theta_i} \dot{\theta}_i$  is in the tangent subspace of  $\{\frac{\partial \rho(\vec{\theta}(t))}{\partial \theta_i}\}$ , we need to project the stochastic evolution equation onto this tangent subspace. Following the Dirac and Frenkel variational principle, we have

$$\text{Tr} \left[ (\delta \rho(\vec{\theta}(t)))^\dagger \left( \frac{d\rho}{dt} - \mathcal{L}(\rho) \right) \right] = \text{Tr} \left[ \left( \sum_i \frac{\partial \rho(\vec{\theta}(t))}{\partial \theta_i} \delta \theta_i \right)^\dagger \left( \sum_j \frac{\partial \rho}{\partial \theta_j} \dot{\theta}_j - \mathcal{L}(\rho) \right) \right] = 0. \quad (89)$$

The evolution of parameters is

$$M_{i,j} \dot{\theta}_j = V_i, \quad (90)$$

with

$$\begin{aligned} M_{i,j} &= \text{Tr} \left[ \left( \frac{\partial \rho(\vec{\theta}(t))}{\partial \theta_i} \right)^\dagger \frac{\partial \rho(\vec{\theta}(t))}{\partial \theta_j} \right], \\ V_i &= \text{Tr} \left[ \left( \frac{\partial \rho(\vec{\theta}(t))}{\partial \theta_i} \right)^\dagger \mathcal{L}(\rho) \right]. \end{aligned} \quad (91)$$

Note that when parameters  $\theta$  are real,  $M$  and  $V$  are also real and the solution  $\dot{\theta}_j$  is also real.

**McLachlan's variational principle.** We can also minimise the distance between the evolution via parameters and the stochastic evolution via McLachlan's variational principle,

$$\delta \|d\rho/dt - \mathcal{L}(\rho)\| = 0, \quad (92)$$

where

$$\begin{aligned}
\|d\rho/dt - \mathcal{L}(\rho)\|^2 &= \text{Tr} \left[ (d\rho/dt - \mathcal{L}(\rho))^\dagger (d\rho/dt - \mathcal{L}(\rho)) \right], \\
&= \text{Tr} \left[ \left( \sum_i \frac{\partial \rho(\vec{\theta}(t))}{\partial \theta_i} \dot{\theta}_i \right)^\dagger \sum_j \frac{\partial \rho(\vec{\theta}(t))}{\partial \theta_j} \dot{\theta}_j - \left( \sum_i \frac{\partial \rho(\vec{\theta}(t))}{\partial \theta_i} \dot{\theta}_i \right)^\dagger \mathcal{L}(\rho) \right] \\
&\quad + \text{Tr} \left[ - \left( \sum_i \frac{\partial \rho(\vec{\theta}(t))}{\partial \theta_i} \dot{\theta}_i \right) \mathcal{L}(\rho) + \mathcal{L}(\rho)^2 \right].
\end{aligned} \tag{93}$$

Note that the variation with respect to  $\|d\rho/dt - \mathcal{L}(\rho)\|$  is equivalent to the variation with respect to  $\|d\rho/dt - \mathcal{L}(\rho)\|^2$ , so we focus on  $\|d\rho/dt - \mathcal{L}(\rho)\|^2$  instead.

Suppose  $\dot{\theta}$  is complex, then

$$\begin{aligned}
\delta \|d\rho/dt - \mathcal{L}(\rho)\|^2 &= \text{Tr} \left[ \left( \frac{\partial \rho(\vec{\theta}(t))}{\partial \theta_i} \right)^\dagger \sum_j \frac{\partial \rho(\vec{\theta}(t))}{\partial \theta_j} \dot{\theta}_j - \left( \frac{\partial \rho(\vec{\theta}(t))}{\partial \theta_i} \right)^\dagger \mathcal{L}(\rho) \right] \delta \dot{\theta}_i^* \\
&\quad + \text{Tr} \left[ \left( \sum_j \frac{\partial \rho(\vec{\theta}(t))}{\partial \theta_j} \dot{\theta}_j \right)^\dagger \frac{\partial \rho(\vec{\theta}(t))}{\partial \theta_i} - \frac{\partial \rho(\vec{\theta}(t))}{\partial \theta_i} \mathcal{L}(\rho) \right] \delta \dot{\theta}_i.
\end{aligned} \tag{94}$$

The evolution to  $\dot{\theta}$  is

$$\sum_j \text{Tr} \left[ \left( \frac{\partial \rho(\vec{\theta}(t))}{\partial \theta_i} \right)^\dagger \frac{\partial \rho(\vec{\theta}(t))}{\partial \theta_j} \right] \dot{\theta}_j = \text{Tr} \left[ \left( \frac{\partial \rho(\vec{\theta}(t))}{\partial \theta_i} \right)^\dagger \mathcal{L}(\rho) \right] \tag{95}$$

When  $\dot{\theta}$  is real, then

$$\begin{aligned}
\delta \|d\rho/dt - \mathcal{L}(\rho)\|^2 &= \text{Tr} \left[ \left( \frac{\partial \rho(\vec{\theta}(t))}{\partial \theta_i} \right) \sum_j \frac{\partial \rho(\vec{\theta}(t))}{\partial \theta_j} \dot{\theta}_j - \frac{\partial \rho(\vec{\theta}(t))}{\partial \theta_i} \mathcal{L}(\rho) \right] \delta \dot{\theta}_i \\
&\quad + \text{Tr} \left[ \left( \sum_j \frac{\partial \rho(\vec{\theta}(t))}{\partial \theta_j} \dot{\theta}_j \right) \frac{\partial \rho(\vec{\theta}(t))}{\partial \theta_i} - \frac{\partial \rho(\vec{\theta}(t))}{\partial \theta_i} \mathcal{L}(\rho) \right] \delta \dot{\theta}_i.
\end{aligned} \tag{96}$$

The evolution to  $\dot{\theta}$  is

$$\sum_j \text{Tr} \left[ \frac{\partial \rho(\vec{\theta}(t))}{\partial \theta_i} \frac{\partial \rho(\vec{\theta}(t))}{\partial \theta_j} \right] \dot{\theta}_j = \text{Tr} \left[ \frac{\partial \rho(\vec{\theta}(t))}{\partial \theta_i} \mathcal{L}(\rho) \right] \tag{97}$$

Then we can get the same equation for parameters

$$M_{i,j} \dot{\theta}_j = V_i. \tag{98}$$

**The time-dependent variational principle.** The Lagrangian for a general stochastic evolution is

$$L = \text{Tr}[\rho^\dagger (d\rho/dt - \mathcal{L}(\rho))]. \tag{99}$$

the Schrödinger equation is obtained by

$$\frac{\partial L}{\partial \rho^\dagger} - \frac{d}{dt} \frac{\partial L}{\partial \frac{\partial \rho^\dagger}{\partial t}} = d\rho/dt - \mathcal{L}(\rho) = 0. \tag{100}$$

Suppose  $\rho = \rho(\vec{\theta}(t))$ , then the Lagrangian is

$$\begin{aligned} L &= \text{Tr}[\rho(\vec{\theta}(t))^\dagger (d\rho(\vec{\theta}(t))/dt - \mathcal{L}(\rho))], \\ &= \text{Tr} \left[ \rho(\vec{\theta}(t))^\dagger \left( \sum_j \frac{\partial \rho(\vec{\theta}(t))}{\partial \theta_j} \dot{\theta}_j - \mathcal{L}(\rho) \right) \right] \end{aligned} \quad (101)$$

Suppose  $\theta$  is complex, then

$$\frac{\partial L}{\partial \theta_i^*} - \frac{d}{dt} \frac{\partial L}{\partial \dot{\theta}_i^*} = \text{Tr} \left[ \frac{\partial \rho(\vec{\theta}(t))^\dagger}{\partial \theta_i^*} \left( \sum_j \frac{\partial \rho(\vec{\theta}(t))}{\partial \theta_j} \dot{\theta}_j - \mathcal{L}(\rho) \right) \right], \quad (102)$$

and the evolution is

$$\sum_j \text{Tr} \left[ \left( \frac{\partial \rho(\vec{\theta}(t))}{\partial \theta_i} \right)^\dagger \frac{\partial \rho(\vec{\theta}(t))}{\partial \theta_j} \right] \dot{\theta}_j = \text{Tr} \left[ \left( \frac{\partial \rho(\vec{\theta}(t))}{\partial \theta_i} \right)^\dagger \mathcal{L}(\rho) \right] \quad (103)$$

Instead, suppose  $\theta$  is real, we cannot reproduce the stochastic evolution equation and hence cannot obtain the evolution equation of the parameters.

#### A.2.2 Reduction to the pure state case

Consider unitary evolution of mixed input states, the von Neumann equation has  $\mathcal{L}(\rho) = -i[H, \rho(t)]$ ,

$$\frac{d\rho(t)}{dt} = -i[H, \rho(t)]. \quad (104)$$

In this case,  $M$  and  $V$  are explicitly given by

$$\begin{aligned} M_{i,j} &= \text{Tr} \left[ \left( \frac{\partial \rho(\vec{\theta}(t))}{\partial \theta_i} \right)^\dagger \frac{\partial \rho(\vec{\theta}(t))}{\partial \theta_j} \right], \\ V_i &= -i \text{Tr} \left[ \left( \frac{\partial \rho(\vec{\theta}(t))}{\partial \theta_i} \right)^\dagger [H, \rho(\vec{\theta}(t))] \right]. \end{aligned} \quad (105)$$

Suppose  $\rho(\vec{\theta}(t))$  is a pure state  $|\phi(\vec{\theta}(t))\rangle \langle \phi(\vec{\theta}(t))|$ , the von Neumann equation becomes the Schrödinger

equation. For the variational evolution equation, we have

$$\begin{aligned}
M_{i,j} &= \text{Tr} \left[ \frac{\partial \rho(\vec{\theta}(t))}{\partial \theta_i} \frac{\partial \rho(\vec{\theta}(t))}{\partial \theta_j} \right], \\
&= \text{Tr} \left[ \left( \frac{\partial |\phi(\vec{\theta}(t))\rangle}{\partial \theta_i} \langle \phi(\vec{\theta}(t))| + |\phi(\vec{\theta}(t))\rangle \frac{\partial \langle \phi(\vec{\theta}(t))|}{\partial \theta_i} \right) \left( \frac{\partial |\phi(\vec{\theta}(t))\rangle}{\partial \theta_j} \langle \phi(\vec{\theta}(t))| + |\phi(\vec{\theta}(t))\rangle \frac{\partial \langle \phi(\vec{\theta}(t))|}{\partial \theta_j} \right) \right], \\
&= \Re \left( \frac{\partial \langle \phi(\vec{\theta}(t))|}{\partial \theta_i} \frac{\partial |\phi(\vec{\theta}(t))\rangle}{\partial \theta_j} + \frac{\partial \langle \phi(\vec{\theta}(t))|}{\partial \theta_i} |\phi(\vec{\theta}(t))\rangle \frac{\partial \langle \phi(\vec{\theta}(t))|}{\partial \theta_j} |\phi(\vec{\theta}(t))\rangle \right), \\
&= A_{i,j}^R + \frac{\partial \langle \phi(\vec{\theta}(t))|}{\partial \theta_i} |\phi(\vec{\theta}(t))\rangle \frac{\partial \langle \phi(\vec{\theta}(t))|}{\partial \theta_j} |\phi(\vec{\theta}(t))\rangle, \\
V_i &= -i \text{Tr} \left[ \frac{\partial \rho(\vec{\theta}(t))}{\partial \theta_i} [H, \rho(\vec{\theta}(t))] \right], \\
&= -i \text{Tr} \left[ \left( \frac{\partial |\phi(\vec{\theta}(t))\rangle}{\partial \theta_i} \langle \phi(\vec{\theta}(t))| + |\phi(\vec{\theta}(t))\rangle \frac{\partial \langle \phi(\vec{\theta}(t))|}{\partial \theta_i} \right) (H |\phi(\vec{\theta}(t))\rangle \langle \phi(\vec{\theta}(t))| - |\phi(\vec{\theta}(t))\rangle \langle \phi(\vec{\theta}(t))| H) \right], \\
&= \Im \left( \frac{\partial \langle \phi(\vec{\theta}(t))|}{\partial \theta_i} H |\phi(\vec{\theta}(t))\rangle + \langle \phi(\vec{\theta}(t))| \frac{\partial |\phi(\vec{\theta}(t))\rangle}{\partial \theta_i} \langle \phi(\vec{\theta}(t))| H |\phi(\vec{\theta}(t))\rangle \right), \\
&= C_i^I + i \frac{\partial \langle \phi(\vec{\theta}(t))|}{\partial \theta_i} |\phi(\vec{\theta}(t))\rangle \langle \phi(\vec{\theta}(t))| H |\phi(\vec{\theta}(t))\rangle
\end{aligned} \tag{106}$$

This equation is equivalent to applying a phase gate to an additional ancilla with an additional parameter in the pure state equation.

*Proof.* Suppose there are  $N$  parameters in the ansatz. Consider an additional ancilla that is rotated via a phase gate  $e^{i\theta_0}$ , the evolution of the  $N + 1$  parameters via the pure state equation is equivalent to the evolution of the  $N$  parameters via the mixed state equation. Under the pure state equation, the evolution of the  $N + 1$  parameters is

$$\begin{aligned}
&\sum_{j=1}^N A_{i,j}^R \dot{\theta}_j + i \frac{\partial \langle \phi(\vec{\theta}(t))|}{\partial \theta_i} |\phi(\vec{\theta}(t))\rangle \dot{\theta}_0 = C_i^I, \\
&-i \sum_{j=1}^N \frac{\partial \langle \phi(\vec{\theta}(t))|}{\partial \theta_i} |\phi(\vec{\theta}(t))\rangle \dot{\theta}_j - \dot{\theta}_0 = \langle \phi(\vec{\theta}(t))| H |\phi(\vec{\theta}(t))\rangle.
\end{aligned} \tag{107}$$

Substituting the last equation into the first  $N$  equations, we get

$$\sum_{j=1}^N A_{i,j}^R \dot{\theta}_j + i \frac{\partial \langle \phi(\vec{\theta}(t))|}{\partial \theta_i} |\phi(\vec{\theta}(t))\rangle \left( -i \sum_{j=1}^N \frac{\partial \langle \phi(\vec{\theta}(t))|}{\partial \theta_i} |\phi(\vec{\theta}(t))\rangle \dot{\theta}_j - \langle \phi(\vec{\theta}(t))| H |\phi(\vec{\theta}(t))\rangle \right) = C_i^I, \tag{108}$$

which it is equivalent to

$$\left( A_{i,j}^R + \frac{\partial \langle \phi(\vec{\theta}(t))|}{\partial \theta_i} |\phi(\vec{\theta}(t))\rangle \frac{\partial \langle \phi(\vec{\theta}(t))|}{\partial \theta_j} |\phi(\vec{\theta}(t))\rangle \right) \dot{\theta}_j = C_i^I + i \frac{\partial \langle \phi(\vec{\theta}(t))|}{\partial \theta_i} |\phi(\vec{\theta}(t))\rangle \langle \phi(\vec{\theta}(t))| H |\phi(\vec{\theta}(t))\rangle. \tag{109}$$

□

## B Imaginary time evolution

### B.1 Imaginary time evolution: pure states

#### B.1.1 The Dirac and Frenkel variational principle

Replace real time with imaginary time  $\tau = it$ , the Wick-rotated Schrödinger equation is,

$$\frac{d|\psi(\tau)\rangle}{d\tau} = -(H - E_\tau)|\psi(\tau)\rangle, \quad (110)$$

where  $E_\tau = \langle\psi(\tau)|H|\psi(\tau)\rangle$ . Consider a normalised parametrised trial state  $|\phi(\vec{\theta}(\tau))\rangle$ , with  $\vec{\theta}(\tau) = (\theta_1(\tau), \theta_2(\tau), \dots, \theta_N(\tau))$ , the imaginary time evolution of the Schrödinger equation on the trial state space is

$$\frac{d|\phi(\vec{\theta}(\tau))\rangle}{d\tau} = \sum_i \frac{\partial|\phi(\vec{\theta}(\tau))\rangle}{\partial\theta_i} \dot{\theta}_i = -(H - E_\tau)|\phi(\vec{\theta}(\tau))\rangle. \quad (111)$$

Here  $\dot{\theta}_i = \frac{d\theta_i}{d\tau}$ . Apply the projector

$$P = \sum_i \frac{\partial|\phi(\vec{\theta}(\tau))\rangle}{\partial\theta_i} \frac{\partial\langle\phi(\vec{\theta}(\tau))|}{\partial\theta_i}, \quad (112)$$

we have

$$\begin{aligned} P \sum_j \frac{\partial|\phi(\vec{\theta}(\tau))\rangle}{\partial\theta_j} \dot{\theta}_j &= -P(H - E_\tau)|\phi(\vec{\theta}(\tau))\rangle, \\ \sum_i \frac{\partial|\phi(\vec{\theta}(\tau))\rangle}{\partial\theta_i} \frac{\partial\langle\phi(\vec{\theta}(\tau))|}{\partial\theta_i} \sum_j \frac{\partial|\phi(\vec{\theta}(\tau))\rangle}{\partial\theta_j} \dot{\theta}_j &= -\sum_i \frac{\partial|\phi(\vec{\theta}(\tau))\rangle}{\partial\theta_i} \frac{\partial\langle\phi(\vec{\theta}(\tau))|}{\partial\theta_i} (H - E_\tau)|\phi(\vec{\theta}(\tau))\rangle. \end{aligned} \quad (113)$$

For each term of  $\frac{\partial|\phi(\vec{\theta}(\tau))\rangle}{\partial\theta_i}$ , the coefficient should satisfy

$$\sum_j \frac{\partial\langle\phi(\vec{\theta}(\tau))|}{\partial\theta_i} \frac{\partial|\phi(\vec{\theta}(\tau))\rangle}{\partial\theta_j} \dot{\theta}_j = -\frac{\partial\langle\phi(\vec{\theta}(\tau))|}{\partial\theta_i} (H - E_\tau)|\phi(\vec{\theta}(\tau))\rangle. \quad (114)$$

Define the matrix elements of  $A$  and  $C$  as

$$\begin{aligned} A_{i,j} &= \frac{\partial\langle\phi(\vec{\theta}(\tau))|}{\partial\theta_i} \frac{\partial|\phi(\vec{\theta}(\tau))\rangle}{\partial\theta_j}, \\ C'_i &= \frac{\partial\langle\phi(\vec{\theta}(\tau))|}{\partial\theta_i} (H - E_\tau)|\phi(\vec{\theta}(\tau))\rangle, \end{aligned} \quad (115)$$

the evolution of parameters is simplified to

$$\sum_j A_{i,j} \dot{\theta}_j = -C'_i. \quad (116)$$

The same equation can be obtained by applying the Dirac and Frenkel variational principle

$$\left\langle \delta\phi(\vec{\theta}(\tau)) \left| \left( \frac{d}{d\tau} + H - E_\tau \right) \right| \phi(\vec{\theta}(\tau)) \right\rangle = 0. \quad (117)$$

This can be verified with

$$\langle \delta\phi(\vec{\theta}(\tau)) | = \sum_i \frac{\partial\langle\phi(\vec{\theta}(\tau))|}{\partial\theta_i} \delta\theta_i, \quad (118)$$

and Eq. (111).



### B.1.2 McLachlan's variational principle

The McLachlan's variational principle [50], applied to imaginary time evolution, is given by

$$\delta \|(d/d\tau + H - E_\tau) |\phi(\vec{\theta}(\tau))\rangle\| = 0 \quad (119)$$

where

$$\begin{aligned} & \|(d/d\tau + H - E_\tau) |\phi(\vec{\theta}(\tau))\rangle\|^2 \\ &= \left( (d/d\tau + H - E_\tau) |\phi(\vec{\theta}(\tau))\rangle \right)^\dagger (d/d\tau + H - E_\tau) |\phi(\vec{\theta}(\tau))\rangle, \\ &= \sum_{i,j} \frac{\partial \langle \phi(\vec{\theta}(t)) |}{\partial \theta_i} \frac{\partial |\phi(\vec{\theta}(t))\rangle}{\partial \theta_j} \dot{\theta}_i^* \dot{\theta}_j + \sum_i \frac{\partial \langle \phi(\vec{\theta}(t)) |}{\partial \theta_i} (H - E_\tau) |\phi(\vec{\theta}(t))\rangle \dot{\theta}_i^* \\ & \quad \sum_i \langle \phi(\vec{\theta}(t)) | (H - E_\tau) \frac{\partial |\phi(\vec{\theta}(t))\rangle}{\partial \theta_i} \dot{\theta}_i + \langle \phi(\vec{\theta}(t)) | (H - E_\tau)^2 |\phi(\vec{\theta}(t))\rangle. \end{aligned} \quad (120)$$

Focusing on  $\|(d/d\tau + H - E_\tau) |\phi(\vec{\theta}(\tau))\rangle\|^2$ , when  $\dot{\theta}_i$  is complex, we have

$$\begin{aligned} & \delta \|(d/d\tau + H - E_\tau) |\phi(\vec{\theta}(\tau))\rangle\|^2 \\ &= \left( \sum_j \frac{\partial \langle \phi(\vec{\theta}(t)) |}{\partial \theta_i} \frac{\partial |\phi(\vec{\theta}(t))\rangle}{\partial \theta_j} \dot{\theta}_j + \frac{\partial \langle \phi(\vec{\theta}(t)) |}{\partial \theta_i} (H - E_\tau) |\phi(\vec{\theta}(t))\rangle \right) \delta \dot{\theta}_i^*, \\ & \quad + \left( \sum_j \frac{\partial \langle \phi(\vec{\theta}(t)) |}{\partial \theta_j} \frac{\partial |\phi(\vec{\theta}(t))\rangle}{\partial \theta_i} \dot{\theta}_j^* + \frac{\partial \langle \phi(\vec{\theta}(t)) |}{\partial \theta_i} (H - E_\tau) |\phi(\vec{\theta}(t))\rangle \right) \delta \dot{\theta}_i. \end{aligned} \quad (121)$$

Then the evolution is

$$\sum_j \frac{\partial \langle \phi(\vec{\theta}(\tau)) |}{\partial \theta_i} \frac{\partial |\phi(\vec{\theta}(\tau))\rangle}{\partial \theta_j} \dot{\theta}_j = - \frac{\partial \langle \phi(\vec{\theta}(\tau)) |}{\partial \theta_i} (H - E_\tau) |\phi(\vec{\theta}(\tau))\rangle. \quad (122)$$

When  $\dot{\theta}_i$  can only take real values, we have

$$\begin{aligned} & \delta \|(d/d\tau + H - E_\tau) |\phi(\vec{\theta}(\tau))\rangle\|^2 \\ &= \sum_j \left( \frac{\partial \langle \phi(\vec{\theta}(t)) |}{\partial \theta_i} \frac{\partial |\phi(\vec{\theta}(t))\rangle}{\partial \theta_j} + \frac{\partial \langle \phi(\vec{\theta}(t)) |}{\partial \theta_j} \frac{\partial |\phi(\vec{\theta}(t))\rangle}{\partial \theta_i} \right) \dot{\theta}_j \delta \theta_i \\ & \quad + \left( \frac{\partial \langle \phi(\vec{\theta}(t)) |}{\partial \theta_i} (H - E_\tau) |\phi(\vec{\theta}(t))\rangle + \langle \phi(\vec{\theta}(t)) | (H - E_\tau) \frac{\partial |\phi(\vec{\theta}(t))\rangle}{\partial \theta_i} \right) \delta \theta_i. \end{aligned} \quad (123)$$

The corresponding evolution equation for parameters is

$$\begin{aligned} & \sum_j \left( \frac{\partial \langle \phi(\vec{\theta}(t)) |}{\partial \theta_i} \frac{\partial |\phi(\vec{\theta}(t))\rangle}{\partial \theta_j} + \frac{\partial \langle \phi(\vec{\theta}(t)) |}{\partial \theta_j} \frac{\partial |\phi(\vec{\theta}(t))\rangle}{\partial \theta_i} \right) \dot{\theta}_j \\ &= - \left( \frac{\partial \langle \phi(\vec{\theta}(t)) |}{\partial \theta_i} (H - E_\tau) |\phi(\vec{\theta}(t))\rangle - \langle \phi(\vec{\theta}(t)) | (H - E_\tau) \frac{\partial |\phi(\vec{\theta}(t))\rangle}{\partial \theta_i} \right), \end{aligned} \quad (124)$$

and it is equivalent to

$$\sum_j A_{i,j}^R \dot{\theta}_j = -C_i^R, \quad (125)$$

where  $A_{i,j}^R$  and  $C_i^R$  are the real parts of  $A_{i,j}$  and  $C_i$ , respectively. Due to the normalisation requirement for  $\phi(\vec{\theta}(\tau))$ , i.e.,  $\langle \phi(\vec{\theta}(\tau)) | \phi(\vec{\theta}(\tau)) \rangle = 1$ , we have  $\Re \left( \frac{\partial \langle \phi(\vec{\theta}(t)) |}{\partial \theta_i} E_\tau | \phi(\vec{\theta}(t)) \rangle \right) = 0$ . Therefore,

$$C_i^R = C_i'^R = \Re \left( \frac{\partial \langle \phi(\vec{\theta}(t)) |}{\partial \theta_i} H | \phi(\vec{\theta}(t)) \rangle \right). \quad (126)$$

### B.1.3 Time-dependent variational principles

To recover the Wick-rotated Schrödinger equation, the corresponding Lagrangian should be modified to be

$$L = \left\langle \psi \left| \frac{d}{d\tau} + H \right| \psi \right\rangle + \langle \psi | H | \psi \rangle (1 - \langle \psi | \psi \rangle). \quad (127)$$

Take  $|\psi\rangle$ ,  $\frac{d|\psi\rangle}{d\tau}$ ,  $\langle\psi|$ , and  $\frac{d\langle\psi|}{d\tau}$  as free parameters, the Euler-Lagrange equation to  $\langle\psi|$  is

$$\begin{aligned} \frac{\partial L}{\partial \langle \psi |} - \frac{d}{d\tau} \frac{\partial L}{\partial \frac{d\langle \psi |}{d\tau}} &= \left( \frac{d}{d\tau} + H \right) |\psi\rangle + H |\psi\rangle (1 - \langle \psi | \psi \rangle) - \langle \psi | H | \psi \rangle |\psi\rangle, \\ &= \left( \frac{d}{d\tau} + H - E_\tau \right) |\psi\rangle, \end{aligned} \quad (128)$$

where the second line makes use of the normalisation condition  $\langle \psi | \psi \rangle = 1$  and the definition of  $E_\tau = \langle \psi | H | \psi \rangle$ .

Now, replace  $\psi$  with the parameterised trial state  $\phi(\vec{\theta}(\tau))$ , the Lagrangian becomes,

$$\begin{aligned} L &= \left\langle \phi(\vec{\theta}(\tau)) \left| \left( \frac{d}{d\tau} + H \right) \right| \phi(\vec{\theta}(\tau)) \right\rangle - \langle \phi(\vec{\theta}(\tau)) | H | \phi(\vec{\theta}(\tau)) \rangle (1 - \langle \phi(\vec{\theta}(\tau)) | \phi(\vec{\theta}(\tau)) \rangle), \\ &= \sum_i \dot{\theta}_i \left\langle \phi(\vec{\theta}(\tau)) \left| \frac{\partial}{\partial \theta_i} \right| \phi(\vec{\theta}(\tau)) \right\rangle - \langle \phi(\vec{\theta}(\tau)) | H | \phi(\vec{\theta}(\tau)) \rangle (2 - \langle \phi(\vec{\theta}(\tau)) | \phi(\vec{\theta}(\tau)) \rangle) \end{aligned} \quad (129)$$

Suppose  $\theta$  is complex, then

$$\begin{aligned} \frac{\partial L}{\partial \theta_i^*} - \frac{d}{dt} \frac{\partial L}{\partial \dot{\theta}_i^*} &= \sum_j \dot{\theta}_j \frac{\partial \langle \phi(\vec{\theta}(t)) |}{\partial \theta_i} \frac{\partial | \phi(\vec{\theta}(t)) \rangle}{\partial \theta_j} - \frac{\partial \langle \phi(\vec{\theta}(t)) |}{\partial \theta_i} H | \phi(\vec{\theta}(t)) \rangle (2 - \langle \phi(\vec{\theta}(\tau)) | \phi(\vec{\theta}(\tau)) \rangle) \\ &\quad + \langle \phi(\vec{\theta}(\tau)) | H | \phi(\vec{\theta}(\tau)) \rangle \frac{\partial \langle \phi(\vec{\theta}(t)) |}{\partial \theta_i} | \phi(\vec{\theta}(t)) \rangle, \end{aligned} \quad (130)$$

and the evolution is

$$\sum_j \frac{\partial \langle \phi(\vec{\theta}(\tau)) |}{\partial \theta_i} \frac{\partial | \phi(\vec{\theta}(\tau)) \rangle}{\partial \theta_j} \dot{\theta}_j = - \frac{\partial \langle \phi(\vec{\theta}(\tau)) |}{\partial \theta_i} (H - E_\tau) | \phi(\vec{\theta}(\tau)) \rangle. \quad (131)$$

When  $\theta_i$  is real, we can have the evolution of parameters with the Euler-Lagrange equation  $\frac{\partial L}{\partial \theta_i} - \frac{d}{d\tau} \frac{\partial L}{\partial \dot{\theta}_i} = 0$ . The first term  $\frac{\partial L}{\partial \theta_i}$

$$\begin{aligned} \frac{\partial L}{\partial \theta_i} &= \sum_j \dot{\theta}_j \frac{\partial}{\partial \theta_i} \left\langle \phi(\vec{\theta}(\tau)) \left| \frac{\partial}{\partial \theta_j} \right| \phi(\vec{\theta}(\tau)) \right\rangle - \frac{\partial}{\partial \theta_i} [\langle \phi(\vec{\theta}(\tau)) | H | \phi(\vec{\theta}(\tau)) \rangle (2 - \langle \phi(\vec{\theta}(\tau)) | \phi(\vec{\theta}(\tau)) \rangle)], \\ &= \sum_j \dot{\theta}_j \left( \frac{\partial \langle \phi(\vec{\theta}(\tau)) |}{\partial \theta_i} \frac{\partial | \phi(\vec{\theta}(\tau)) \rangle}{\partial \theta_j} + \langle \phi(\vec{\theta}(\tau)) | \frac{\partial^2 | \phi(\vec{\theta}(\tau)) \rangle}{\partial \theta_i \partial \theta_j} \right) \\ &\quad - \frac{\partial \langle \phi(\vec{\theta}(\tau)) |}{\partial \theta_i} H | \phi(\vec{\theta}(\tau)) \rangle - \langle \phi(\vec{\theta}(\tau)) | H \frac{\partial | \phi(\vec{\theta}(\tau)) \rangle}{\partial \theta_i} \end{aligned} \quad (132)$$

Here, we applied the normalisation condition  $\langle \phi(\vec{\theta}(\tau)) | \phi(\vec{\theta}(\tau)) \rangle = 1$  and  $\frac{\partial}{\partial \theta_i} (2 - \langle \phi(\vec{\theta}(\tau)) | \phi(\vec{\theta}(\tau)) \rangle) = 0$ . The second term  $\frac{d}{dt} \frac{\partial L}{\partial \dot{\theta}_i}$  is

$$\begin{aligned} \frac{d}{d\tau} \frac{\partial L}{\partial \dot{\theta}_i} &= \frac{d}{d\tau} \left\langle \phi(\vec{\theta}(\tau)) \left| \frac{\partial}{\partial \theta_i} \right| \phi(\vec{\theta}(\tau)) \right\rangle, \\ &= \sum_j \dot{\theta}_j \left( \frac{\partial \langle \phi(\vec{\theta}(\tau)) |}{\partial \theta_j} \frac{\partial |\phi(\vec{\theta}(\tau))\rangle}{\partial \theta_i} + \langle \phi(\vec{\theta}(\tau)) | \frac{\partial^2 |\phi(\vec{\theta}(\tau))\rangle}{\partial \theta_i \partial \theta_j} \right), \end{aligned} \quad (133)$$

and

$$\begin{aligned} \frac{\partial L}{\partial \theta_i} - \frac{d}{d\tau} \frac{\partial L}{\partial \dot{\theta}_i} &= \sum_j \dot{\theta}_j \left( \frac{\partial \langle \phi(\vec{\theta}(\tau)) |}{\partial \theta_i} \frac{\partial |\phi(\vec{\theta}(\tau))\rangle}{\partial \theta_j} - \frac{\partial \langle \phi(\vec{\theta}(\tau)) |}{\partial \theta_j} \frac{\partial |\phi(\vec{\theta}(\tau))\rangle}{\partial \theta_i} \right) \\ &\quad - \frac{\partial \langle \phi(\vec{\theta}(\tau)) |}{\partial \theta_i} H |\phi(\vec{\theta}(\tau))\rangle - \langle \phi(\vec{\theta}(\tau)) | H \frac{\partial |\phi(\vec{\theta}(\tau))\rangle}{\partial \theta_i} \end{aligned} \quad (134)$$

The evolution of parameters is

$$\begin{aligned} &\sum_j \dot{\theta}_j \left( \frac{\partial \langle \phi(\vec{\theta}(\tau)) |}{\partial \theta_i} \frac{\partial |\phi(\vec{\theta}(\tau))\rangle}{\partial \theta_j} - \frac{\partial \langle \phi(\vec{\theta}(\tau)) |}{\partial \theta_j} \frac{\partial |\phi(\vec{\theta}(\tau))\rangle}{\partial \theta_i} \right) \\ &= \frac{\partial \langle \phi(\vec{\theta}(\tau)) |}{\partial \theta_i} H |\phi(\vec{\theta}(\tau))\rangle + \langle \phi(\vec{\theta}(\tau)) | H \frac{\partial |\phi(\vec{\theta}(\tau))\rangle}{\partial \theta_i}, \end{aligned} \quad (135)$$

and it is equivalent to

$$i \sum_j A_{i,j}^I \dot{\theta}_j = C_i^R. \quad (136)$$

Here,  $A_{i,j}^I$  and  $C_i^R$  is the imaginary and real parts of  $A_{i,j}$  and  $C_i$ , respectively. Note that the equation always lead to an incorrect imaginary solution of  $\dot{\theta}_j$ .

## B.2 Imaginary time evolution of mixed state

### B.2.1 Evolution of mixed states

Under the imaginary time evolution for Hamiltonian  $H$ , the state  $\rho(\tau)$  at imaginary time  $\tau$  should be

$$\rho(\tau) = \frac{e^{-H\tau} \rho(0) e^{-H\tau}}{\text{Tr}[e^{-2H\tau} \rho(0)]}, \quad (137)$$

where  $\rho(0)$  is the initial state. It can be easily checked that  $\rho(\tau)$  follows the time derivative equation

$$\frac{d\rho(\tau)}{d\tau} = - \left( \{H, \rho(\tau)\} - 2 \langle H \rangle \rho(\tau) \right), \quad (138)$$

where  $\{H, \rho(\tau)\} = H\rho(\tau) + \rho(\tau)H$ . Consider a parametrised state, and we can obtain the evolution of  $\vec{\theta}$  that simulates Eq. (138),

$$\sum_i \frac{\partial \rho(\vec{\theta})}{\partial \theta_i} \dot{\theta}_i = - \left( \{H, \rho(\vec{\theta}(t))\} - 2 \langle H \rangle \rho(\vec{\theta}(t)) \right) \quad (139)$$

**The Dirac and Frenkel variational principle** Following the Dirac and Frenkel variational principle, we have

$$\text{Tr} \left[ \delta \rho(\vec{\theta}) \left( \sum_i \frac{\partial \rho(\vec{\theta})}{\partial \theta_i} \dot{\theta}_i + \left( \{H, \rho(\vec{\theta}(t))\} - 2 \langle H \rangle \rho(\vec{\theta}(t)) \right) \right) \right] = 0 \quad (140)$$

The evolution of parameters is

$$\sum_j M_{i,j} \dot{\theta}_j = Y_i \quad (141)$$

with

$$\begin{aligned} M_{i,j} &= \text{Tr} \left[ \left( \frac{\partial \rho(\vec{\theta})}{\partial \theta_i} \right)^\dagger \frac{\partial \rho(\vec{\theta}(t))}{\partial \theta_j} \right] \\ Y_i &= -\text{Tr} \left[ \left( \frac{\partial \rho(\vec{\theta})}{\partial \theta_i} \right)^\dagger \left( \{H, \rho(\vec{\theta}(t))\} - 2 \langle H \rangle \rho(\vec{\theta}(t)) \right) \right] \end{aligned} \quad (142)$$

Note that when the parameter  $\vec{\theta}$  is real,  $M$  and  $Y$  are also real, and

$$V'_i = -\text{Tr} \left[ \frac{\partial \rho(\vec{\theta}(t))}{\partial \theta_i} \{H, \rho(\vec{\theta}(t))\} \right]. \quad (143)$$

In this case, the solution  $\vec{\theta}$  is also real.

**MacLachlan's variational principle** We can also minimise the distance between the evolution via parameters and the evolution via Eq. (138).

$$\delta \|d\rho/d\tau + (\{H, \rho\} - 2 \langle H \rangle \rho)\| = 0 \quad (144)$$

where

$$\begin{aligned} &\delta \|d\rho/d\tau + (\{H, \rho\} - 2 \langle H \rangle \rho)\|^2 \\ &= \text{Tr} \left[ \left( d\rho/d\tau + (\{H, \rho\} - 2 \langle H \rangle \rho) \right)^\dagger \left( d\rho/d\tau + (\{H, \rho\} - 2 \langle H \rangle \rho) \right) \right] \\ &= \text{Tr} \left[ \left( \sum_i \frac{\partial \rho}{\partial \theta_i} \dot{\theta}_i \right)^\dagger \sum_j \frac{\partial \rho}{\partial \theta_j} \dot{\theta}_j + \left( \sum_i \frac{\partial \rho}{\partial \theta_i} \dot{\theta}_i \right)^\dagger (\{H, \rho\} - 2 \langle H \rangle \rho) \right] \\ &\quad + \text{Tr} \left[ (\{H, \rho\} - 2 \langle H \rangle \rho) \sum_j \frac{\partial \rho}{\partial \theta_j} \dot{\theta}_j + (\{H, \rho\} - 2 \langle H \rangle \rho)^2 \right] \end{aligned} \quad (145)$$

Focusing on  $\delta \|d\rho/d\tau + (\{H, \rho\} - 2 \langle H \rangle \rho)\|^2$ , suppose  $\vec{\theta}$  is complex, then

$$\begin{aligned} &\delta \|d\rho/d\tau + (\{H, \rho\} - 2 \langle H \rangle \rho)\|^2 \\ &= \sum_i \left( \text{Tr} \left[ \left( \frac{\partial \rho(\vec{\theta})}{\partial \theta_i} \right)^\dagger \sum_j \frac{\partial \rho(\vec{\theta})}{\partial \theta_j} \dot{\theta}_j + \left( \frac{\partial \rho(\vec{\theta})}{\partial \theta_i} \right)^\dagger (\{H, \rho(\vec{\theta}(t))\} - 2 \langle H \rangle \rho(\vec{\theta})) \right] \delta \theta_i^* \right. \\ &\quad \left. + \text{Tr} \left[ \left( \sum_j \frac{\partial \rho(\vec{\theta})}{\partial \theta_j} \dot{\theta}_j \right)^\dagger \frac{\partial \rho(\vec{\theta})}{\partial \theta_i} + (\{H, \rho(\vec{\theta}(t))\} - 2 \langle H \rangle \rho(\vec{\theta})) \frac{\partial \rho(\vec{\theta})}{\partial \theta_i} \right] \delta \theta_i \right) \end{aligned} \quad (146)$$

The evolution to  $\vec{\theta}$  is

$$\sum_j \text{Tr} \left[ \left( \frac{\partial \rho(\vec{\theta})}{\partial \theta_i} \right)^\dagger \frac{\partial \rho(\vec{\theta})}{\partial \theta_j} \right] \dot{\theta}_j = -\text{Tr} \left[ \left( \frac{\partial \rho(\vec{\theta})}{\partial \theta_i} \right)^\dagger (\{H, \rho(\vec{\theta}(t))\} - 2 \langle H \rangle \rho(\vec{\theta})) \right] \quad (147)$$

When  $\vec{\theta}$  is real, then

$$\begin{aligned} & \delta \|d\rho/d\tau + (\{H, \rho\} - 2 \langle H \rangle \rho)\|^2 \\ &= \sum_i \left( \text{Tr} \left[ \frac{\partial \rho(\vec{\theta})}{\partial \theta_i} \sum_j \frac{\partial \rho(\vec{\theta})}{\partial \theta_j} \dot{\theta}_j + \frac{\partial \rho(\vec{\theta})}{\partial \theta_i} \left( \{H, \rho(\vec{\theta}(t))\} - 2 \langle H \rangle \rho(\vec{\theta}) \right) \right] \delta \theta_i \right. \\ & \quad \left. + \text{Tr} \left[ \left( \sum_j \frac{\partial \rho(\vec{\theta})}{\partial \theta_j} \dot{\theta}_j \right) \frac{\partial \rho(\vec{\theta})}{\partial \theta_i} + \left( \{H, \rho(\vec{\theta}(t))\} - 2 \langle H \rangle \rho(\vec{\theta}) \right) \frac{\partial \rho(\vec{\theta})}{\partial \theta_i} \right] \delta \theta_i \right) \end{aligned} \quad (148)$$

The evolution to  $\vec{\theta}$  is

$$\begin{aligned} \sum_j \text{Tr} \left[ \frac{\partial \rho(\vec{\theta}(t))}{\partial \theta_i} \frac{\partial \rho(\vec{\theta}(t))}{\partial \theta_j} \right] \dot{\theta}_j &= -\text{Tr} \left[ \frac{\partial \rho(\vec{\theta}(t))}{\partial \theta_i} \left( \{H, \rho(\vec{\theta}(t))\} - 2 \langle H \rangle \rho(\vec{\theta}(t)) \right) \right], \\ &= -\text{Tr} \left[ \frac{\partial \rho(\vec{\theta}(t))}{\partial \theta_i} \{H, \rho(\vec{\theta}(t))\} \right]. \end{aligned} \quad (149)$$

Then we can obtain the same equation as derived by using the time-dependent variational principle.

**The time-dependent variational principle** The Lagrangian for Eq. (138) is

$$L = \text{Tr}[\rho^\dagger (d\rho/d\tau + (\{H, \rho\} - 2 \langle H \rangle \rho))] \quad (150)$$

Eq. (138) is obtained by

$$\frac{\partial L}{\partial \rho^\dagger} - \frac{d}{d\tau} \frac{\partial L}{\partial \dot{\rho}^\dagger} = d\rho/d\tau + (\{H, \rho\} - 2 \langle H \rangle \rho) = 0 \quad (151)$$

Suppose  $\rho = \rho(\vec{\theta}(t))$ , then the Lagrangian is

$$L = \text{Tr} \left[ \rho(\vec{\theta}(t))^\dagger \left( \left( \sum_j \frac{\partial \rho(\vec{\theta}(t))}{\partial \theta_j} \dot{\theta}_j + \left( \{H, \rho(\vec{\theta}(t))\} - 2 \langle H \rangle \rho(\vec{\theta}(t)) \right) \right) \right) \right] \quad (152)$$

Suppose  $\vec{\theta}$  is complex, then

$$\frac{\partial L}{\partial \theta_i^*} - \frac{d}{d\tau} \frac{\partial L}{\partial \dot{\theta}_i^*} = \text{Tr} \left[ \frac{\partial \rho(\vec{\theta}(t))^\dagger}{\partial \theta_i^*} \left( \sum_j \frac{\partial \rho(\vec{\theta}(t))}{\partial \theta_j} \dot{\theta}_j + \left( \{H, \rho(\vec{\theta}(t))\} - 2 \langle H \rangle \rho(\vec{\theta}(t)) \right) \right) \right], \quad (153)$$

and the evolution is

$$\sum_j \text{Tr} \left[ \left( \frac{\partial \rho(\vec{\theta}(t))}{\partial \theta_i} \right)^\dagger \frac{\partial \rho(\vec{\theta}(t))}{\partial \theta_j} \right] \dot{\theta}_j = -\text{Tr} \left[ \left( \frac{\partial \rho(\vec{\theta}(t))}{\partial \theta_i} \right)^\dagger \left( \{H, \rho(\vec{\theta}(t))\} - 2 \langle H \rangle \rho(\vec{\theta}(t)) \right) \right] \quad (154)$$

Instead, suppose  $\vec{\theta}$  is real, we cannot reproduce the imaginary time evolution as we have  $\rho = \rho^\dagger$ . Therefore, we cannot the equation for evaluating  $\vec{\theta}$ .

### B.2.2 Reduction to the pure state case

Suppose  $\rho(\vec{\theta}(t))$  is a pure state  $|\phi(\vec{\theta})\rangle\langle\phi(\vec{\theta})|$ , and Eq. (138) should become the Wicked-rotated Schrödinger equation. For the variational evolution equation, we have

$$\begin{aligned}
M_{i,j} &= \text{Tr} \left[ \frac{\partial \rho(\vec{\theta}(t))}{\partial \theta_i} \frac{\partial \rho(\vec{\theta}(t))}{\partial \theta_j} \right], \\
&= \text{Tr} \left[ \left( \frac{\partial |\phi(\vec{\theta}(\tau))\rangle}{\partial \theta_i} \langle\phi(\vec{\theta}(\tau))| + |\phi(\vec{\theta}(\tau))\rangle \frac{\partial \langle\phi(\vec{\theta}(\tau))|}{\partial \theta_i} \right) \left( \frac{\partial |\phi(\vec{\theta}(\tau))\rangle}{\partial \theta_j} \langle\phi(\vec{\theta}(\tau))| + |\phi(\vec{\theta}(\tau))\rangle \frac{\partial \langle\phi(\vec{\theta}(\tau))|}{\partial \theta_j} \right) \right], \\
&= \Re \left( \frac{\partial \langle\phi(\vec{\theta}(\tau))|}{\partial \theta_i} \frac{\partial |\phi(\vec{\theta}(\tau))\rangle}{\partial \theta_j} + \frac{\partial \langle\phi(\vec{\theta}(\tau))|}{\partial \theta_i} |\phi(\vec{\theta}(\tau))\rangle \frac{\partial \langle\phi(\vec{\theta}(\tau))|}{\partial \theta_j} |\phi(\vec{\theta}(\tau))\rangle \right), \\
&= A_{i,j}^R + \frac{\partial \langle\phi(\vec{\theta}(\tau))|}{\partial \theta_i} |\phi(\vec{\theta}(\tau))\rangle \frac{\partial \langle\phi(\vec{\theta}(\tau))|}{\partial \theta_j} |\phi(\vec{\theta}(\tau))\rangle, \\
V_i' &= -\text{Tr} \left[ \frac{\partial \rho(\vec{\theta}(t))}{\partial \theta_i} \left( \{H, \rho(\vec{\theta}(t))\} \right) \right], \\
&= -\text{Tr} \left[ \left( \frac{\partial |\phi(\vec{\theta}(\tau))\rangle}{\partial \theta_i} \langle\phi(\vec{\theta}(\tau))| + |\phi(\vec{\theta}(\tau))\rangle \frac{\partial \langle\phi(\vec{\theta}(\tau))|}{\partial \theta_i} \right) \left( H |\phi(\vec{\theta}(\tau))\rangle \langle\phi(\vec{\theta}(\tau))| + |\phi(\vec{\theta}(\tau))\rangle \langle\phi(\vec{\theta}(\tau))| H \right) \right], \\
&= -C_i^R
\end{aligned} \tag{155}$$

This equation is equivalent to applying a phase gate to an additional ancilla with an additional parameter in the pure state equation.

*Proof.* Suppose there are  $N$  parameters in the ansatz. Consider an additional ancilla that is rotated via a phase gate  $e^{i\theta_0}$ , the evolution of the  $N + 1$  parameters via the pure state equation is equivalent to the evolution of the  $N$  parameters via the mixed state equation. Under the pure state equation, the evolution of the  $N + 1$  parameters is

$$\begin{aligned}
\sum_{j=1}^N A_{i,j}^R \dot{\theta}_j + i \frac{\partial \langle\phi(\vec{\theta}(\tau))|}{\partial \theta_i} |\phi(\vec{\theta}(\tau))\rangle \dot{\theta}_0 &= -C_i^R, \\
-i \sum_{j=1}^N \frac{\partial \langle\phi(\vec{\theta}(\tau))|}{\partial \theta_i} |\phi(\vec{\theta}(\tau))\rangle \dot{\theta}_j - \dot{\theta}_0 &= 0.
\end{aligned} \tag{156}$$

Substituting the last equation into the first  $N$  equations, we get

$$\sum_{j=1}^N A_{i,j}^R \dot{\theta}_j + i \frac{\partial \langle\phi(\vec{\theta}(\tau))|}{\partial \theta_i} |\phi(\vec{\theta}(\tau))\rangle \left( -i \sum_{j=1}^N \frac{\partial \langle\phi(\vec{\theta}(\tau))|}{\partial \theta_i} |\phi(\vec{\theta}(\tau))\rangle \dot{\theta}_j \right) = -C_i^R, \tag{157}$$

which it is equivalent to

$$\left( A_{i,j}^R + \frac{\partial \langle\phi(\vec{\theta}(\tau))|}{\partial \theta_i} |\phi(\vec{\theta}(\tau))\rangle \frac{\partial \langle\phi(\vec{\theta}(\tau))|}{\partial \theta_j} |\phi(\vec{\theta}(\tau))\rangle \right) \dot{\theta}_j = -C_i^R. \tag{158}$$

□

## References

- [1] Roger Balian and Marcel Veneroni. Static and dynamic variational principles for expectation values of observables. *Annals of Physics*, 187(1):29 – 78, 1988. ISSN 0003-4916. DOI: [https://doi.org/10.1016/0003-4916\(88\)90280-1](https://doi.org/10.1016/0003-4916(88)90280-1). URL <http://www.sciencedirect.com/science/article/pii/0003491688902801>.
- [2] Víctor M. Pérez-García, Humberto Michinel, J. I. Cirac, M. Lewenstein, and P. Zoller. Dynamics of bose-einstein condensates: Variational solutions of the gross-pitaevskii equations. *Phys. Rev. A*, 56:1424–1432, Aug 1997. DOI: <https://doi.org/10.1103/PhysRevA.56.1424>. URL <https://link.aps.org/doi/10.1103/PhysRevA.56.1424>.
- [3] Franco Dalfovo, Stefano Giorgini, Lev P. Pitaevskii, and Sandro Stringari. Theory of bose-einstein condensation in trapped gases. *Rev. Mod. Phys.*, 71:463–512, Apr 1999. DOI: <https://doi.org/10.1103/RevModPhys.71.463>. URL <https://link.aps.org/doi/10.1103/RevModPhys.71.463>.
- [4] Jutho Haegeman, J. Ignacio Cirac, Tobias J. Osborne, Iztok Pizorn, Henri Verschelde, and Frank Verstraete. Time-dependent variational principle for quantum lattices. *Phys. Rev. Lett.*, 107:070601, Aug 2011. DOI: <https://doi.org/10.1103/PhysRevLett.107.070601>. URL <https://link.aps.org/doi/10.1103/PhysRevLett.107.070601>.
- [5] F. Verstraete, J. J. García-Ripoll, and J. I. Cirac. Matrix product density operators: Simulation of finite-temperature and dissipative systems. *Phys. Rev. Lett.*, 93:207204, Nov 2004. DOI: <https://doi.org/10.1103/PhysRevLett.93.207204>. URL <https://link.aps.org/doi/10.1103/PhysRevLett.93.207204>.
- [6] Tao Shi, Eugene Demler, and J. Ignacio Cirac. Variational study of fermionic and bosonic systems with non-gaussian states: Theory and applications. *Annals of Physics*, 390:245 – 302, 2018. ISSN 0003-4916. DOI: <https://doi.org/10.1016/j.aop.2017.11.014>. URL <http://www.sciencedirect.com/science/article/pii/S0003491617303251>.
- [7] Laurens Vanderstraeten, Jutho Haegeman, and Frank Verstraete. Tangent-space methods for uniform matrix product states. *SciPost Phys. Lect. Notes*, page 7, 2019. DOI: <https://doi.org/10.21468/SciPostPhysLectNotes.7>. URL <https://scipost.org/10.21468/SciPostPhysLectNotes.7>.
- [8] Hans Feldmeier and Jürgen Schnack. Molecular dynamics for fermions. *Rev. Mod. Phys.*, 72:655–688, Jul 2000. DOI: <https://doi.org/10.1103/RevModPhys.72.655>. URL <https://link.aps.org/doi/10.1103/RevModPhys.72.655>.
- [9] A. Szabo and N.S. Ostlund. *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*. Dover Books on Chemistry. Dover Publications, 2012. ISBN 9780486134598. URL <https://books.google.co.uk/books?id=KQ3DAgAAQBAJ>.
- [10] T. Helgaker, P. Jorgensen, and J. Olsen. *Molecular Electronic-Structure Theory*. Wiley, 2013. ISBN 9781118531471. DOI: <https://doi.org/10.1002/9781119019572>. URL <https://books.google.co.uk/books?id=APjLWFFxWkQC>.
- [11] F. Verstraete, V. Murg, and J. I. Cirac. Matrix product states, projected entangled pair states, and variational renormalization group methods for quantum spin systems. *Advances in Physics*, 57(2):143–224, 03 2008. DOI: <https://doi.org/10.1080/14789940801912366>. URL <https://doi.org/10.1080/14789940801912366>.
- [12] Yuto Ashida, Tao Shi, Mari Carmen Bañuls, J. Ignacio Cirac, and Eugene Demler. Variational principle for quantum impurity systems in and out of equilibrium: Application to kondo problems. *Phys. Rev. B*, 98:024103, Jul 2018. DOI: <https://doi.org/10.1103/PhysRevB.98.024103>. URL <https://link.aps.org/doi/10.1103/PhysRevB.98.024103>.
- [13] R. Jackiw and A. Kerman. Time-dependent variational principle and the effective action.



- Physics Letters A*, 71(2):158 – 162, 1979. ISSN 0375-9601. DOI: [https://doi.org/10.1016/0375-9601\(79\)90151-8](https://doi.org/10.1016/0375-9601(79)90151-8). URL <http://www.sciencedirect.com/science/article/pii/S0375960179901518>.
- [14] L. Lehtovaara, J. Toivanen, and J. Eloranta. Solution of time-independent schrödinger equation by the imaginary time propagation method. *Journal of Computational Physics*, 221(1):148 – 157, 2007. ISSN 0021-9991. DOI: <https://doi.org/10.1016/j.jcp.2006.06.006>. URL <http://www.sciencedirect.com/science/article/pii/S0021999106002798>.
  - [15] P Kramer. A review of the time-dependent variational principle. *Journal of Physics: Conference Series*, 99:012009, feb 2008. DOI: <https://doi.org/10.1088/1742-6596/99/1/012009>. URL <https://doi.org/10.1088%2F1742-6596%2F99%2F1%2F012009>.
  - [16] Christina V Kraus and J Ignacio Cirac. Generalized hartree–fock theory for interacting fermions in lattices: numerical methods. *New Journal of Physics*, 12(11):113004, nov 2010. DOI: <https://doi.org/10.1088/1367-2630/12/11/113004>. URL <https://doi.org/10.1088%2F1367-2630%2F12%2F11%2F113004>.
  - [17] Aram W. Harrow and Ashley Montanaro. Quantum computational supremacy. *Nature*, 549:203 EP –, 09 2017. URL <https://doi.org/10.1038/nature23458>.
  - [18] Sergio Boixo, Sergei V. Isakov, Vadim N. Smelyanskiy, Ryan Babbush, Nan Ding, Zhang Jiang, Michael J. Bremner, John M. Martinis, and Hartmut Neven. Characterizing quantum supremacy in near-term devices. *Nature Physics*, 14(6):595–600, 2018. DOI: <https://doi.org/10.1038/s41567-018-0124-x>. URL <https://doi.org/10.1038/s41567-018-0124-x>.
  - [19] C. Neill, P. Roushan, K. Kechedzhi, S. Boixo, S. V. Isakov, V. Smelyanskiy, A. Megrant, B. Chiaro, A. Dunsworth, K. Arya, R. Barends, B. Burkett, Y. Chen, Z. Chen, A. Fowler, B. Foxen, M. Giustina, R. Graff, E. Jeffrey, T. Huang, J. Kelly, P. Klimov, E. Lucero, J. Mutus, M. Neeley, C. Quintana, D. Sank, A. Vainsencher, J. Wenner, T. C. White, H. Neven, and J. M. Martinis. A blueprint for demonstrating quantum supremacy with superconducting qubits. *Science*, 360(6385):195–199, 2018. ISSN 0036-8075. DOI: <https://doi.org/10.1126/science.aao4309>. URL <https://science.sciencemag.org/content/360/6385/195>.
  - [20] Richard P. Feynman. Simulating physics with computers. *International Journal of Theoretical Physics*, 21(6):467–488, Jun 1982. ISSN 1572-9575. DOI: <https://doi.org/10.1007/BF02650179>. URL <https://doi.org/10.1007/BF02650179>.
  - [21] Seth Lloyd. Universal quantum simulators. *Science*, 273(5278):1073–1078, 1996. ISSN 0036-8075. DOI: <https://doi.org/10.1126/science.273.5278.1073>. URL <http://science.sciencemag.org/content/273/5278/1073>.
  - [22] Daniel S. Abrams and Seth Lloyd. Simulation of many-body fermi systems on a universal quantum computer. *Phys. Rev. Lett.*, 79:2586–2589, Sep 1997. DOI: <https://doi.org/10.1103/PhysRevLett.79.2586>. URL <https://link.aps.org/doi/10.1103/PhysRevLett.79.2586>.
  - [23] Joe O’Gorman and Earl T. Campbell. Quantum computation with realistic magic-state factories. *Phys. Rev. A*, 95:032338, Mar 2017. DOI: <https://doi.org/10.1103/PhysRevA.95.032338>. URL <https://link.aps.org/doi/10.1103/PhysRevA.95.032338>.
  - [24] Earl T. Campbell, Barbara M. Terhal, and Christophe Vuillot. Roads towards fault-tolerant universal quantum computation. *Nature*, 549:172 EP –, 09 2017. URL <https://doi.org/10.1038/nature23460>.
  - [25] Markus Reiher, Nathan Wiebe, Krysta M. Svore, Dave Wecker, and Matthias Troyer. Elucidating reaction mechanisms on quantum computers. *Proceedings of the National Academy of Sciences*, 2017. ISSN 0027-8424. DOI: <https://doi.org/10.1073/pnas.1619152114>. URL <https://www.pnas.org/content/early/2017/06/30/1619152114>.

- [26] James Wooten. Benchmarking of quantum processors with random circuits. *arXiv preprint 1806.02736*, 2018.
- [27] John Preskill. Quantum Computing in the NISQ era and beyond. *Quantum*, 2:79, August 2018. ISSN 2521-327X. DOI: <https://doi.org/10.22331/q-2018-08-06-79>. URL <https://doi.org/10.22331/q-2018-08-06-79>.
- [28] Edward Farhi, Jeffrey Goldstone, and Sam Gutmann. A quantum approximate optimization algorithm. *arXiv preprint arXiv:1411.4028*, 2014.
- [29] Alberto Peruzzo, Jarrod McClean, Peter Shadbolt, Man-Hong Yung, Xiao-Qi Zhou, Peter J. Love, Alán Aspuru-Guzik, and Jeremy L. O’Brien. A variational eigenvalue solver on a photonic quantum processor. *Nature Communications*, 5:4213, 07 2014. URL <https://doi.org/10.1038/ncomms5213>.
- [30] Ya Wang, Florian Dolde, Jacob Biamonte, Ryan Babbush, Ville Bergholm, Sen Yang, Ingmar Jakobi, Philipp Neumann, Alán Aspuru-Guzik, James D. Whitfield, and Jörg Wrachtrup. Quantum simulation of helium hydride cation in a solid-state spin register. *ACS Nano*, 9(8): 7769–7774, 08 2015. DOI: <https://doi.org/10.1021/acsnano.5b01651>. URL <https://doi.org/10.1021/acsnano.5b01651>.
- [31] 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, and J. M. Martinis. Scalable quantum simulation of molecular energies. *Phys. Rev. X*, 6:031007, Jul 2016. DOI: <https://doi.org/10.1103/PhysRevX.6.031007>. URL <https://link.aps.org/doi/10.1103/PhysRevX.6.031007>.
- [32] Yangchao Shen, Xiang Zhang, Shuaining Zhang, Jing-Ning Zhang, Man-Hong Yung, and Kihwan Kim. Quantum implementation of the unitary coupled cluster for simulating molecular electronic structure. *Phys. Rev. A*, 95:020501, Feb 2017. DOI: <https://doi.org/10.1103/PhysRevA.95.020501>. URL <https://link.aps.org/doi/10.1103/PhysRevA.95.020501>.
- [33] Jarrod R McClean, Jonathan Romero, Ryan Babbush, and Alán Aspuru-Guzik. The theory of variational hybrid quantum-classical algorithms. *New Journal of Physics*, 18(2):023023, feb 2016. DOI: <https://doi.org/10.1088/1367-2630/18/2/023023>. URL <https://doi.org/10.1088/1367-2630/18/2/023023>.
- [34] S. Paesani, A. A. Gentile, R. Santagati, J. Wang, N. Wiebe, D. P. Tew, J. L. O’Brien, and M. G. Thompson. Experimental bayesian quantum phase estimation on a silicon photonic chip. *Phys. Rev. Lett.*, 118:100503, Mar 2017. DOI: <https://doi.org/10.1103/PhysRevLett.118.100503>. URL <https://link.aps.org/doi/10.1103/PhysRevLett.118.100503>.
- [35] J. I. Colless, V. V. Ramasesh, D. Dahlen, M. S. Blok, M. E. Kimchi-Schwartz, J. R. McClean, J. Carter, W. A. de Jong, and I. Siddiqi. Computation of molecular spectra on a quantum processor with an error-resilient algorithm. *Phys. Rev. X*, 8:011021, Feb 2018. DOI: <https://doi.org/10.1103/PhysRevX.8.011021>. URL <https://link.aps.org/doi/10.1103/PhysRevX.8.011021>.
- [36] Raffaele Santagati, Jianwei Wang, Antonio A. Gentile, Stefano Paesani, Nathan Wiebe, Jarrod R. McClean, Sam Morley-Short, Peter J. Shadbolt, Damien Bonneau, Joshua W. Silverstone, David P. Tew, Xiaoqi Zhou, Jeremy L. O’Brien, and Mark G. Thompson. Witnessing eigenstates for quantum simulation of hamiltonian spectra. *Science Advances*, 4(1), 2018. DOI: <https://doi.org/10.1126/sciadv.aap9646>. URL <http://advances.sciencemag.org/content/4/1/eaap9646>.
- [37] Abhinav Kandala, Antonio Mezzacapo, Kristan Temme, Maika Takita, Markus Brink, Jerry M.

- Chow, and Jay M. Gambetta. Hardware-efficient variational quantum eigensolver for small molecules and quantum magnets. *Nature*, 549:242 EP –, 09 2017. URL <https://doi.org/10.1038/nature23879>.
- [38] Abhinav Kandala, Kristan Temme, Antonio D. Córcoles, Antonio Mezzacapo, Jerry M. Chow, and Jay M. Gambetta. Error mitigation extends the computational reach of a noisy quantum processor. *Nature*, 567(7749):491–495, 2019. DOI: <https://doi.org/10.1038/s41586-019-1040-7>. URL <https://doi.org/10.1038/s41586-019-1040-7>.
- [39] Cornelius Hempel, Christine Maier, Jonathan Romero, Jarrod McClean, Thomas Monz, Heng Shen, Petar Jurcevic, Ben P. Lanyon, Peter Love, Ryan Babbush, Alán Aspuru-Guzik, Rainer Blatt, and Christian F. Roos. Quantum chemistry calculations on a trapped-ion quantum simulator. *Phys. Rev. X*, 8:031022, Jul 2018. DOI: <https://doi.org/10.1103/PhysRevX.8.031022>. URL <https://link.aps.org/doi/10.1103/PhysRevX.8.031022>.
- [40] C. Kokail, C. Maier, R. van Bijnen, T. Brydges, M. K. Joshi, P. Jurcevic, C. A. Muschik, P. Silvi, R. Blatt, C. F. Roos, and P. Zoller. Self-verifying variational quantum simulation of lattice models. *Nature*, 569(7756):355–360, May 2019. DOI: <https://doi.org/10.1038/s41586-019-1177-4>.
- [41] Ying Li and Simon C. Benjamin. Efficient variational quantum simulator incorporating active error minimization. *Phys. Rev. X*, 7:021050, Jun 2017. DOI: <https://doi.org/10.1103/PhysRevX.7.021050>. URL <https://link.aps.org/doi/10.1103/PhysRevX.7.021050>.
- [42] Ken M Nakanishi, Kosuke Mitarai, and Keisuke Fujii. Subspace-search variational quantum eigensolver for excited states. *arXiv preprint arXiv:1810.09434*, 2018.
- [43] David Poulin, Angie Qarry, Rolando Somma, and Frank Verstraete. Quantum simulation of time-dependent hamiltonians and the convenient illusion of hilbert space. *Phys. Rev. Lett.*, 106:170501, Apr 2011. DOI: <https://doi.org/10.1103/PhysRevLett.106.170501>. URL <https://link.aps.org/doi/10.1103/PhysRevLett.106.170501>.
- [44] I. M. Georgescu, S. Ashhab, and Franco Nori. Quantum simulation. *Rev. Mod. Phys.*, 86:153–185, Mar 2014. DOI: <https://doi.org/10.1103/RevModPhys.86.153>. URL <https://link.aps.org/doi/10.1103/RevModPhys.86.153>.
- [45] S. Kais, K.B. Whaley, A.R. Dinner, and S.A. Rice. *Quantum Information and Computation for Chemistry*. Advances in Chemical Physics. Wiley, 2014. ISBN 9781118742600. DOI: <https://doi.org/10.1002/9781118742631>. URL <https://books.google.co.uk/books?id=dCXPagAAQBAJ>.
- [46] Sam McArdle, Suguru Endo, Alan Aspuru-Guzik, Simon Benjamin, and Xiao Yuan. Quantum computational chemistry. *arXiv e-prints*, art. arXiv:1808.10402, Aug 2018.
- [47] Yudong Cao, Jonathan Romero, Jonathan P. Olson, Matthias Degroote, Peter D. Johnson, Mária Kieferová, Ian D. Kivlichan, Tim Menke, Borja Peropadre, Nicolas P. D. Sawaya, Sukin Sim, Libor Veis, and Alán Aspuru-Guzik. Quantum chemistry in the age of quantum computing. *Chemical Reviews*, 08 2019. DOI: <https://doi.org/10.1021/acs.chemrev.8b00803>. URL <https://doi.org/10.1021/acs.chemrev.8b00803>.
- [48] P. A. M. Dirac. Note on exchange phenomena in the thomas atom. In *Mathematical Proceedings of the Cambridge Philosophical Society*, volume 26, pages 376–385. Cambridge University Press, 1930. ISBN 0305-0041. DOI: <https://doi.org/10.1017/S0305004100016108>. URL <https://www.cambridge.org/core/article/note-on-exchange-phenomena-in-the-thomas-atom/6C5FF7297CD96F49A8B8E9E3EA50E412>.
- [49] J. Frenkel. *Wave mechanics: advanced general theory*. Clarendon Press Oxford, 1934.
- [50] A.D. McLachlan. A variational solution of the time-dependent schrodinger equation. *Molecular Physics*, 8(1):39–44, 1964. DOI: <https://doi.org/10.1080/00268976400100041>.

- [51] PH Kramer and Marcos Saraceno. *Geometry of the time-dependent variational principle in quantum mechanics*. Springer, 1981. DOI: <https://doi.org/10.1007/3-540-10579-4>.
- [52] J. Broeckhove, L. Lathouwers, E. Kesteloot, and P. Van Leuven. On the equivalence of time-dependent variational principles. *Chemical Physics Letters*, 149(5):547 – 550, 1988. ISSN 0009-2614. DOI: [https://doi.org/10.1016/0009-2614\(88\)80380-4](https://doi.org/10.1016/0009-2614(88)80380-4). URL <http://www.sciencedirect.com/science/article/pii/0009261488803804>.
- [53] Jutho Haegeman, Tobias J. Osborne, and Frank Verstraete. Post-matrix product state methods: To tangent space and beyond. *Phys. Rev. B*, 88:075133, Aug 2013. DOI: <https://doi.org/10.1103/PhysRevB.88.075133>. URL <https://link.aps.org/doi/10.1103/PhysRevB.88.075133>.
- [54] Kentaro Heya, Ken M Nakanishi, Kosuke Mitarai, and Keisuke Fujii. Subspace variational quantum simulator. *arXiv preprint arXiv:1904.08566*, 2019.
- [55] Sam McArdle, Tyson Jones, Suguru Endo, Ying Li, Simon C. Benjamin, and Xiao Yuan. Variational ansatz-based quantum simulation of imaginary time evolution. *npj Quantum Information*, 5(1):75, 2019. DOI: <https://doi.org/10.1038/s41534-019-0187-2>. URL <https://doi.org/10.1038/s41534-019-0187-2>.
- [56] Tyson Jones, Suguru Endo, Sam McArdle, Xiao Yuan, and Simon C. Benjamin. Variational quantum algorithms for discovering hamiltonian spectra. *Phys. Rev. A*, 99:062304, Jun 2019. DOI: <https://doi.org/10.1103/PhysRevA.99.062304>. URL <https://link.aps.org/doi/10.1103/PhysRevA.99.062304>.
- [57] Ming-Cheng Chen, Ming Gong, Xiao-Si Xu, Xiao Yuan, Jian-Wen Wang, Can Wang, Chong Ying, Jin Lin, Yu Xu, Yulin Wu, Shiyu Wang, Hui Deng, Futian Liang, Cheng-Zhi Peng, Simon C. Benjamin, Xiaobo Zhu, Chao-Yang Lu, and Jian-Wei Pan. Demonstration of Adiabatic Variational Quantum Computing with a Superconducting Quantum Coprocessor. *arXiv e-prints*, art. arXiv:1905.03150, May 2019.
- [58] Kosuke Mitarai and Keisuke Fujii. Methodology for replacing indirect measurements with direct measurements. *Phys. Rev. Research*, 1:013006, Aug 2019. DOI: <https://doi.org/10.1103/PhysRevResearch.1.013006>. URL <https://link.aps.org/doi/10.1103/PhysRevResearch.1.013006>.
- [59] Artur K. Ekert, Carolina Moura Alves, Daniel K. L. Oi, Michał Horodecki, Paweł Horodecki, and L. C. Kwak. Direct estimations of linear and nonlinear functionals of a quantum state. *Phys. Rev. Lett.*, 88:217901, May 2002. DOI: <https://doi.org/10.1103/PhysRevLett.88.217901>. URL <https://link.aps.org/doi/10.1103/PhysRevLett.88.217901>.
- [60] Suguru Endo, Ying Li, Simon Benjamin, and Xiao Yuan. Variational quantum simulation of general processes. *arXiv preprint arXiv:1812.08778*, 2018.
- [61] J.R. Johansson, P.D. Nation, and Franco Nori. Qutip: An open-source python framework for the dynamics of open quantum systems. *Computer Physics Communications*, 183(8):1760 – 1772, 2012. ISSN 0010-4655. DOI: <https://doi.org/10.1016/j.cpc.2012.02.021>. URL <http://www.sciencedirect.com/science/article/pii/S0010465512000835>.
- [62] J.R. Johansson, P.D. Nation, and Franco Nori. Qutip 2: A python framework for the dynamics of open quantum systems. *Computer Physics Communications*, 184(4):1234 – 1240, 2013. ISSN 0010-4655. DOI: <https://doi.org/10.1016/j.cpc.2012.11.019>. URL <http://www.sciencedirect.com/science/article/pii/S0010465512003955>.
- [63] Jarrod R. McClean, Mollie E. Kimchi-Schwartz, Jonathan Carter, and Wibe A. de Jong. Hybrid quantum-classical hierarchy for mitigation of decoherence and determination of excited states. *Phys. Rev. A*, 95:042308, Apr 2017. DOI: <https://doi.org/10.1103/PhysRevA.95.042308>. URL <https://link.aps.org/doi/10.1103/PhysRevA.95.042308>.

- [64] Kristan Temme, Sergey Bravyi, and Jay M. Gambetta. Error mitigation for short-depth quantum circuits. *Phys. Rev. Lett.*, 119:180509, Nov 2017. DOI: <https://doi.org/10.1103/PhysRevLett.119.180509>. URL <https://link.aps.org/doi/10.1103/PhysRevLett.119.180509>.
- [65] Suguru Endo, Simon C. Benjamin, and Ying Li. Practical quantum error mitigation for near-future applications. *Phys. Rev. X*, 8:031027, Jul 2018. DOI: <https://doi.org/10.1103/PhysRevX.8.031027>. URL <https://link.aps.org/doi/10.1103/PhysRevX.8.031027>.
- [66] J. I. Colless, V. V. Ramasesh, D. Dahlen, M. S. Blok, M. E. Kimchi-Schwartz, J. R. McClean, J. Carter, W. A. de Jong, and I. Siddiqi. Computation of molecular spectra on a quantum processor with an error-resilient algorithm. *Phys. Rev. X*, 8:011021, Feb 2018. DOI: <https://doi.org/10.1103/PhysRevX.8.011021>. URL <https://link.aps.org/doi/10.1103/PhysRevX.8.011021>.
- [67] Matthew Otten and Stephen K. Gray. Recovering noise-free quantum observables. *Phys. Rev. A*, 99:012338, Jan 2019. DOI: <https://doi.org/10.1103/PhysRevA.99.012338>. URL <https://link.aps.org/doi/10.1103/PhysRevA.99.012338>.
- [68] Sam McArdle, Xiao Yuan, and Simon Benjamin. Error-mitigated digital quantum simulation. *Phys. Rev. Lett.*, 122:180501, May 2019. DOI: <https://doi.org/10.1103/PhysRevLett.122.180501>. URL <https://link.aps.org/doi/10.1103/PhysRevLett.122.180501>.
- [69] X. Bonet-Monroig, R. Sagastizabal, M. Singh, and T. E. O’Brien. Low-cost error mitigation by symmetry verification. *Phys. Rev. A*, 98:062339, Dec 2018. DOI: <https://doi.org/10.1103/PhysRevA.98.062339>. URL <https://link.aps.org/doi/10.1103/PhysRevA.98.062339>.
- [70] Jarrod R McClean, Zhang Jiang, Nicholas C Rubin, Ryan Babbush, and Hartmut Neven. Decoding quantum errors with subspace expansions. *arXiv preprint arXiv:1903.05786*, 2019.