國立臺灣大學理學院物理學研究所

碩士論文

Graduate Institute of Physics
College of Science
National Taiwan University

Master Thesis

量子邏輯閘 (CNOT) 之時間最佳化研究

Optimal Control of CNOT gate

蔡東邦

Dung-Bang Tsai

指導教授：管希聖 博士

Advisor: Hsi-Sheng Goan, Ph.D.

中華民國九十七年七月

July, 2008

國立臺灣大學碩士學位論文
口試委員會審定書

量子邏輯閘 (CNOT gate) 之時間最佳化研究
Optimal control of CNOT gate

本論文係蔡東邦君（R95222004）在國立臺灣大學物理學系、所完成之碩士學位論文，於民國 97 年 6 月 20 日承下列考試委員審查通過及口試及格，特此證明

口試委員：

管希聖 （簽名）

（指導教授）

謝　辭

猶記得兩年前炎炎夏日，我來到了台大物理所修習碩士學位。雖然這兩年看似不長，只有七百多個日子，但是當中遇到不少貴人並從中學習到不少知識，讓我獲益不淺。

首先感謝黃詩姍同學，在我本來決定要直升原來母校時，鼓勵我換個環境，看看外面的世界。如果沒有她，或許這篇論文就不會出現。在論文快完成時，感謝林平同學，幫我仔細修正英文上的錯誤，並給我在句子和文章結構上很大的建議。若沒有這些勞心勞力的逐字檢查錯誤，大概論文會貽笑大方，只有自己看得懂吧。

感謝所有同研究團隊的同儕，在遇到瓶頸的時候互相打氣。和他們討論，讓我很多盲點豁然開朗。尤其是陳柏文、黃上瑜學長、饒海威和柯百謙同學，和他們的討論獲益良多。在生活上，饒海威、柯百謙、王浚帆同學和我常常在課後之餘，出去外面踏踏青，調劑研究的壓力，這讓研究做得更加順利。

感謝我的指導老師管希聖博士，在百忙的學術研究中，花了如此多的心力來指導我們。包含每週一次的一對一討論，可以很仔細的討論研究方向，遇到瓶頸時，也可以很深入的討論解決。這樣的指導方式，給我很大的幫助。在最後論文定稿時，再次又再次的檢查論文內容，力求內容的正確與通暢，充分展現出學者的風範。

最後感謝我的家人，在過去二十五年來的關心與栽培，成就了我的學識與人品。並在這段學習的過程中，讓我過著衣食無缺的生活。

謹將此論文獻給我摯愛的家人。

摘　要

我們探討在矽半導體上參雜磷原子架構下的量子電腦內，如何使用磷的電子自旋來當量子位元並做邏輯計算。我們採用脈衝序列控制來實作 CNOT 量子邏輯閘，並透過 gradient ascent pulse engineering (GRAPE) 演算法來尋找量子邏輯閘的時間最佳化解。

首先我們使用 reduced Hamiltonian 來尋找脈衝式控制序列之時間最佳解，其可控制的變數為電子與核子自旋的交互作用力(hyperfine A interactions) 和兩相鄰電子之間的交互作用力(exchange J interactions). 我們嘗試不同的控制步數，並數值計算其錯誤率是否低於我們的要求。我們發現，在時間為$100ns$, 控制步數為 30 個分段時，其錯誤率大約為1.11×10^{-16}. 接下來我們用完整的 Hamiltonian 來模擬先前找到的時間最佳化之操作序列，我們發現錯誤率約為10^{-6}, 小於量子計算中容許的錯誤率閥值(10^{-4}).

使用 GRAPE 實作此 CNOT 邏輯閘的運作時間為$100ns$. 在同樣的硬體上使用 globally controlled electron spin scheme 來實作 CNOT 邏輯閘需要$297ns$. 使用 GRAPE 比先前的方法約快了三倍。在我們的構想中還有一大優勢，我們不需要很強的電子和電子交互作用力。在我們的理論計算中最大的交互作用能量只需要$J/h = 20$MHz,傳統上該作用力需要10.2GHz. 先前若要達到如此強的作用力，兩個量子位元約需間隔10nm, 但因為在我們的架構上不需要如此強的交互作用力，我們可以把兩個量子位元的距離拉開到30nm, 如此有可能可以解決當前製造技術上的難題。

Abstract

We investigate how pulse-sequences and operation times of elementary quantum gates can be optimized for silicon-based donor electron spin quantum computer architecture, complementary to the original Kane's nuclear spin proposal. This gate-sequence-optimal or time-optimal quantum gate control in a quantum circuit is in addition to the more conventional concept of optimality in terms of the number of elementary gates needed in a quantum transformation.

The optimal control method we use is the so-called gradient ascent pulse engineering (GRAPE) scheme. We focus on the high fidelity controlled-NOT (CNOT) gate and explicitly find the digitized control sequences by optimizing the effective, reduced donor electron spin Hamiltonian, with external controls over the hyperfine *A* and exchange *J* interactions. We first try different piecewise constant control steps and numerically calculate the fidelity (error) against the time needed to implement a CNOT gate with stopping criteria of error in the optimizer set to 10^{-9} in order to economize the simulation time. Here, the error is defined as $1-F$, where F is fidelity. The error is less than 10^{-8} for times longer than $100ns$, and it is found that 30 piecewise constant control steps for the CNOT gate operation will be sufficient to meet the required fidelity (error), and the performance would not be improved further with more steps.

With operation time $t = 100ns$ and stopping criteria of error set to 10^{-16}, we can find that the near time-optimal, high-fidelity CNOT gate control sequence has an error of 1.11×10^{-16}. We then simulate the control sequences of the CNOT gate, obtained from reduced Hamiltonian simulations, with the full spin Hamiltonian. We find the error of about 10^{-6} which is below the error threshold required for fault-tolerant (10^{-4}) quantum computation. The CNOT gate operation time of $100ns$ is 3 times faster than the globally controlled electron spin scheme of $297ns$. One of the great advantages of this near optimal-time high fidelity CNOT gate is that the exchange interaction is not required to be strong (the maximum value is $\mathrm{J/h} = 20\mathrm{MHz}$ compared to the typical value of $10.2\mathrm{GHz}$. This relaxes significantly the stringent distance constraint of two neighboring donor atoms of about $10\mathrm{nm}$ as reported in the original Kane's proposal to be about $30\mathrm{nm}$ which is within the reach of the current fabrication technology.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Thanks to quantum Shor's algorithm [1, 2] and quantum Grover's algorithm [3–5], quantum computation has drawn much attention and interest in recent years. That makes it possible to efficiently factorize semi-prime integers and that enables searches within unsorted database . These algorithms are the key motivators for people to develop controllable quantum systems to construct universal quantum gates [6] which can be used to implement arbitrary unitary operations. Due to decoherence and lack of precision in quantum control, the error rate can be defined as $1 - q$, where the quality $q \approx Fe^{-T/T_2}$, F denotes the fidelity of a gate operation of duration T, and T_2 is the pertinent overall decay time. One of the most important theoretical challenge is how to make the device work fast enough to meet the error correction threshold. The threshold for fault-tolerant quantum computation is estimated to be likely about 10^{-4}. Since CNOT gate is conventionally the most important gate in the universal gates, this work focuses on finding control parameter sequence in time optimal way for a CNOT gate in Si:P based electron spin quantum computer architectures [7,8] where the electron spin is defined as qubit. It's also required that the error of applying the control parameter sequence should be below the threshold for fault-tolerant quantum

computation.

Up to now, there are different approaches in optimal operation time problems [9–11]. One of these approaches uses the variational principle to find the optimal control sequence with parameters which are continuous in time [9, 10]. One of the constrains of their theory is the total energy of the system they study is a finite constant value, and they find the shortest path in time is the geodesic line in the Fubini-Study manifold. It's analogous to finding an optimal time path from one place to another place with specific energy (constant speed) on a sphere by using variational principle, and the result is the geodesic line on the sphere. The advantage of this approach is that the optimal time could be calculated naturally by solving the differential equation obtained by the variational method. However, analytical solutions of the differential equations are hard to find for somewhat complicate systems. In addition, the constrain of total energy has to be constant with total energy may not be practical for real world systems. We nevertheless investigate the Controlled-Z gate for a model Hamiltonian for quantum computation in Chapter 4.

Another approach called gradient ascent pulse engineering (GRAPE) [11] partitions a given time into several equal time steps, and in each time step of the sequence, the amplitudes of control parameters are constant. For a desired operation, we can define the trace fidelity between the desired operation and the unitary operation from the sequence. Since we can calculate the derivative of fidelity with respect to the control amplitudes in each step, we will be able to obtain, given the required fidelity, the near time optimal control sequence numerically. The advantage of this approach is that the constraint of a constant total energy is not required. It requires only that the energy is smaller than some certain fixed value. It's more flexible and practical in real world systems - that is to say, we could turn all the control parameters in the Hamiltonian off or turn them all on simultaneously. Although this method can not

obtain the optimal time formally, we can still optimize the fidelity in a given time. If the fidelity does not meet the required threshold, we simply extend the time and optimize the fidelity again and see if the fidelity is acceptable or not. The minimum time sequence that meets the required fidelity (error) threshold is the near time-optimal control sequence.

On the other hand, it has been shown that the gate fidelity of Kane's QC is limited primarily by the electron decoherence time where the typical gate operation timescale was expected to be closer to O(μs) [12–14]. Some experiments [15] indicate that the decoherence time for phosphorous donor electron spin in purified Si is considerably longer than 60 ms at $T = 4$K. Since the quality is related to the operation time, a considerable error may be still caused by the decoherence even though the fidelity in an ideal case. In Ref. [16], a globally controlled electron spin quantum computing scheme is proposed. In that proposal, the CNOT gate operation can be done around $296.8ns$ (in the paper [16], the indicated CNOT time is 148ns which is due to a factor of 2 missing in the denominator of the σ_z term of their Hamiltonian). In addition to the error from fidelity, the error rate from decoherence can be estimated to be $1 - e^{-T/T_2} = 5 \times 10^{-6}$ (by assuming $F = 1$), if the total error is around $1 - q \lesssim 10^{-4}$, then we can use error correction algorithms to correct the error.

Recently, the GRAPE algorithm has been applied to the coupled Josephson qubit quantum computing architecture [17], and the numerically optimal control time for a CNOT gate decompositions is found to be $55ps$ [18] instead of the $255ps$ in Ref. [17]. The GRAPE technique can be extended to pseudospin systems that have Hamiltonian expressed with a closed Lie algebra, e.g., SU(2^N) for a system of N qubits. Hence, we will use the GRAPE algorithm to investigate the time-optimal CNOT gate in the Kane's quantum computing scheme and discuss the influence of noise that may decrease fidelity.

This thesis is organized as follows: Chapter 2 provides an overview of quantum computing and the Kane's quantum computer architecture. Chapter 3 describes the background material of GRAPE algorithm that is useful in later parts. In Chapter 4, we study the time-optimal unitary operation using the variational principle approach for a model Hamiltonian for quantum computation. Specifically, we construct the controlled-Z gate using the variational principle approach and canonical gate decomposition method, respectively. Chapter 5 describes how to implement CNOT gate using GRAPE algorithm, and discusses the results when the control gate voltages are influenced by Gaussian noise. Finally, Chapter 6 summarizes the findings of this thesis.

Chapter 2

The Si based Quantum Computer and Quantum Computing

2.1 The architecture and the Hamiltonian

The architecture of the Kane's Si-based quantum computer is schematically in Fig. 2.1. It is composed of ^{31}P atoms doped in a array in a purified ^{28}Si ($I = 0$) host where each phosphorous has a nuclear spin ($I = \frac{1}{2}$) Because each ^{31}P atom has five valence electrons, as a first approximation, four of these electrons form covalent bounds with neighboring Si atoms, and the fifth electron forms a hydrogen-like S-orbital structure around each $^{31}P^{+}$ atom with effective Bohr radius and bound state energy levels given by

$$a_B^* = \epsilon \frac{m_e}{m^*} a_B \ , \tag{2.1}$$

$$E_n = \frac{m^*}{\epsilon^2 m_e} E_n^H \ , \tag{2.2}$$

where $\epsilon = 11.7$ is the static dielectric constant in Si, and the value of the effective mass m^* in Si is $m^* \approx m_T^* = 0.2 m_e$ (where m_e is the free elecrton mass, and m_T^* is the

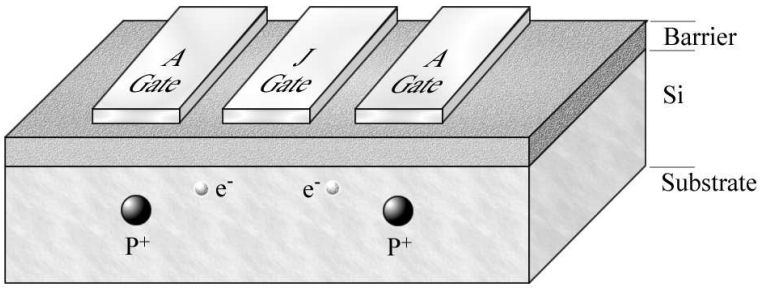

Figure 2.1 A schematic diagram of two qubits in Kane's architecture containing ^{31}P donors and electrons in a Si host where A gate controls the hyperfine coupling strength between electron and nuclei, and J gate controls the electron-mediated coupling electrons.

transverse effective mass in Si). With the Bohr radius and the bound state energies for a hydrogen atom : $a_B = 0.53\mathring{A}$ and $E_n^H = -13.6\text{eV/n}^2$, we obtain effective $a_B^* \approx 30\mathring{A}$ and $E_1 \approx -20\text{meV}$.

At sufficient low temperatures, the extra electron donor will only occupy the lowest energy bound state so that the electron wave function will be 1s orbital and concentrated at the donor nucleus, yielding a large hyperfine coupling energy. The strength A of the contact hyperfine interaction is proportional to the value of the electron probability density evaluated at the nucleus,

$$A = \frac{8\pi}{3}\mu_B g_n \mu_n |\psi(0)|^2 \, . \tag{2.3}$$

If we apply a negative voltage on A gate, the wave function would be distorted and repelled away from the nucleus, therefore reducing the strength of the hyperfine interaction. The typical value for the hyperfine interaction is $A = 1.21 \times 10^{-7}\text{eV}$. According to numerical calculations [19], it may be possible to vary the hyperfine coupling by up to $\approx 50\%$ before the donor electron is ionized.

Now, we apply an external magnetic field $\mathbf{B}_0$ to break the shallow donor electron ground state two-fold spin degeneracy such that we can use the two level electron

spin system as qubits. The single-qubit Hamiltonian can be written as

$$H = \frac{1}{2} g_e \mu_B B_0 \sigma_z^e - \frac{1}{2} g_n \mu_n B_0 \sigma_z^n + A \boldsymbol{\sigma}^e \cdot \boldsymbol{\sigma}^n \ , \tag{2.4}$$

where the effective g-factor of an electron in Si is $g_e = 2$, the g-factor for a ^{31}P nuclear spin is $g_n = 2.26$, $\mu_B = 5.788 \times 10^{-5}$eV/T is the electron magneton, and $\mu_n = 3.152 \times 10^{-8}$eV/T is the nuclear magneton. A typical value of $\mathbf{B}_0$ is 2.0T, giving the Zeeman energy for the electron of $\frac{1}{2} g_e \mu_B B_0 = 1.159 \times 10^{-4}$eV, and for the nucleus of $\frac{1}{2} g_n \mu_n B_0 = 7.124 \times 10^{-8}$eV. Under the influence of a constant magnetic field $\mathbf{B}_0$, electrons and nucleus will undergo a Larmor precession around the Z axis. The electron spin and nuclear spin may flip due to part of the terms of the hyperfine interaction, $A\sigma_x \otimes \sigma_x + A\sigma_y \otimes \sigma_y$. Once the electron spin flips, the nuclear spin will flip. Contrariwise, the electron spin will flips, once the nuclear spin flips. Because the Zeeman energy of the electron spin is 1000 times larger than the Zeeman energy of the nuclear spin and the energy conservation, the probabilities that the electron spin and nuclear spin flip are very small. Therefor, if we initiate the nuclear spin in the lowest energy orientation - spin up, we might change the effective Lamor precession frequency of a selected electron through tuning the hyperfine interaction strength achieved by applying a voltage on its surface of A gate. Since the energy difference between the spin-up and spin-down state of the targeted electron could be controlled, the qubit can be selectively addressed.

First, we diagonalize the Hamiltonian in Eq. (2.4) to analyze the energy levels of the system. Through direct diagonization of this Hamiltonian, we obtain analytically

the eigen-energies and eigen-states as follows :

$$E_{|\uparrow_e\uparrow_n\rangle'} = \frac{1}{2}g_e\mu_B B_0 - \frac{1}{2}g_n\mu_n B_0 + A \,, \tag{2.5a}$$

$$E_{|\uparrow_e\downarrow_n\rangle'} = \sqrt{(\frac{g_e\mu_B B_0 + g_n\mu_n B_0}{2})^2 + 4A^2} - A \,, \tag{2.5b}$$

$$E_{|\downarrow_e\uparrow_n\rangle'} = -\sqrt{(\frac{g_e\mu_B B_0 + g_n\mu_n B_0}{2})^2 + 4A^2} - A, \tag{2.5c}$$

$$E_{|\downarrow_e\downarrow_n\rangle'} = -\frac{1}{2}g_e\mu_B B_0 + \frac{1}{2}g_n\mu_n B_0 + A \,, \tag{2.5d}$$

and

$$|\uparrow_e\uparrow_n\rangle' = |\uparrow_e\uparrow_n\rangle \,, \tag{2.6a}$$

$$|\uparrow_e\downarrow_n\rangle' = \cos(\frac{\phi}{2})\,|\uparrow_e\downarrow_n\rangle + \sin(\frac{\phi}{2})\,|\downarrow_e\uparrow_n\rangle \,, \tag{2.6b}$$

$$|\downarrow_e\uparrow_n\rangle' = -\sin(\frac{\phi}{2})\,|\uparrow_e\downarrow_n\rangle + \cos(\frac{\phi}{2})\,|\downarrow_e\uparrow_n\rangle \,, \tag{2.6c}$$

$$|\downarrow_e\downarrow_n\rangle' = |\downarrow_e\downarrow_n\rangle \,, \tag{2.6d}$$

where $\phi = \tan^{-1}\left(\frac{2A}{\frac{1}{2}g_e\mu_B B_0 + \frac{1}{2}g_n\mu_n B_0}\right)$. Due to the hyperfine interaction, the eigen-states are not in computational basis: $|\uparrow_e\uparrow_n\rangle$, $|\uparrow_e\downarrow_n\rangle$, $|\downarrow_e\uparrow_n\rangle$, and $|\downarrow_e\downarrow_n\rangle$. Yet the typical hyperfine interaction $2A$ is about 500 times smaller than the Zeenman energy $\frac{1}{2}g_e\mu_B B_0 + \frac{1}{2}g_n\mu_n B_0$, and we thus have $\phi \approx 0$. Therefore, we can treat these states as approximatively splitting into electric and nuclear spin direct product states with small energy perturbation from the hyperfine interaction. As we calculate the energy states perturbatively by treating the hyperfine interaction A as a perturbation, or we

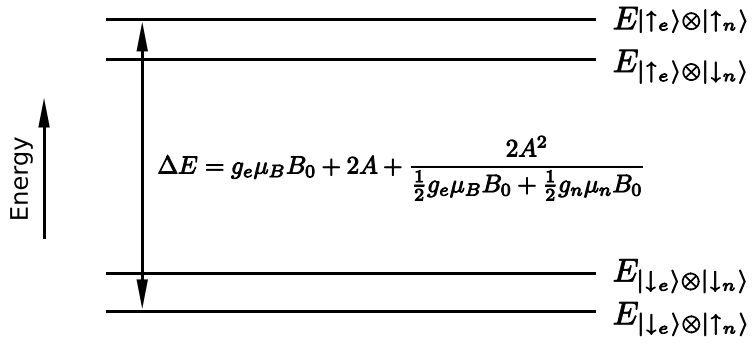

Figure 2.2 Energy levels of the donor electron-nucleus system obtained by using 2nd order approximation with magnetic field B and hyperfine coupling A.

simply expand the result of Eq. (2.5) up to 2nd order, the following can be obtained

$$E_{|\uparrow_e\uparrow_n\rangle} = \frac{1}{2}g_e\mu_B B_0 - \frac{1}{2}g_n\mu_n B_0 + A \,, \tag{2.7a}$$

$$E_{|\uparrow_e\downarrow_n\rangle} = \frac{1}{2}g_e\mu_B B_0 + \frac{1}{2}g_n\mu_n B_0 - A + \frac{2A^2}{\frac{1}{2}g_e\mu_B B_0 + \frac{1}{2}g_n\mu_n B_0} \,, \tag{2.7b}$$

$$E_{|\downarrow_e\uparrow_n\rangle} = -\frac{1}{2}g_e\mu_B B_0 - \frac{1}{2}g_n\mu_n B_0 - A - \frac{2A^2}{\frac{1}{2}g_e\mu_B B_0 + \frac{1}{2}g_n\mu_n B_0} \,, \tag{2.7c}$$

$$E_{|\downarrow_e\downarrow_n\rangle} = -\frac{1}{2}g_e\mu_B B_0 + \frac{1}{2}g_n\mu_n B_0 + A \,, \tag{2.7d}$$

This is shown schematically in Fig. 2.2.

Since the eigen-states are not perfectly simply the spin-up and spin-down direct product states, one might think that the hyperfine interaction would cause both electron spin and the nuclear spin to flip. But it can be proven [20] that this effect is minor, and in our full Hamiltonian simulation latter, it will be shown again that this is the case.

In order to demonstrate the ability to perform quantum computation, we have

to construct some universal gate, such as a CNOT gate and arbitrary single-qubit operators. As a result, we apply an extra background rf field B_{ac}, and we might tune the hyperfine interaction to bring the selected electron qubits into resonance and giving us the ability to control the qubits rotating between the spin-up and spin-down states. Because it's not easy to control and turn on and off the rf magnetic field quickly at the precise times in an experiment, in this setup, we always let the rf magnetic field be on. But it has been shown [16] that for typical parameters expected for the Kane architecture, if the canonical scheme works, the energy difference due to the hyperfine interaction must be large compared to the full width half maximum (FWHM) of the resonance given by $4\mu_B B_{ac}/\hbar$ which leads to small B_{ac} field and long gate operation time of $1.7\mu s$. To solve this issue, they [16] introduce the global addressing proposal.

In our scheme, we reserve the rf magnetic field to be always on, but we use the GRAPE algorithm to find the optimal control parameter sequence. One might think we could optimize the full Hamiltonian with B_{ac} directly - in other words, to find a control parameter sequence that leads to the desired unitary operation on the electron spin, with the evolution of nuclear spin being from the spin-up state to spin-up state again at the end of the operation. It has been discussed in the Ref. [11] that through finding the optimal control parameters, it's only possible to implement either a desired unitary operation, or to implement an evolution from a specific state to another. As a result, there exits a theoretical bottleneck to find a set of parameters that allows both the implementation of a unitary operation on one part of the system and the evolution of particular states in the other part of the system. In order to solve this problem, we will introduce the reduced Hamiltonian approximated from assuming the nuclear spins always to be up. Then we will be able to obtain a set of control parameters by optimizing quantum gate operations using the reduced Hamiltonian.

2.2 The reduced Hamiltonian for a single qubit

If we initialize the nuclear spin up and apply a dc magnetic field B_0, then the energy difference between electron spin up and down state from Eq. (2.7a) and Eq. (2.7c) is given to 2nd order in A by

$$\Delta E(A) = g_e\mu_B B_0 + 2A + \frac{2A^2}{\frac{1}{2}g_e\mu_B B_0 + \frac{1}{2}g_n\mu_n B_0} \ . \tag{2.8}$$

By defining $\omega(A) = \Delta E(A)/\hbar$, the effective Hamiltonian can be written as $H = (\hbar/2)\omega\sigma_z^e$. The Hamiltonian with the rf B_{ac} field, is then

$$H = \frac{\hbar\omega}{2}\sigma_z + \frac{1}{2}g_e\mu_B B_{ac}(\cos\omega_{ac}t\sigma_x + \sin\omega_{ac}t\sigma_y) \ , \tag{2.9}$$

where ω_{ac} is the angular frequency of rf magnetic field. It is much simpler to understand the control processes, once we go into the frame rotating with the rf field. We make the substitution,

$$\tilde{\rho} = U_{rot}\rho U_{rot}^{\dagger} \ , \tag{2.10}$$

where

$$U_{rot} = e^{\frac{i}{2}\omega_{ac}\sigma_z t} \ . \tag{2.11}$$

Inserting Eq. (2.10) and Eq. (2.11) into Liouville-Von Neumann equation, Eq. (3.3), which will be described later, and with the identity $e^{\frac{i}{2}\omega_{ac}\sigma_z t}(\cos\omega_{ac}t\sigma_x + \sin\omega_{ac}t\sigma_y)e^{-\frac{i}{2}\omega_{ac}\sigma_z t} = \sigma_z$, we have the reduced Hamiltonian in the rotating frame,

$$\tilde{H} = \frac{\hbar}{2}(\omega(A) - \omega_{ac})\sigma_z + \frac{1}{2}g_e\mu_B B_{ac}\sigma_x \ . \tag{2.12}$$

We tune the angular frequency of rf magnetic field B_{ac} to the electron spin resonance frequency obtained when no voltage is applied to the corresponding A gate, that is $\omega_{ac} = \omega(A_0)$. If we define $\Delta\omega = \omega(A) - \omega(A_0)$, then the qubits will effectively rotate

around the x-axis when $\Delta\omega = 0$, or equivalently $A = A_0$, and around an axis which is slightly shifted with respect to this axis (when $\Delta\omega \neq 0$, or $A \neq A_0$) described by Eq. (2.12). The typical value of the energy difference due to the detuning of changing the hyperfine interaction is $(\hbar/2)\Delta\omega = -6.065 \times 10^{-8}$eV.

2.3 Two-qubit system

The spins of two adjacent electrons can be coupled by the exchange energy J. The donor electrons are comparatively loosely bound to the P donor allowing neighboring electrons' wave-functions to overlap. The strength of the exchange coupling can be approximated using the Herring-Flicker approximation [21],

$$J(a^*, d) \approx \frac{1.6e^2}{\hbar\epsilon a^*}(\frac{d}{a^*})^{5/2} \exp(-2\frac{d}{a^*}) \,, \tag{2.13}$$

where ϵ is the dielectric constant of the semiconductor. The approximated formula is valid when the inter-donor spacing, d, is much greater than the effective Bohr radius, a^*. The exchange energy drops off exponentially, and it is thought of as a short range interaction. If a positive voltage is applied on a J gate, the electron wave-functions will be attracted toward the center, and the overlap will be increased, therefore increasing the strength of exchange energy. Conversely, applying a negative voltage will decrease the exchange energy. In a typical Kane quantum computer's scheme, the typical value of J is 4.23×10^{-5}eV that requires the separation between two neighboring donors is roughly about $100 - 200\mathring{A}$ that sets a stringent fabrication condition to fabricate surface A and J gates within such a short distance. In the globally addressing proposal [16], during the time (300ns) of CNOT operation, the J gate is only on for 0.02ns with a value of 4.23×10^{-5}eV. In that proposal, the energy scale of J is much larger than the energy splitting in the σ_z and σ_x terms, if there is any slight noise on J gate when it's supposed to be turned off, it will damage the

reliability and fidelity of the computing. If one would like to avoid this situation, one could increase the separation distance of two donors. However, this will increase the operation time dramatically, since it is hard to perform the electron spin exchange interaction and other interactions in parallel in the conventional gate decomposition scheme. Another solution is using the robust controlled method [22] to correct the error from the variations of exchange interaction, but it requires very long operation time (751ns) for 2nd level corrected CNOT gate. In our scheme, we can numerically calculate the optimal path. In other words, we can perform spin exchange interaction and other interactions in parallel so that we can use a smaller exchange interaction energy without increasing the operation time. Here, our maximum exchange energy will be $J \approx 8.3 \times 10^{-8}$eV corresponding to a donor separation around $300\mathring{A}$ in which to fabricate surface gates is within the reach of the current fabrication technology. This is one of the great advantages in our scheme. The Hamiltonian of the exchange interaction can be written as

$$H_J = J\boldsymbol{\sigma}^{1e} \cdot \boldsymbol{\sigma}^{2e} \,, \tag{2.14}$$

Because the rotating operator U_{rot} of Eq. (2.11) commutes with the exchange interaction Hamiltonian H_J, the effective reduced two-qubit Hamiltonian with the exchange interaction can be simply written as

$$\tilde{H} = \frac{\hbar}{2}\Delta\omega_1\sigma_z^1 + \frac{\hbar}{2}\Delta\omega_2\sigma_z^2 + \frac{1}{2}g_e\mu_B B_{ac}\left(\sigma_x^1 + \sigma_x^2\right) + J\boldsymbol{\sigma}^{1e} \cdot \boldsymbol{\sigma}^{2e} \,. \tag{2.15}$$

Finally, we can write down the full two-qubit Hamiltonian as

$$\begin{aligned} H = \quad & \tfrac{1}{2}g_e\mu_B B_0\left(\sigma_z^{1e} + \sigma_z^{2e}\right) - \tfrac{1}{2}g_n\mu_n B_0\left(\sigma_z^{1n} + \sigma_z^{2n}\right) \\ & + \tfrac{1}{2}g_e\mu_B B_{ac}\left(\cos\omega_{ac}t(\sigma_x^{1e} + \sigma_x^{2e}) + \sin\omega_{ac}t(\sigma_y^{1e} + \sigma_y^{2e})\right) \\ & - \tfrac{1}{2}g_n\mu_n B_{ac}\left(\cos\omega_{ac}t(\sigma_x^{1n} + \sigma_x^{2n}) + \sin\omega_{ac}t(\sigma_y^{1n} + \sigma_y^{2n})\right) \\ & + A_1\boldsymbol{\sigma}^{1e} \cdot \boldsymbol{\sigma}^{1n} + A_2\boldsymbol{\sigma}^{2e} \cdot \boldsymbol{\sigma}^{2n} + J\boldsymbol{\sigma}^{1e} \cdot \boldsymbol{\sigma}^{2e} \,, \end{aligned} \tag{2.16}$$

Table 2.1 Typical parameters used for numerical calculations.

Description	Term	Typical Value
Planck Constant ($\frac{\hbar}{2\pi}$)	$\hbar$	$6.58211889(26) \times 10^{-16}$ eV s
Electron Mass	m_e	$9.10938188(72) \times 10^{-31}$ kg
Proton Mass	m_n	$1.67262158(13) \times 10^{-27}$ kg
Elementary Charge	e	$1.602176462(63) \times 10^{-19}$ C
Bohr Magneton ($\frac{e\hbar}{2m_e}$)	μ_B	$5.788381749(43) \times 10^{-5}$ eV T^{-1}
Proton Bohr Magneton ($\frac{e\hbar}{2m_e}$)	μ_n	$3.152451241 \times 10^{-8}$ eV T^{-1}
Electron g-factor ($2\mu_e\mu_B$)	g_e	2.0023193043737(82)
Effective Proton g-factor in Si ($2\mu_e\mu_B$)	g_n	2.26
Unperturbed Hyperfine Interaction	A_0	1.211×10^{-7} eV
Minimum Varied Hyperfine Interaction	A_p	0.606×10^{-7} eV
Constant Magnetic Field Strength	B_0	2.0 T
Electron Zeeman Energy ($\frac{1}{2}g_e\mu_B B_0$) at B_0		$1.159018851 \times 10^{-4}$ eV
Nuclear Zeeman Energy ($\frac{1}{2}g_n\mu_n B_0$) at B_0		$7.124539805 \times 10^{-8}$ eV
Maximum Exchange Interaction	J	8.3×10^{-8} eV
Energy Difference in Reduced Hamiltonian	$\frac{\hbar}{2}\omega\sigma_z$	-6.065×10^{-8} eV

which would be used in the simulation later.

Chapter 3

GRAPE algorithm

In this chapter, we will briefly introduce some background material of GRAPE algorithm. Basically, the theoretical derivation follows Navin Khaneja's work [11].

3.1 Density Matrix Formalism

For a closed quantum system, each possible microstate $|\psi_i\rangle$ obeys Schrödinger equation,

$$\frac{\partial}{\partial t}|\psi_i\rangle = -\frac{i}{\hbar}H|\psi_i\rangle, \tag{3.1}$$

where H is the total Hamiltonian. The most general density matrix can be defined as

$$\rho = \sum_i P_i |\psi_i\rangle \langle\psi_i|, \tag{3.2}$$

where the coefficients P_j are non-negative and add up to one. This represents a statistical mixture of pure states. One can think of a mixed state as representing a single system that the mixed state represents an ensemble of systems, i.e. a large number of copies of the system in question, where P_j is the probability of the ensemble being in the microstate $|\psi_i\rangle$. An ensemble is described by a pure state if every copy

of the system in that ensemble is in the same state, i.e. it is a pure ensemble. Taking the time derivative of ρ and inserting Eq. (3.1), we have

$$\begin{aligned}
\dot{\rho} &= \sum_i P_i \left(|\dot{\psi}_i\rangle \langle\psi_i| + |\psi_i\rangle \langle\dot{\psi}_i| \right) \\
&= \sum_i P_i \left(-\frac{i}{\hbar} H |\psi_i\rangle \langle\psi_i| + \frac{i}{\hbar} |\psi_i\rangle \langle\psi_i| H \right) \\
&= -\frac{i}{\hbar} \left(H \sum_i P_i |\psi_i\rangle \langle\psi_i| - \sum_i P_i |\psi_i\rangle \langle\psi_i| H \right) \\
&= -\frac{i}{\hbar} (H\rho - \rho H) \\
&= -\frac{i}{\hbar} [H, \rho], \qquad (3.3)
\end{aligned}$$

where P_i is time independent as we have mentioned before. Equation (3.3) is called Liouville-Von Neumann equation of motion for the density matrix. Note that Liouville equation is only valid when the system is closed, so it's not valid for the subsystem of a composite system whose subsystems have interaction to each other. The equation can only describe the whole closed composite system which includes a subsystem in which we are interested and the rest of the system.

3.2 Optimal Unitary transformations

Assume we have m control parameters and a chosen transfer time T discretized in N equal steps of duration $\Delta t = T/N$, then the time dependent Hamiltonian can be written as

$$H(t) = H_0 + \sum_{k=1}^{m} u_{kj} H_k \,, \qquad (3.4)$$

where u_{kj} is the control amplitude of the kth control parameter during the time between $(j-1)t$ and jt, and each kth parameter has an upper bound and a lower bound. The state of the closed quantum system can be characterized by the density

operator $\rho(t)$ with an equation of motion of Liouville-von Neumann mentioned in Eq. (3.3). Solving Eq. (3.3) with the Hamiltonian defined in Eq. (3.4) gives that the final density operator at time $t = T$ is

$$\rho(T) = U_N \dots U_1 \rho_0 U_1^\dagger \dots U_N^\dagger \ , \tag{3.5}$$

where the propagator during a time step j is given by

$$U_j(\Delta t) = \exp\left\{ -\frac{i}{\hbar}\Delta t \left(H_0 + \sum_{k=1}^{m} u_{kj} H_k \right) \right\} \ , \tag{3.6}$$

and the propagator at final time T is

$$U_F = U_N \dots U_1 \ . \tag{3.7}$$

To calculate U_j numerically, the irreducible (p, p) degree rational Padé approximation (see Appendix A and the Refs. [23, 24]) has been used for computing matrix exponential.

Let us consider a problem to create a desired unitary operator U_D in a given time T by applying a pulse sequence u'_{kj}. At first, we guess all of the parameters u_{kj} randomly and name the parameters as initial parameters such that at the final time T, the unitary operator will be U_F. Define a performance function $\Phi(U_D, U_F)$ such that when U_F is equal to U_D up to an arbitrary phase factor $\exp(i\phi)$, we have $\Phi(U_D, U_F) = 2^N$ in a system of N qubits. Such function can be defined as

$$\begin{aligned} \Phi(U_D, U_F) &= |\mathrm{Tr}\{U_D^\dagger U_F\}|^2 \\ &= |\mathrm{Tr}\{\underbrace{U_D^\dagger \dots U_{j+1}}_{B_j^\dagger} \underbrace{U_j \dots U_1}_{F_j}\}|^2 \\ &= \mathrm{Tr}\{B_j^\dagger F_j\}\mathrm{Tr}\{F_j^\dagger B_j\} \ , \end{aligned} \tag{3.8}$$

where F_j is the forward propagation such that the density operator is $\rho(t) = F_j \rho_0 F_j^\dagger$ at time $t = j\Delta t$, and B_j is the backward propagation such that the prime density

operator $\rho'(t) = B_j \rho_0 B_j^\dagger = U_{j+1}^\dagger \dots U_N^\dagger \rho_D U_N \dots U_{j+1}$ at the same time $t = j\Delta t$, where ρ_D is desired final state evolved from your input state ρ_0. Since the initial trying parameters are produced by guessing randomly, the performance function should not be equal to 2^N. But if we can know the derivative of Φ with respect to the control amplitudes, u_{kj}, we may use multi-dimesion optimization technique to find the control parameters from the initial guessing parameters such that the performance function is maximum, that is finding our desired control sequence. Using the standard formula

$$\frac{\mathrm{d}}{\mathrm{d}x}\mathrm{e}^{A(x)} = \int_0^1 \mathrm{e}^{\tau A(x)} \frac{\mathrm{d}A(x)}{\mathrm{d}x} \mathrm{e}^{(1-\tau)A(x)} \mathrm{d}\tau \ , \tag{3.9}$$

where $A(x)$ is an operator, we have

$$\begin{aligned} \frac{\mathrm{d}U_j(\Delta t)}{\mathrm{d}u_{kj}} &= -\frac{i}{\hbar}\Delta t \left(\int_0^1 U_j(\Delta t\tau) H_k U_j(-\Delta t\tau) \mathrm{d}\tau \right) U_j(\Delta t) \ , \\ &= -\frac{i}{\hbar}\Delta t \bar{H}_k U_j(\Delta t) \ , \end{aligned} \tag{3.10}$$

where $\bar{H}_k = \int_0^1 U_j(\Delta t\tau) H_k U_j(-\Delta t\tau) \mathrm{d}\tau$.

For small Δt (when $\Delta t \ll ||\frac{1}{\hbar}(H_0 + \sum_{k=1}^m u_{kj} H_k)||^{-1}$), $\bar{H}_k \approx H_k$. Take the derivative Φ with respect to u_{kj} with Eq. (3.10), we have

$$\begin{aligned} \frac{\partial \Phi}{\partial u_{kj}} &= \mathrm{Tr}\left[\left(\frac{\partial B_j^\dagger}{\partial u_{kj}} \right) F_j + B_j^\dagger \left(\frac{\partial F_j}{\partial u_{kj}} \right) \right] \mathrm{Tr}\{F_j^\dagger B_j\} + c.c. \\ &= -2\mathrm{Re}\left\{ \frac{i}{\hbar}\Delta t \mathrm{Tr}\{B_j^\dagger H_k F_j\} \mathrm{Tr}\{F_j^\dagger B_j\} \right\} \ , \end{aligned}$$

where $B_j = U_{j+1}^\dagger \dots U_N^\dagger U_D$, and $F_j = U_j \dots U_1$.

Now, it becomes an optimization problem with constraints on the boundary which limits the range of the control amplitudes. The optimizer we implemented here is the spectral projected gradient method [25]. Since this algorithm is based on the gradient ascent procedure, there is no guarantee that it will converge to a global maximum - it may also converge to a local maximum. Nevertheless, we will try different guessed initial control sequences to see if they all fall into the same converged point or not.

3.3 Optimal Transfer between Hermitian density operators

In previous section, we describe the case of implement a desired unitary operation. In this section, let us consider a problem to create a transformation between initial Hermitian density operator, $\rho(0)$ at $T = 0$, and desired Hermitian density operator, ρ_D, in a given time T by applying a pulse sequence u'_{kj}. Using the same technique as in the pervious section, we guess at first all of the parameters u_{kj} randomly and name the parameters as initial parameters such that at the final time T, the final density operator will be $\rho_F = \rho(T) = U_N \dots U_1 \rho_0 U_1^\dagger \dots U_N^\dagger$. The overlap of two Hermitian operators ρ_F and ρ_D can be measured by the standard inner product, $\mathrm{Tr}\{\rho_D^\dagger \rho_F\}$. Therefore, we can define a performance function which is a real number as

$$\begin{aligned}
\Phi(\rho_D, \rho_F) &= \mathrm{Tr}\{\rho_D^\dagger \rho_F\} \\
&= \mathrm{Tr}\{\rho_D^\dagger U_N \dots U_1 \rho_0 U_1^\dagger \dots U_N^\dagger\} \\
&= \mathrm{Tr}\{\underbrace{U_{j+1}^\dagger \dots U_N^\dagger \rho_D^\dagger U_N \dots U_{j+1}}_{\lambda_j^\dagger} \underbrace{U_j \dots U_1 \rho_F U_1^\dagger \dots U_j^\dagger}_{\rho_j}\} \\
&= \mathrm{Tr}\{\lambda_j^\dagger \rho_j\} \,, \qquad (3.11)
\end{aligned}$$

, where ρ_j is the density operator $\rho(t)$ at time $t = j\Delta t$ and λ_j is the backward propagated target operator ρ_D at the same time $t = j\Delta t$.

Now, like the pervious section, if we know the derivative of Φ with respect to the control amplitudes, u_{kj}, we may optimize the performance function $\Phi(\rho_D, \rho_F) \approx 1$.

For small Δt (when $\Delta t \ll ||\frac{1}{\hbar}(H_0 + \sum_{k=1}^m u_{kj} H_k)||^{-1}$), $\bar{H}_k \approx H_k$. Take the derivative Φ with respect to u_{kj} with Eq. (3.10), we have

$$\begin{aligned}
\frac{\partial \Phi}{\partial u_{kj}} &= \mathrm{Tr}\left\{\lambda_j^\dagger \left(\frac{\partial \rho_j}{\partial u_{kj}}\right)\right\} \\
&= -\frac{i}{\hbar}\Delta t \mathrm{Tr}\{\lambda_j^\dagger [H_k, \rho_j]\} \,, \qquad (3.12)
\end{aligned}$$

where $\lambda_j = U_{j+1} \dots U_N \rho_D^\dagger U_N^\dagger \dots U_{j+1}^\dagger$, and $\rho_j = U_j \dots U_1 \rho_F U_1^\dagger \dots U_j^\dagger$.

Finally, we can use an optimizer to find the parameters which maximum the performance function. There are several different kinds of applications that transfer between two hermitian density operators. For example, if we have two spin-down qubits which we want to entangle them into a Bell state, we can implement it by using the method of optimal transformation between Hermitian density operators since we know the final density operator and initial density operator.

Chapter 4

Variational Principle Approach of Time-Optimal Evolution

In contrast to the digitized time sequence of the control parameters found in the GRAPE technique, the variational principle approach uses the variational principle to find the time-optimal control sequence and parameters which are continuous in time. In their first paper [9], Carlini et al. present a general framework for finding the time-optimal evolution and the optimal Hamiltonian for a quantum system with a given set of initial and final state. They find that the time-optimal solution to constrained Hamiltonian (finite constant energy) is the geodesic equation for the Fubini-Study metric on $\mathbb{C}P^{n-1}$. In other words, with the constrain of a finite constant total energy, they find the shortest path in time is the geodesic line in the Fubini-Study manifold and the time-optimal Hamiltonian is time-independent. It's analogous to finding the time-optimal path between two points with a specific energy on a sphere by using variational principle, and the result is the geodesic line on the sphere with constant speed.

In their second paper [10], Carlini et al. extend their previous work into finding the

time-optimal realization of a target unitary operation using the variational principle. The main concept here is that they replace the projective space representing quantum state vectors with the space of unitary operators. In their first work [9], the time-optimal realization depends on the initial and final states. However, most applications in quantum computation need to implement a quantum gate - that is an unitary operation; therefore, this extension is of great importance and more directly relevant to subroutines in quantum computation.

In Sec. 4.1, we will derive the geodesic equation for the Fubini-Study metric on $\mathbb{C}P^{n-1}$ in more detail than the original paper [9], and we will also give the solution to the geodesic equation. In Sec. 4.2, we will basically follow the original paper [10] to define an action principle for the time-optimal unitary operator, then derive the fundamental equations of motion. In the last section of this chapter, Sec. 4.3, we will implement the controlled-Z gate by using this time-optimal approach and using the canonical decomposition approach, respectively. The controlled-Z gate is one of the most important unitary operations for implementing quantum algorithms. In particular, we can implement the CNOT gate from a controlled-Z gate conjugated by $I \otimes H$. Furthermore, one of the simplest ways to implement quantum Fourier transformations (QFTs) uses multiple controlled-Z gates (see for example [6]). The controlled-Z gate may be used in the construction of controlled-X and controlled-Y gates. The reason why we only realize the time-opmtial controlled-Z gate instead of CNOT gate here is that we encounter some difficulty in solving the equation of motions analytically, and we will discuses it in this section. Finally, we will compare the operation time between the time-optimal approach and the conventional canonical decomposition approach.

4.1 Time-optimal evolution between a given set of initial and final states

The problem which we are going to study here is analogous to the classical brachistochrone mechanic problem. In the classical brachistochrone problem, one has to find the shape of the curve down which a bead sliding from rest and accelerated by gravity will slip (without friction) from one point to another in the least time. The brachistochrone problem was one of the earliest problems posed in the calculus of variations. The time to travel from a point to another point is given by the integral

$$T = \int \frac{\mathrm{d}s}{v} \,, \tag{4.1}$$

where the parameter $s(t)$ specifies the length from initial point to the current position $x(t)$ of the particle, v is the speed of the current position defined by $v = \frac{\mathrm{d}s(t)}{\mathrm{d}t} = \sqrt{2[E - V(x)]/m}$, where E is the conserved energy and V is the gravitational potential. Using the calculus of variations, the problem can be solved, and the solution is a segment of a cycloid curve which does not depend on the body's mass or on the strength of the gravitational constant.

In the quantum version, one would like to find a time-dependent Hamiltonian such that the state evolves from a given initial state $|\psi_i\rangle$ to a given final state $|\psi_f\rangle$ belonging to a n-dimensional Hilbert space in the least time. Here, the state $|\psi(t)\rangle$ and $H(t)$ are dynamical variables, and the Lagrange in n-dimensional Hilbert space can be defined as

$$\begin{aligned} L(\psi, \dot{\psi}, H, \phi, \dot{\phi}, \lambda) =& \frac{\sqrt{\left\langle \frac{\mathrm{d}}{\mathrm{d}t}\psi \right| (1-P) \left| \frac{\mathrm{d}}{\mathrm{d}t}\psi \right\rangle}}{\Delta E} + \left(i \left\langle \frac{\mathrm{d}}{\mathrm{d}t}\phi \middle| \psi \right\rangle + \langle \phi \,|\, H \,|\, \psi \rangle + \text{c.c.} \right) \\ &+ \lambda \left(\frac{\mathrm{Tr}\tilde{H}^2}{2} - \omega^2 \right) \,, \end{aligned} \tag{4.2}$$

where $P(t) = |\psi\rangle \langle\psi|$ is the projection operator, $\tilde{H} = H - (\mathrm{Tr}H)/n$ is the traceless

part of the Hamiltonian, $(\Delta E)^2 = \langle\psi\,|\,H^2\,|\,\psi\rangle - \langle\psi\,|\,H\,|\,\psi\rangle^2$ is the energy variance, and ω is a given nonzero constant. The action can be defined as

$$\begin{aligned} S(\psi,\dot{\psi},H,\phi,\dot{\phi}) &= \int \mathrm{d}t L(\psi,\dot{\psi},H,\phi,\dot{\phi},\lambda) \\ &= \int \mathrm{d}t \left[\frac{\sqrt{\left\langle \frac{\mathrm{d}}{\mathrm{d}t}\psi\right| (1-P) \left|\frac{\mathrm{d}}{\mathrm{d}t}\psi\right\rangle}}{\Delta E} + \left(i\left\langle \frac{\mathrm{d}}{\mathrm{d}t}\phi \,\middle|\, \psi \right\rangle + \langle\phi\,|\,H\,|\,\psi\rangle + \text{c.c.} \right) \right. \\ &\quad \left. + \lambda\left(\frac{\mathrm{Tr}\tilde{H}^2}{2} - \omega^2 \right) \right] , \end{aligned} \tag{4.3}$$

where ψ, $\dot{\psi}$, H, ϕ and $\dot{\phi}$ are the dynamic variables. The first term in the action, Eq. (4.3), is interpreted as the time duration of the evolution expressed in terms of the Fubini-Study line element $\mathrm{d}s^2 = \langle \mathrm{d}\psi|\,(1-P)\,|\mathrm{d}\psi\rangle$. The second term guarantees that $|\psi(t)\rangle$ and $H(t)$ obey the Schrödinger equation through the Lagrange multiplier $|\phi(t)\rangle$. The third term constrains the constant of the energy through the Lagrange multiplier, λ. The finite energy constraint is necessary, otherwise one could rescale the Hamiltonian such that the duration time is arbitrarily small.

The equations of motion satisfies the Euler equation,

$$\frac{\partial L}{\partial q} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{q}} = 0 \,. \tag{4.4}$$

Since L is symmetrical, as a result, we will obtain the same equation of motions by the variation respect to $\langle\phi|$ or $|\phi\rangle$. The variation of Eq. (4.2) respect to $\langle\phi|$ by using Euler-Lagrange equation, Eq. (4.4), will gives us the Schrödinger Equation,

$$\begin{aligned} \frac{\partial L}{\partial \langle\phi|} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \left\langle \frac{\mathrm{d}}{\mathrm{d}t}\phi\right|} &= 0 \\ \therefore H\,|\psi\rangle - i\frac{\mathrm{d}}{\mathrm{d}t}\,|\psi\rangle &= 0 \,. \end{aligned} \tag{4.5}$$

The variation of Eq. (4.2) respect to λ leads to the constraint of Hamiltonian, finite

constant energy,

$$\begin{aligned}\frac{\partial L}{\partial \lambda}-\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{\lambda}}&=0\\ \therefore \frac{\mathrm{Tr}\tilde{H}^2}{2}&=\omega^2\ .\end{aligned}\tag{4.6}$$

The variation of Eq. (4.2) respect to $\langle\psi|$ with Eq. (4.5) implies

$$\begin{aligned}\frac{\partial L}{\partial \langle\psi|}-\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \left\langle\frac{\mathrm{d}}{\mathrm{d}t}\psi\right|}&=0\\ \therefore H|\phi\rangle-i\left|\frac{\mathrm{d}}{\mathrm{d}t}\phi\right\rangle+\left(\frac{\partial}{\partial\langle\psi|}-\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial}{\partial\left\langle\frac{\mathrm{d}}{\mathrm{d}t}\psi\right|}\right)\left(\frac{\sqrt{\left\langle\frac{\mathrm{d}}{\mathrm{d}t}\psi\right|(1-P)\left|\frac{\mathrm{d}}{\mathrm{d}t}\psi\right\rangle}}{\Delta E}\right)&=0\ ,\end{aligned}$$

where

$$\begin{aligned}&\frac{\partial}{\partial\langle\psi|}\frac{\sqrt{\left\langle\frac{\mathrm{d}}{\mathrm{d}t}\psi\right|(1-P)\left|\frac{\mathrm{d}}{\mathrm{d}t}\psi\right\rangle}}{\Delta E}\\ &=\frac{-\left\langle\frac{\mathrm{d}}{\mathrm{d}t}\psi\,\middle|\,\psi\right\rangle\left|\frac{\mathrm{d}}{\mathrm{d}t}\psi\right\rangle}{2\Delta E\sqrt{\left\langle\frac{\mathrm{d}}{\mathrm{d}t}\psi\right|(1-P)\left|\frac{\mathrm{d}}{\mathrm{d}t}\psi\right\rangle}}-\frac{\Delta E\left(\langle\psi|H^2|\psi\rangle-\langle H\rangle^2\right)^{-3/2}\left(H^2|\psi\rangle-2\langle H\rangle H|\psi\rangle\right)}{2}\\ &=\frac{-1}{2(\Delta E)^2}\langle H\rangle H|\psi\rangle-\frac{1}{2(\Delta E)^2}\left(H^2|\psi\rangle-2\langle H\rangle|\psi\rangle\right)\\ &=\frac{1}{2(\Delta E)^2}\left(\langle H\rangle H|\psi\rangle-H^2|\psi\rangle\right)\ ,\end{aligned}$$

and

$$\begin{aligned}&\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial}{\partial\left\langle\frac{\mathrm{d}}{\mathrm{d}t}\psi\right|}\frac{\sqrt{\left\langle\frac{\mathrm{d}}{\mathrm{d}t}\psi\right|(1-P)\left|\frac{\mathrm{d}}{\mathrm{d}t}\psi\right\rangle}}{\Delta E}\\ &=\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{1}{2\Delta E\sqrt{\left\langle\frac{\mathrm{d}}{\mathrm{d}t}\psi\right|(1-P)\left|\frac{\mathrm{d}}{\mathrm{d}t}\psi\right\rangle}}(1-P)\left|\frac{\mathrm{d}}{\mathrm{d}t}\psi\right\rangle\right)\\ &=-\frac{i}{2}\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{1}{(\Delta E)^2}H|\psi\rangle-\langle H\rangle|\psi\rangle\right)\\ &=i\left[\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\langle H\rangle-H}{2(\Delta E)^2}\right)\right]|\psi\rangle+i\left(\frac{\langle H\rangle-H}{2(\Delta E)^2}\right)\left|\frac{\mathrm{d}}{\mathrm{d}t}\psi\right\rangle\\ &=i\left[\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\langle H\rangle-H}{2(\Delta E)^2}\right)\right]|\psi\rangle+\left(\frac{\langle H\rangle-H}{2(\Delta E)^2}\right)H|\psi\rangle\ ;\end{aligned}$$

finally, we obtain

$$H\left|\phi\right\rangle - i\frac{\mathrm{d}}{\mathrm{d}t}\left|\phi\right\rangle + i\left[\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\langle H\rangle - H}{2(\Delta E)^2}\right)\right]\left|\psi\right\rangle = 0\ . \tag{4.7}$$

The variation with respect to H gives

$$\frac{\{\ H, P\ \} - 2\left\langle H\right\rangle P}{2\left(\Delta E\right)^2} - \lambda\tilde{H} - \left(\left|\psi\right\rangle\left\langle\phi\right| + \left|\phi\right\rangle\left\langle\psi\right|\right) = 0\ . \tag{4.8}$$

Let's take trace over Eq. (4.8). After use of $\mathrm{Tr}\left(\left|\psi\right\rangle\left\langle\psi\right|\right) = 1$, $\mathrm{Tr}\tilde{H} = 0$, and $\mathrm{Tr}(HP) = \langle H\rangle$, we obtain

$$\begin{aligned}&\frac{\mathrm{Tr}(HP) - \langle H\rangle\,\mathrm{Tr}(P)}{\left(\Delta E\right)^2} - \lambda\mathrm{Tr}\tilde{H} - \mathrm{Tr}\left(\left|\psi\right\rangle\left\langle\phi\right| + \left|\phi\right\rangle\left\langle\psi\right|\right)\\ &= -\mathrm{Tr}\left(\left|\psi\right\rangle\left\langle\phi\right| + \left|\phi\right\rangle\left\langle\psi\right|\right)\ ,\end{aligned}$$

and therefore

$$\mathrm{Tr}\left(\left|\psi\right\rangle\left\langle\phi\right| + \left|\phi\right\rangle\left\langle\psi\right|\right) = 0\ . \tag{4.9}$$

Substituting $\left|\psi\right\rangle$ and $\left|\phi\right\rangle$ in Eq. (4.9) by $\left|\psi\right\rangle = \sum_i A_i\left|i\right\rangle$ and $\left|\phi\right\rangle = \sum_j B_j\left|j\right\rangle$, we have

$$\begin{aligned}&\sum_{ij} A_i B_j^*\left|i\right\rangle\left\langle j\right| + \sum_{ij} A_i^* B_j\left|i\right\rangle\left\langle j\right| = 0\\ &\Rightarrow \sum_{ij} A_i B_j^* = -\sum_{ij} A_i^* B_j\\ &\therefore \left\langle\psi\,|\,\phi\right\rangle = -\left\langle\phi\,|\,\psi\right\rangle\ ;\end{aligned} \tag{4.10}$$

thus, $\left\langle\psi\,|\,\phi\right\rangle$ is purely imaginary.

The expectation value of Eq. (4.8) after using Eq. (4.10) gives

$$\left\langle\psi\right|\frac{\{\ H, P\ \} - 2\left\langle H\right\rangle P}{2\left(\Delta E\right)^2}\left|\psi\right\rangle - \lambda\left\langle\psi\right|\tilde{H}\left|\psi\right\rangle - \underbrace{\left(\left\langle\phi\,|\,\psi\right\rangle + \left\langle\psi\,|\,\phi\right\rangle\right)}_{0} = 0$$

$$\begin{aligned}\therefore\ \left\langle\psi\right|\tilde{H}\left|\psi\right\rangle &= \frac{1}{2\lambda(\Delta E)^2}\left\langle\psi\,|\,\{\ H, P\ \} - 2\left\langle H\right\rangle P\,|\,\psi\right\rangle\\ &= \frac{1}{2\lambda(\Delta E)^2}\left(\langle H\rangle + \langle H\rangle - \langle H\rangle\right)\\ &= 0\ ;\end{aligned} \tag{4.11}$$

finally, we get

$$\left\langle \tilde{H} \right\rangle = \langle H \rangle - \frac{\mathrm{Tr}H}{n}\mathbf{I} = 0$$

$$\therefore \ \langle H \rangle = \frac{\mathrm{Tr}H}{n}\mathbf{I} \, . \tag{4.12}$$

Applying Eq. (4.8) to $|\psi\rangle$ and after using Eq. (4.10) and $\tilde{H} = H - \langle H \rangle$, we have

$$\frac{\{ H, P \} |\psi\rangle - 2 \langle H \rangle P |\psi\rangle}{2 (\Delta E)^2} - \lambda \tilde{H} |\psi\rangle - (|\psi\rangle \langle \phi | \psi\rangle + |\phi\rangle) = 0$$

$$\begin{aligned} \therefore \ |\phi\rangle &= \frac{\{ H, P \} |\psi\rangle - 2 \langle H \rangle P |\psi\rangle}{2 (\Delta E)^2} - \lambda \tilde{H} |\psi\rangle - \langle \phi | \psi\rangle |\psi\rangle \\ &= \frac{H |\psi\rangle - \langle H \rangle |\psi\rangle}{2 (\Delta E)^2} - \lambda \tilde{H} |\psi\rangle + \langle \psi | \phi\rangle |\psi\rangle \\ &= \left[\left(\frac{1}{2(\Delta E)^2} - \lambda \right) \tilde{H} + \langle \psi | \phi \rangle \right] |\psi\rangle \ . \end{aligned} \tag{4.13}$$

Using $\tilde{H} = H - \mathrm{Tr}H/n$, Eq. (4.8) and Eq. (4.13), we have

$$\begin{aligned} \left\{ \tilde{H}, P \right\} &= (H - \mathrm{Tr}H/n) P + P (H - \mathrm{Tr}H/n) \\ &= \{ H, P \} - 2 \langle H \rangle P \\ &= 2(\Delta E)^2 \left(\lambda \tilde{H} + |\psi\rangle \langle \phi| + |\phi\rangle \langle \psi| \right) \\ &= 2(\Delta E)^2 \left(\lambda \tilde{H} + \left\{ |\psi\rangle \langle \psi|, \left(\frac{1}{2(\Delta E)^2} - \lambda \right) \tilde{H} \right\} \right) \\ &= \left\{ \tilde{H}, P \right\} + 2\lambda (\Delta E)^2 \tilde{H} - 2\lambda (\Delta E)^2 \left\{ \tilde{H}, P \right\} \\ \therefore \ \tilde{H} &= \left\{ \tilde{H}, P \right\} \ . \end{aligned} \tag{4.14}$$

Substituting Eq. (4.13) into Eq. (4.7) and using the properties that $\langle \psi | \phi \rangle$ is constant

in time and $[H, \tilde{H}] = 0$, we obtain

$$\begin{aligned}
&(H - i\frac{\mathrm{d}}{\mathrm{d}t}) \left[\left(\frac{1}{2(\Delta E)^2} - \lambda \right) \tilde{H} + \langle \psi \,|\, \phi \rangle \right] |\psi\rangle + i \left[\frac{\mathrm{d}}{\mathrm{d}t} \left(\frac{\langle H \rangle - H}{2(\Delta E)^2} \right) \right] |\psi\rangle \\
&= \left[\left(\frac{1}{2(\Delta E)^2} - \lambda \right) \tilde{H} + \langle \psi \,|\, \phi \rangle \right] H \,|\psi\rangle - i \left[\frac{\mathrm{d}}{\mathrm{d}t} \left(\frac{1}{2(\Delta E)^2} - \lambda \right) \tilde{H} \right] |\psi\rangle \\
&\quad - i \left[\left(\frac{1}{2(\Delta E)^2} - \lambda \right) \tilde{H} + \langle \psi \,|\, \phi \rangle \right] \frac{\mathrm{d}}{\mathrm{d}t} |\psi\rangle + i \left[\frac{\mathrm{d}}{\mathrm{d}t} \left(\frac{\tilde{H}}{2(\Delta E)^2} \right) \right] |\psi\rangle \\
&= i \left(\frac{\mathrm{d}}{\mathrm{d}t} \lambda \tilde{H} \right) |\psi\rangle = i \left(\frac{\mathrm{d}}{\mathrm{d}t} \lambda \right) \tilde{H} \,|\psi\rangle + i\lambda \left(\frac{\mathrm{d}}{\mathrm{d}t} \tilde{H} \right) |\psi\rangle = 0 \,. \qquad (4.15)
\end{aligned}$$

Then multiplying $\langle\psi|$ to Eq. (4.15), we have

$$\left(\frac{\mathrm{d}}{\mathrm{d}t} \lambda \right) \left\langle \psi \,\middle|\, \tilde{H} \,\middle|\, \psi \right\rangle + \lambda \left\langle \psi \,\middle|\, \left(\frac{\mathrm{d}}{\mathrm{d}t} \tilde{H} \right) \middle|\, \psi \right\rangle = 0 \,. \qquad (4.16)$$

Let's calculate the second term of Eq. (4.16) first. Using

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t} \tilde{H} &= \frac{\mathrm{d}}{\mathrm{d}t} H - \frac{\mathrm{d}}{\mathrm{d}t} \langle \psi | H | \psi \rangle \, \mathrm{I} \\
&= \frac{\mathrm{d}}{\mathrm{d}t} H - \left(\langle \dot{\psi} | H | \psi \rangle + \langle \psi | \dot{H} | \psi \rangle + \langle \psi | H | \dot{\psi} \rangle \right) \mathrm{I} \\
&= \frac{\mathrm{d}}{\mathrm{d}t} H - \left(-i \langle \psi | H^2 | \psi \rangle + \langle \psi | \dot{H} | \psi \rangle + i \langle \psi | H^2 | \psi \rangle \right) \mathrm{I} \\
&= \frac{\mathrm{d}}{\mathrm{d}t} H - \langle \psi | \dot{H} | \psi \rangle \, \mathrm{I} \,, \qquad (4.17)
\end{aligned}$$

the second term of Eq. (4.16) can be written as

$$\left\langle \psi \,\middle|\, \left(\frac{\mathrm{d}}{\mathrm{d}t} \tilde{H} \right) \middle|\, \psi \right\rangle = \left\langle \psi \,\middle|\, \left(\frac{\mathrm{d}}{\mathrm{d}t} H \right) \middle|\, \psi \right\rangle - \langle \psi | \dot{H} | \psi \rangle \, \langle \psi | \psi \rangle \, \mathrm{I} = 0 \,,$$

and therefore Eq. (4.16) can be written as

$$\left(\frac{\mathrm{d}}{\mathrm{d}t} \lambda \right) \left\langle \psi \,\middle|\, \tilde{H} \,\middle|\, \psi \right\rangle = 0 \,. \qquad (4.18)$$

Eq. (4.18) implies that λ is constant in time, and finally, we have

$$\left(\frac{\mathrm{d}}{\mathrm{d}t} \tilde{H} \right) |\psi\rangle = 0 \,. \qquad (4.19)$$

Using Eq. (4.19), we can derive

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\tilde{H} &= \frac{\mathrm{d}}{\mathrm{d}t}\left(\tilde{H}P + P\tilde{H}\right)\\ &= \left(\frac{\mathrm{d}}{\mathrm{d}t}H\right)P + \tilde{H}\left(\frac{\mathrm{d}}{\mathrm{d}t}P\right) + P\left(\frac{\mathrm{d}}{\mathrm{d}t}H\right) + \left(\frac{\mathrm{d}}{\mathrm{d}t}P\right)\tilde{H}\\ &= \tilde{H}\left(|\dot{\psi}\rangle\langle\psi| + |\psi\rangle\langle\dot{\psi}|\right) + \left(|\dot{\psi}\rangle\langle\psi| + |\psi\rangle\langle\dot{\psi}|\right)\tilde{H}\\ &= i\tilde{H}H|\psi\rangle\langle\psi| - i|\psi\rangle\langle\psi|H\tilde{H}\\ &= 0\,, \end{aligned} \tag{4.20}$$

which implies $\tilde{H}$ is constant in time. Let's introduce $|\psi(t)\rangle = \exp\left(-i\int_0^t \langle H\rangle\,\mathrm{dt}\right)|\tilde{\psi}(t)\rangle$, as a result, we have $\tilde{H}|\tilde{\psi}(t)\rangle = i\frac{\partial}{\partial \mathrm{t}}|\tilde{\psi}(t)\rangle$. Now, the Hamiltonian $\tilde{H}$ could be written as

$$\begin{aligned}\tilde{H} &= \tilde{H}P + P\tilde{H}\\ &= \tilde{H}\tilde{P} + \tilde{P}\tilde{H}\\ &= i\left(|\dot{\tilde{\psi}}\rangle\langle\tilde{\psi}| - |\tilde{\psi}\rangle\langle\dot{\tilde{\psi}}|\right)\,. \end{aligned} \tag{4.21}$$

Because $\langle\tilde{H}\rangle = \langle\tilde{\psi}|\tilde{H}|\tilde{\psi}\rangle = \langle\tilde{\psi}|\frac{\partial}{\partial t}|\tilde{\psi}\rangle = 0$, the derivative $|\frac{\partial}{\partial t}|\tilde{\psi}\rangle$ is orthogonal to $|\tilde{\psi}\rangle$. Since $\frac{\mathrm{d}\tilde{H}}{\mathrm{d}t} = 0$, using Eq. (4.21), we have

$$\begin{aligned}\left(\frac{\mathrm{d}\tilde{H}}{\mathrm{d}t}\right)|\tilde{\psi}\rangle &= i\left(|\ddot{\tilde{\psi}}\rangle\langle\tilde{\psi}| + |\dot{\tilde{\psi}}\rangle\langle\dot{\tilde{\psi}}| - |\dot{\tilde{\psi}}\rangle\langle\dot{\tilde{\psi}}| - |\tilde{\psi}\rangle\langle\ddot{\tilde{\psi}}|\right)|\tilde{\psi}\rangle\\ &= i\left(\frac{\mathrm{d}^2}{\mathrm{d}t^2}|\tilde{\psi}\rangle - |\tilde{\psi}\rangle\langle\ddot{\tilde{\psi}}|\tilde{\psi}\rangle\right)\\ &= i\left(\frac{\mathrm{d}^2}{\mathrm{d}t^2}|\tilde{\psi}\rangle - |\tilde{\psi}\rangle\langle\tilde{\psi}|\ddot{\tilde{\psi}}\rangle\right)\\ &\therefore\quad \left(1-\tilde{P}\right)\frac{\mathrm{d}^2}{\mathrm{d}t^2}|\tilde{\psi}\rangle = 0\,, \end{aligned} \tag{4.22}$$

where

$$\begin{aligned}\langle\ddot{\tilde{\psi}}|\tilde{\psi}\rangle &= -i\left(\frac{\mathrm{d}}{\mathrm{d}t}\langle\tilde{\psi}|\,\tilde{H}\right)|\tilde{\psi}\rangle \\ &= -i\left(\frac{\mathrm{d}}{\mathrm{d}t}\langle\tilde{\psi}|\right)\tilde{H}\,|\tilde{\psi}\rangle \\ &= -\langle\tilde{\psi}|\tilde{H}^2|\tilde{\psi}\rangle = \langle\tilde{\psi}|\ddot{\tilde{\psi}}\rangle\ .\end{aligned}$$

Eq. (4.22) is the geodesic equation for the Fubini-Study metric on $\mathbb{C}P^{n-1}$. The solution to Eq. (4.22) can be written as

$$|\tilde{\psi}(t)\rangle = \cos\omega t\,|\tilde{\psi}(0)\rangle + \sin\omega t\,|\tilde{\psi}_\perp(0)\rangle\ , \tag{4.23}$$

where $|\tilde{\psi}(0)\rangle$ is the initial state - $|\tilde{\psi}_i\rangle$, and $|\tilde{\psi}_\perp(0)\rangle = \frac{1}{\omega}\,|\frac{\mathrm{d}}{\mathrm{d}t}\tilde{\psi}(0)\rangle$ is a chosen state which is normalized and orthogonal to $|\tilde{\psi}(0)\rangle$. Since $|\tilde{\psi}_\perp(0)\rangle$ can be chosen, it's reasonable to suppose that our final state at time T - $|\tilde{\psi}_f(T)\rangle$ is spanned by $|\tilde{\psi}(0)\rangle$ and $|\tilde{\psi}_\perp(0)\rangle$; that is

$$|\tilde{\psi}_f(T)\rangle = \alpha\,|\tilde{\psi}(0)\rangle + \sqrt{1-\alpha^2}\,|\tilde{\psi}_\perp(0)\rangle\ . \tag{4.24}$$

Comparing the coefficients between Eq. (4.23) and Eq. (4.24), we find $\alpha = \langle\tilde{\psi}_i|\tilde{\psi}_f\rangle$ and the optimal time - $T = \frac{1}{\omega}\cos^{-1}|\,\langle\tilde{\psi}_i|\tilde{\psi}_f\rangle\,|$. The whole Hamiltonian is given by $H(t) = \tilde{H} + \langle H(t)\rangle$, where $\langle H(t)\rangle$ is an arbitrary real function corresponding to the degree of freedom.

4.2 Time-optimal realization of unitary operators

The realization of time-optimal unitary operation using the variational principle approach can be extended very naturally from the previous section. Carlini et al. replace the projective space representing quantum state vectors with the space of unitary operators. The known initial condition and final condition here are the identity matrix,

and the final unitary operator which replace the initial state and the final state in the previous section. As a result, this approach is more useful in quantum computation when the input may be unknown. The derivation of the equations of motion for the unitary operation case is similar to that for the quantum state case. Therefore , we will roughly derive the theory, and discuses the difficulty that the theory may encounter.

Now, one would like to find a time dependent Hamiltonian such that the unitary operator evolves from identity matrix, I, to a given unitary operator U_f which satisfies the Schröding equation ,$i\frac{\mathrm{d}}{\mathrm{d}t}U(t) = H(t)U(t)$, and belongs to a SU($n$) (modulo overall phases) in the least time. Here, the unitary operator $U(t)$ and $H(t)$ are dynamical variables, and the action and Lagrange can be defined as

$$S(U, \dot{U}, H, \dot{H}, \Lambda, \dot{\Lambda}, \lambda_j, \dot{\lambda_j}) = \int \mathrm{d}t L_T + L_S + L_C \ , \tag{4.25}$$

$$L_T = \sqrt{\frac{\langle \frac{\mathrm{d}}{\mathrm{d}t}U, (1-P_U)\frac{\mathrm{d}}{\mathrm{d}t}U\rangle}{\langle HU, (1-P_U)HU\rangle}} \ , \tag{4.26}$$

$$L_S = \left\langle \Lambda, i\frac{\mathrm{d}U}{\mathrm{d}t}U^\dagger - H \right\rangle \ , \tag{4.27}$$

$$L_C = \sum_j \lambda_j f_j(H) \ , \tag{4.28}$$

where the Hilbert-Schmidt norm $\langle A, B\rangle = \mathrm{Tr}(A^\dagger B)$ has been introduced, and the projection operator is defined as $P_U A = \frac{1}{N}\mathrm{Tr}(AU^\dagger)U$. The Hermitian operator $\Lambda(t)$ and scalars $\lambda_j(t)$ are Lagrange multipliers.

Since the derivation is similar to the quantum state version, we will merely go through the important equations that we will use to implement controlled-Z gate in the next subsection. The variation of L_S respect to Λ will give us the Schrödinger Equation,

$$i\frac{\mathrm{d}}{\mathrm{d}t}U(t) = H(t)U(t) \ . \tag{4.29}$$

The variation of action respect to H will give us

$$-\frac{(1-P_1)H}{\langle H,(1-P_1)H\rangle}-\Lambda+F=0\ , \tag{4.30}$$

where $F=\frac{\partial L_C}{\partial H}$ and $P_1A=\frac{1}{N}(\mathrm{Tr}A)I$. Note that $P_U(A)=(P_1AU^\dagger)U$. With the condition $(1-P_1)(H)=\tilde{H}$, Eq. (4.30) can be rewritten as

$$\Lambda=F-\frac{\tilde{H}}{\mathrm{Tr}\tilde{H}^2}\ . \tag{4.31}$$

Performing the variation of action respect to U and after some trivial calculations, we get

$$D\left[L_T\frac{(1-P_1)(\frac{\mathrm{d}U}{\mathrm{d}t}U^\dagger)}{\left\langle\frac{\mathrm{d}U}{\mathrm{d}t}U^\dagger,(1-P_1)\left(\frac{\mathrm{d}U}{\mathrm{d}t}U^\dagger\right)\right\rangle}+i\Lambda\right]=0\ , \tag{4.32}$$

where $D[A]=\frac{\mathrm{d}}{\mathrm{d}t}A+\left[A,\frac{\mathrm{d}U}{\mathrm{d}t}U^\dagger\right]$. Using Eq. (4.29), Eq. (4.31), and Eq. (4.32), we can derive the quantum brachistochrone equation as

$$i\frac{\mathrm{d}F}{\mathrm{d}t}=[H,F]\ . \tag{4.33}$$

Assume that we have another constrain, the finite constant energy condition, which can be written as

$$f(H)=\frac{1}{2}(\mathrm{Tr}\tilde{H}^2-N\omega^2)\ , \tag{4.34}$$

where ω is a constant. The constraint part of the Lagrangian L_C will have an extra term, and can be written as

$$L_C=\lambda f(H)+L'_C\ , \tag{4.35}$$

where λ is a Lagrange multiplier and L'_C is sum of the other constraints. Now, $F=\lambda\tilde{H}+F'$ where $F=\frac{\partial L'_C}{\partial H}$. If the constrains of the system are linear and homogeneous in $\tilde{H}$,

$$L'_C=\mathrm{Tr}\tilde{H}F'\ , \tag{4.36}$$

where $F' = \sum_j \lambda_j g_j$ with $g_j \in SU(N)$, we have

$$\mathrm{Tr} g_j \tilde{H} = 0 \, . \tag{4.37}$$

With the above conditions, Eq. (4.33) can be written as

$$\frac{\mathrm{d}}{\mathrm{d}t}\tilde{H} + \sum_j g_j \frac{\mathrm{d}\lambda_j}{\mathrm{d}t} = -i \sum_j \lambda_j [\tilde{H}, g_j] \, , \tag{4.38}$$

where the Lagrange multipliers $\lambda_j(t)$ and the Hamiltonian are still to be determined. Once the differential equations, Eq. (4.38), have been solved, we can compare the coefficients between the desired unitary operation and the unitary operation constructed from the solution of Eq. (4.38) to find the control sequences.

4.3 Compare the Controlled Z gate implemented by the time-optimal approach and the canonical decomposition respectively

The time-optimal approach using the variational principle has more geometric and physical meaning than the others, and the optimal time can be obtained from solving the differential equations. But in the differential equations of the motion, Eq. (4.38), the only known conditions are the initial unitary operator, $U(0) = I$, and the final $U(T) = U_f$, and we need to use this conditions to find the time dependent parameters, $H(t)$, $U(t)$, and the Lagrange multipliers $\lambda_j(t)$. The typical numerical solvable differential equation is that we know all the initial values of the dynamic parameters, and thus we can numerically obtain the dynamic values at given time t. In the case that we deal now, we don't know the initial values of the Lagrange multipliers, $\lambda_j(t)$, and what's worse, we do not know the time, T, at which the unitary operator will evolve to $U_f(T)$.

It seems that the only possible method to numerically solve the atypical differential equations is to guess the initial values of $\lambda_j(t)$ and the duration of the operation time, T, then put into the differential equation to see if the guessing is correct. It turns out that the numerical solution may be very difficult. On the other hand, if we can analytically solve the differential equations, Eq. (4.38), we may compare the solution $U(t)$ of the given Hamiltonian with the desired gate operation to find out the dynamic values and the optimal time. But it may still be a problem that the coupled differential equations are often very complex and can not be solved analytically at most of time.

We have tried to solve the time-opmtial equations of motion of the Kane's system, but it's too complex to find the analytical solution. As a result, we try to use the system Hamiltonian which had been analytically solved [10]. The Hamiltonian has this form:

$$H(t) = -\sum_j J_j(t)\sigma_j \otimes \sigma_j + B_1(t)\sigma_z \otimes I + B_2(t) I \otimes \sigma_z \ , \tag{4.39}$$

which contains controllable anisotopic couplings and controllable local terms. Since the generator of the system does not constitute a minimal generating set of the Lie algebra SU(4), the system is not fully controllable; that is, we can not implement all of the unitary operation in SU(4). Meanwhile, we have tried to add another terms such that the system constitute a minimal generating set of the Lie algebra SU(4), but it turns out that the equations are hard to solve analytically. This system can implement the controlled-Z gate, and the controlled-Z gate is one of the most important unitary operations for implementing quantum algorithms. In particular, we can implement the CNOT gate from a controlled-Z gate conjugated by $I \otimes H$ which is schematically shown in Fig 4.1. Therefore, we will implement the controlled-Z gate using the time-optimal variational principle approach in the following subsection 4.3.1, and using the canonical decomposition approach in subsection 4.3.2. Finally, we will compare

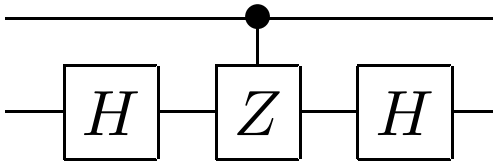

Figure 4.1 Circuit diagram of the CNOT gate constructed by controlled Z gate and Hadamard gate.

the results using these two approaches and give brief conclusion in subsection 4.3.3.

4.3.1 The time-optimal of the controlled Z gate using variational principle approach

The Hamiltonian in Eq. (4.39) can be written in computational basis as

$$H(t) = \begin{pmatrix} -J_z + B_+ & 0 & 0 & -J_- \\ 0 & J_z + B_- & -J_+ & 0 \\ 0 & -J_+ & J_z - B_- & 0 \\ -J_- & 0 & 0 & -J_z - B_+ \end{pmatrix}, \tag{4.40}$$

where $B_\pm = B_1(t) \pm B_2(t)$ and $J_\pm(t) = J_x(t) \pm J_y(t)$. From Eqs. (4.36) and (4.37), the physical Hamiltonian is guaranteed by

$$F' = \sum_{j \neq k} \lambda_{jk} \sigma_j \otimes \sigma_k + \sum_{j=x,y} \lambda_j^1 \sigma_j \otimes I + \lambda_j^2 I \otimes \sigma_j \,, \tag{4.41}$$

where the λ_{jk} and λ_j^i are the Lagramge multipliers. As a result, we can write down all the g_j as

$$\begin{aligned} g_1 = \sigma_1 \otimes \sigma_2 \ , \qquad & g_2 = \sigma_1 \otimes \sigma_3 \ , \qquad g_3 = \sigma_2 \otimes \sigma_1 \ , \\ g_4 = \sigma_2 \otimes \sigma_3 \ , \qquad & g_5 = \sigma_3 \otimes \sigma_1 \ , \qquad g_6 = \sigma_3 \otimes \sigma_2 \ , \\ & g_7 = \sigma_x \otimes I \ , \qquad g_8 = \sigma_y \otimes I \ , \\ & g_9 = I \otimes \sigma_x \ , \qquad g_{10} = I \otimes \sigma_y \ . \end{aligned} \tag{4.42}$$

Inserting Eq. (4.42) and Eq. (4.40) into Eq. (4.38) and comparing the coefficients of the generators of SU(4) on both sides, we find that λ_{xy}, λ_{yz}, and J_z are constant. Then $B_\pm$ and $J_\pm$ can be solved, and the solution is

$$\begin{aligned} B_\pm(t) &= B_{0\pm} \cos 2(\gamma_\pm t + \psi_\pm) \ , \\ J_\pm(t) &= \mp B_{0\mp} \sin 2(\gamma_\mp t + \psi_\mp) \ , \end{aligned} \tag{4.43}$$

where $\gamma_\pm = \lambda_{xy} \pm \lambda_{yx}$, $B_{0\pm}$ and $\psi_\pm$ are constant. Now, we have the time-depentant Hamiltonian, and therefore we can solve the Schrödinger equation, Eq. (4.29), to obtain the evolution of the unitary operator. After solving the Schrödinger equation, the optimal unitary operator of the system will be given by

$$U(t) = \begin{pmatrix} (\alpha_{0+} + i\alpha_{z+})e^{iJ_z t} & 0 & 0 & (\alpha_{y+} + i\alpha_{x+})e^{iJ_z t} \\ 0 & (\alpha_{0-} + i\alpha_{z-})e^{-iJ_z t} & (\alpha_{y-} + i\alpha_{x-})e^{-iJ_z t} & 0 \\ 0 & (-\alpha_{y-} + i\alpha_{x-})e^{-iJ_z t} & (\alpha_{0-} - i\alpha_{z-})e^{-iJ_z t} & 0 \\ (-\alpha_{y+} + i\alpha_{x+})e^{iJ_z t} & 0 & 0 & (\alpha_{0+} - i\alpha_{z+})e^{iJ_z t} \end{pmatrix} , \tag{4.44}$$

where $U(0) = 1$ is chosen, and

$$\Omega_{\pm} = \sqrt{B_{0\pm}^2 + \gamma_{\pm}^2} \,, \tag{4.45}$$

$$\alpha_{0\pm}(t) = \cos\gamma_{\pm}t\cos\Omega_{\pm}t + \frac{\gamma_{\pm}}{\Omega_{\pm}}\sin\gamma_{\pm}t\sin\Omega_{\pm}t \,, \tag{4.46}$$

$$\alpha_{z\pm}(t) = -\frac{B_{0\pm}}{\Omega_{\pm}}\sin\Omega_{\pm}t\cos(\gamma_{\pm}t + 2\psi_{\pm}) \,, \tag{4.47}$$

$$\alpha_{x\pm}(t) = \pm\frac{B_{0\pm}}{\Omega_{\pm}}\sin\Omega_{\pm}t\sin(\gamma_{\pm}t + 2\psi_{\pm}) \,, \tag{4.48}$$

$$\alpha_{y\pm}(t) = \pm\left(\sin\gamma_{\pm}t\cos\Omega_{\pm}t - \frac{\gamma_{\pm}}{\Omega_{\pm}}\cos\gamma_{\pm}t\sin\Omega_{\pm}t\right) \,. \tag{4.49}$$

The controlled Z gate is defined in the computational basis by

$$U_{\Lambda Z} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \,, \tag{4.50}$$

comparing the coefficients between Eq. (4.49) and Eq. (4.50), then we find that $\alpha_{x\pm} = \alpha_{y\pm} = 0$ giving $\gamma_{\pm} = \psi_{\pm} = 0$ which imply from Eq. (4.43) that $B_{\pm}$ are constant of time and $J_{\pm}$ are equal to zero. Eq. (4.44) can thus be written as

$$U(T) = \begin{pmatrix} e^{i(J_z - B_{0+})T} & 0 & 0 & 0 \\ 0 & e^{-i(J_z + B_{0-})T} & 0 & 0 \\ 0 & 0 & e^{-i(J_z - B_{0-})T} & 0 \\ 0 & 0 & 0 & e^{i(J_z + B_{0+})T} \end{pmatrix} \,. \tag{4.51}$$

Supposing that at given time, T, $U(T)$ is equal to $U_{\Lambda Z}$ up to a global phase - $U(T) = e^{ix}U_{\Lambda Z}$, we obtain the set of relation of parameters

$$(J_z - B_{0+})T = x \,, \tag{4.52}$$

$$-(J_z + B_{0-})T = x + 2n\pi \,, \tag{4.53}$$

$$-(J_z - B_{0-})T = x + 2p\pi \,, \tag{4.54}$$

$$(J_z + B_{0+})T = x + (2q+1)\pi \,, \tag{4.55}$$

where n, p, and q are arbitrary integers and T is still to be determined. Solving Eq. (4.52) to Eq. (4.55) , we obtain

$$J_z T = \frac{1}{2}(q - n - p + \frac{1}{2})\pi \ , \tag{4.56}$$

$$x = -\frac{1}{2}(n + p - q - \frac{1}{2})\pi \ , \tag{4.57}$$

$$B_+ T = (q + \frac{1}{2})\pi \ , \tag{4.58}$$

$$B_- T = (p - n)\pi \ . \tag{4.59}$$

Since $B_+ = B_1 + B_2$ and $B_- = B_1 - B_2$, Eq. (4.58) and Eq. (4.59) can be written as

$$\begin{aligned} T &= \frac{1}{2B_1}(p + q - n + \frac{1}{2})\pi \\ T &= \frac{1}{2B_2}(q - p + n + \frac{1}{2})\pi \ . \end{aligned} \tag{4.60}$$

The time-optimal duration $T_{\Lambda Z}$ can then be found by minimizing Eq. (4.60). We find the solution is $p = q = n = 0$. Using Eq. (4.56), we then have $T = \frac{\pi}{4B}$ and $J_z = B$, where $B_1 = B_2$.

The Hamiltonian of the time-optimal evolution can be writes as

$$H = -B\sigma_z \otimes \sigma_z + B\sigma_z \otimes I + BI \otimes \sigma_z \ , \tag{4.61}$$

where the value of B is determined by the real physical system. Since the Hamiltonian is time independent, the evolution of the unitary operator, Eq. (4.50), will be along a geodesic curves on the $SU(4)$ manifold endowed with the metric $\mathrm{d}s_U^2$.

4.3.2 The canonical decomposition of the controlled Z gate

The canonical decomposition [26,27] decomposes any two qubit unitary operator into a product of four single qubit unitaries and one entangling unitary -

$$U = (V_1 \otimes V_2) U_{\mathrm{can}} (V_3 \otimes V_4) \ , \tag{4.62}$$

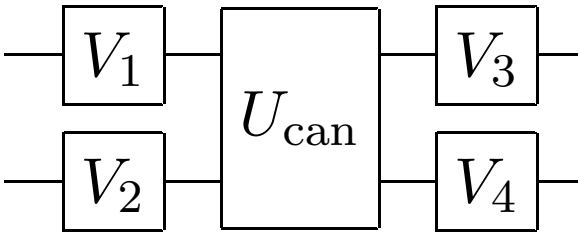

Figure 4.2 Circuit diagram of the canonical decomposition.

where V_1, V_2, V_3, and V_4 are single qubit unitaries, and U_{can} is the two qubit interaction. The schematic diagram has been shown in Fig. 4.2. The U_{can} only involves three parameters, θ_x, θ_y, and θ_z:

$$U_{\text{can}} = e^{i\theta_x\sigma_x\otimes\sigma_x} e^{i\theta_y\sigma_y\otimes\sigma_y} e^{i\theta_z\sigma_z\otimes\sigma_z} \ , \tag{4.63}$$

where the purely non-local terms - $e^{i\theta_x\sigma_x\otimes\sigma_x}$, $e^{i\theta_y\sigma_y\otimes\sigma_y}$, and $e^{i\theta_z\sigma_z\otimes\sigma_z}$ are known as the interaction contents of the gate, and it's not difficult to show that each of the terms commute with each other.

In principle, each of the terms - $e^{i\theta_x\sigma_x\otimes\sigma_x}$, $e^{i\theta_y\sigma_y\otimes\sigma_y}$, and $e^{i\theta_z\sigma_z\otimes\sigma_z}$ can be interpreted physically as a type of controlled rotation. For example [28],

$$\begin{aligned} e^{i\theta_z\sigma_z\otimes\sigma_z} &= \cos\theta_z I\otimes I + i\sin\theta_z\sigma_z\otimes\sigma_z \\ &= \cos\theta_z(|0\rangle\langle 0| + |1\rangle\langle 1|)\otimes I + i\sin\theta_z(|0\rangle\langle 0| - |1\rangle\langle 1|)\otimes\sigma_z \\ &= |0\rangle\langle 0|\otimes e^{i\theta_z\sigma_z} + |1\rangle\langle 1|\otimes e^{-i\theta_z\sigma_z} \\ &= \left(I\otimes e^{i\theta_z\sigma_z}\right)\left(|0\rangle\langle 0|\otimes I + |1\rangle\langle 1|\otimes e^{-i2\theta_z\sigma_z}\right) \end{aligned}$$

$$\begin{aligned} \therefore\ \left(I\otimes e^{-i\theta_z\sigma_z}\right) e^{i\theta_z\sigma_z\otimes\sigma_z} &= |0\rangle\langle 0|\otimes I + |1\rangle\langle 1|\otimes e^{-i2\theta_z\sigma_z} \\ &= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & e^{-i2\theta_z} & 0 \\ 0 & 0 & 0 & e^{i2\theta_z} \end{pmatrix} . \end{aligned} \tag{4.64}$$

As a result, $\left(I \otimes e^{-i\theta_z\sigma_z}\right) e^{i\theta_z\sigma_z\otimes\sigma_z}$ is corresponding to an operation consisting of both a controlled-Z and a phase rotation. Similarly, the other two terms have the same physically meaning -

$$\left(I \otimes e^{-i\theta_x\sigma_x}\right)e^{i\theta_x\sigma_x\otimes\sigma_x} = |x_+\rangle\langle x_+| \otimes I + |x_-\rangle\langle x_-| \otimes e^{-i2\theta_x\sigma_x} , \tag{4.65}$$

$$\left(I \otimes e^{-i\theta_y\sigma_y}\right)e^{i\theta_y\sigma_y\otimes\sigma_y} = |y_+\rangle\langle y_+| \otimes I + |y_-\rangle\langle y_-| \otimes e^{-i2\theta_y\sigma_y} . \tag{4.66}$$

Since the controlled-Z rotation of angle θ is defined in the computational basis by

$$U_{\Lambda Z}(\theta) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & e^{i\theta} \end{pmatrix} , \tag{4.67}$$

we can define a controlled phase rotation -

$$\begin{aligned} U_{\Lambda\text{phase}}(\theta) &= |0\rangle\langle 0| \otimes I + |1\rangle\langle 1| \otimes e^{i\theta I} \\ &= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & e^{i\theta} & 0 \\ 0 & 0 & 0 & e^{i\theta} \end{pmatrix} , \end{aligned} \tag{4.68}$$

together with Eq. (4.64) to construct controlled-Z rotation. Now controlled Z rotation could be decomposed as

$$\begin{aligned} U_{\Lambda Z}(\theta) &= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & e^{i\frac{1}{2}\theta} & 0 \\ 0 & 0 & 0 & e^{i\frac{1}{2}\theta} \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & e^{-i\frac{1}{2}\theta} & 0 \\ 0 & 0 & 0 & e^{i\frac{1}{2}\theta} \end{pmatrix} \\ &= \left(|0\rangle\langle 0| \otimes I + |1\rangle\langle 1| \otimes e^{i\frac{\theta}{2}I}\right)\left(I \otimes e^{-i\frac{\theta}{4}\sigma_z}\right) e^{i\frac{\theta}{4}\sigma_z\otimes\sigma_z} . \end{aligned} \tag{4.69}$$

In the case of controlled Z gate, $\theta = \pi$, we have

$$\begin{aligned} U_{\Lambda Z} &= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \\ &= \left(|0\rangle\langle 0| \otimes I + |1\rangle\langle 1| \otimes e^{i\frac{\pi}{2}I}\right)\left(I \otimes e^{-i\frac{\pi}{4}\sigma_z}\right) e^{i\frac{\pi}{4}\sigma_z\otimes\sigma_z} \\ &= \left(\begin{pmatrix} 1 & 0 \\ 0 & i \end{pmatrix} \otimes \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right)\left(I \otimes R_z\left(\frac{\pi}{2}\right)\right) e^{i\frac{\pi}{4}\sigma_z\otimes\sigma_z} \\ &= \left(S \otimes R_z\left(\frac{\pi}{2}\right)\right) e^{i\frac{\pi}{4}\sigma_z\otimes\sigma_z} , \end{aligned} \tag{4.70}$$

where $R_z(\theta)$ is rotation operator around $\hat{z}$ axis, defined by

$$R_z(\theta) = e^{-i\frac{\theta}{2}\sigma_z} = \cos\frac{\theta}{2}I - i\sin\frac{\theta}{2}\sigma_z , \tag{4.71}$$

and S is the phase gate which can be constructed by

$$S = \begin{pmatrix} 1 & 0 \\ 0 & i \end{pmatrix} = e^{i\frac{\pi}{4}} \begin{pmatrix} e^{-i\frac{\pi}{4}} & 0 \\ 0 & e^{i\frac{\pi}{4}} \end{pmatrix} = e^{i\frac{\pi}{4}} R_z(\frac{\pi}{2}) . \tag{4.72}$$

Therefore, Eq. (4.70) can be write as

$$\begin{aligned} U_{\Lambda Z} &= \left(e^{i\frac{\pi}{4}I} \otimes I\right)\left(R_z\left(\frac{\pi}{2}\right) \otimes R_z\left(\frac{\pi}{2}\right)\right) e^{i\frac{\pi}{4}\sigma_z\otimes\sigma_z} \\ &= \left(e^{-i\frac{\pi}{4}I} \otimes I\right)\left(R_z\left(-\frac{\pi}{2}\right) \otimes R_z\left(-\frac{\pi}{2}\right)\right) e^{-i\frac{\pi}{4}\sigma_z\otimes\sigma_z} , \end{aligned} \tag{4.73}$$

where $U_{\Lambda Z} = U_{\Lambda Z}^{\dagger}$, and $e^{-i\frac{\pi}{4}I} \otimes I$ can be regarded as global phase which can be ignored. The controlled Z gate is schematically shown in Fig. 4.3.

For purpose of comparing with the result of pervious subsection 4.3.1, suppose we have a system with the following Hamiltonian,

$$H = -J_z\sigma_z \otimes \sigma_z + B_1\sigma_z \otimes I + B_2 I \otimes \sigma_z , \tag{4.74}$$

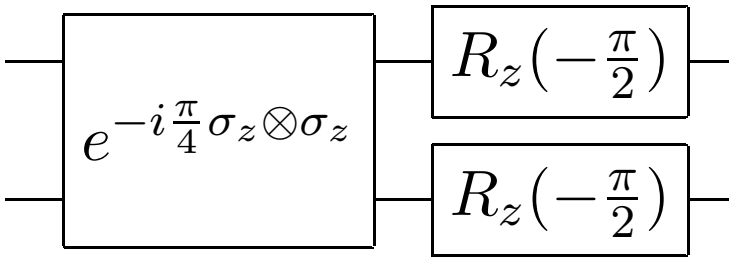

Figure 4.3 Circuit diagram of the controlled Z gate.

where the coupling strength J_z, B_1, and B_2 are controllable. For example, the single-qubit Z rotation around $\hat{z}$ axis of qubit 1 can be constructed by turning B_1 on and keeping J_z and B_2 off, so $H = B_1\sigma_z \otimes I$, and

$$R_z(\theta) = e^{-i\frac{\theta}{2}\sigma_z\otimes I} = e^{-i(2n\pi - B_1 t)\sigma_z\otimes I} \ . \tag{4.75}$$

We then find the duration of Z rotation is $t = 2n\pi - \frac{\theta}{2}$, where n is an arbitrary integer. The unitary gate $e^{-i\theta\sigma_z\otimes\sigma_z}$ can be easily implemented by turning on $H = -J_z\sigma_z \otimes \sigma_z$ and other terms off. We thus can find the duration is given by $t = \frac{\theta}{J_z}$. The two single qubit operations $R_z\left(-\frac{\pi}{2}\right) \otimes I$ and $I \otimes R_z\left(-\frac{\pi}{2}\right)$ could be computed in parallel. As a result, the duration of controlled Z gate is $T = \frac{\pi}{4B} + \frac{\pi}{4J}$.

4.3.3 Conclusion

In the time-optimal variational principle approach, we find that the total operation time of a controlled-Z gate is $T = \frac{\pi}{4B}$. On the other hand, we find the total operation time using the canonical decomposition is $T = \frac{\pi}{4B} + \frac{\pi}{4J}$. The variational principle approach can find the optimal time very naturally and guaranteed that it's the least time. In this case, variational principle gives us the one step unitary operation. We can turn on the coupling term J_z, B_1, B_2 at the same time, and set their strength equal to each other,$J_z = B_1 = B_2 = B$. In the canonical decomposition, we have two steps. In the first step, we perform the coupling interaction. In the second

step, we perform the single-qubit Z-rotations. The canonical decomposition approach has one extra step compared with the variational principle approach. Luckily, since $[e^{i\theta Z\otimes Z}, R_z(\theta)\otimes I] = 0$, we can artificially combine the two steps into one step, i.e. doing them altogether in parallel. With the choice of $J_z = B$, the controlled-Z operation is exactly the same as the result of time-opmtial variational principle approach. Although we can deliberately obtain the same result, the time-optimal variational principle approach is more powerful. It guarantees that the evolution this approach finds is the shortest path in time, and it's not merely coincident.

The authors of the paper [10] point out that it's possible and no conceptual difficulty in extending the variational methods to the more realistic case that we have mention in Chapter 1 when similar constraints are given in terms of inequalities instead of that the constraints are expressed as equality conditions. We have derive the equations of the motion with the inequality constrains following the same concept, but it turns out that the equations are too complex to find analytical solution.

Possible further work is to develop a numerical technique to the quantum brachistochrone equations. We can then use the variation methods to find time-optimal gate operations numerically for more complex system which can not be analytically solved. The application of this method is we can evaluate the time-dependent Hamiltonian and the evolution for a given final unitary operator by using a classical computer, and perform this control sequence in the quantum experiment and computation.

Chapter 5

Optimal CNOT Gate

5.1 Control sequence obtained from reduced Hamiltonian

In this chapter, we investigate and find the digitized control sequence of the near time-optimal, high-fidelity CNOT gate for the Silicon-based electron spin quantum computing architecture discussed in Chapter 3. Since the rotating magnetic field is always on in this scheme, electrons will undergo a rotation around the x-axis when there are no voltages applied on A gates, i.e. $\Delta\omega = 0$ with an angular frequency of $\Omega_0 = g_e\mu_B B_{ac}/\hbar$. While target electrons will perform particular unitary operation within time t, every spectator qubit will rotate around the x-axis with an angle of

$$\theta_x = 2\pi - \Omega_0 t \ . \tag{5.1}$$

If θ_x does not equal to $2n\pi$, where n is integral, another correction step will be required for the spectator qubits. Therefore, it will be more convenient to choose the

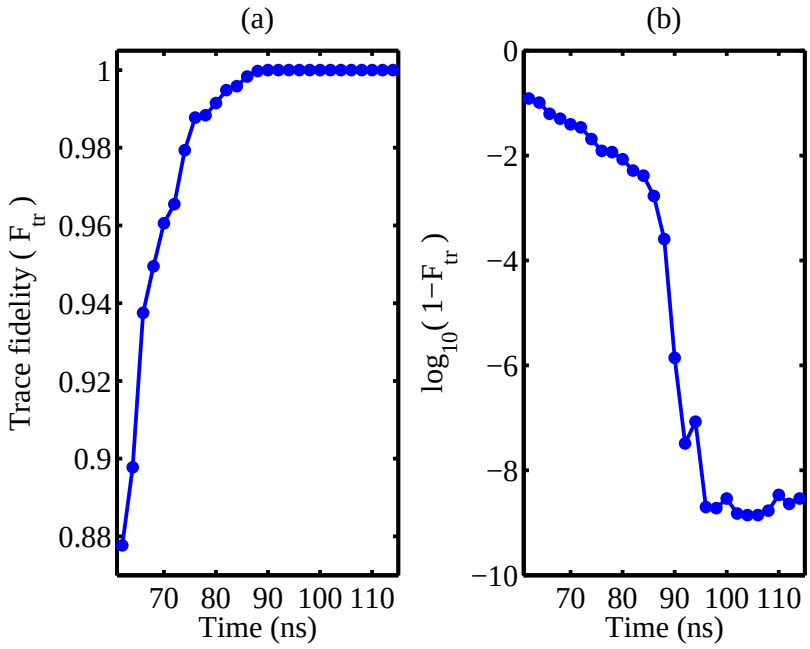

Figure 5.1 Fidelity versus time for implementing CNOT gate. (a) Gives the trace fidelities against time, while (b) shows deviation $\log_{10}(1 - F_{tr})$ from fidelities.

operation time,

$$\begin{aligned} t &= 2n\pi/\Omega_0 \\ &= \frac{2n\hbar\pi}{g_e\mu_B B_{ac}} \end{aligned} \tag{5.2}$$

such that there is no need for correction for spectator qubits. For a chosen time t, a larger n will require a larger B_{ac} field . Yet in the approximation of the reduced Hamiltonian, we assume that the B_{ac} filed is very small compared with dc magnetic field B_0. As a result, we chose $n = 1$ for higher fidelity in full Hamiltonian simulation.

We focus on the high fidelity controlled-NOT (CNOT) gate by optimizing the effective, reduced donor electron spin Hamiltonian, with external controls over the hyperfine A and exchange J interactions. We first try different piecewise constant control steps and numerically calculate the fidelity (error) against the time needed to implement a CNOT gate with stopping criteria of error in the optimizer set to

10^{-9} in order to economize the simulation time. Here, the error is defined as $1 - F$, where F is fidelity. For each trying value of time t, we divide the sequence into 30 piecewise steps, starting with initial control amplitudes u_{kj} by assigning a random value to every five point and using cubic spline to fill in the amplitudes u_{kj} of the intermediate time points. The values of the control amplitudes are constrained to be within certain range in the parameter space. Fig. 5.1 shows the fidelity against time. As shown in Fig. 5.1 (a), the error is less than 10^{-8} for times longer than $100ns$, and it is found that 30 piecewise constant control steps for the CNOT gate operation will be sufficient to meet the required fidelity (error), and the performance would not be improved further with more steps. With operation time $t = 100$ns and stopping criteria of error set to 10^{-16}, we can find that the near time-optimal, high-fidelity CNOT gate control sequence has an error of 1.11×10^{-16}. The sequence of controls is schematically shown in Fig. 5.2.

5.2 Full Hamiltonian simulation

In this section, we simulate the control sequence of the CNOT gate for the full spin Hamiltonian simulations, Eq. (2.16), and the control sequence is obtained from the reduced Hamiltonian, Eq. (2.15), optimization. Define $\epsilon = \frac{1}{2}g_e\mu_B B_0 + \frac{1}{2}g_n\mu_n B_0$, and solve Eq. (2.8), then we will have

$$A = \frac{-\epsilon + \sqrt{\epsilon^2 + 4\left[A_0^2 + \epsilon\left(\frac{\hbar}{2}\Delta\omega + A_0\right)\right]}}{2} . \tag{5.3}$$

Using the full Hamiltonian, Eq. (2.16), and Eq. (5.3), we simulate the CNOT gate numerically with initialized spin-up nuclear states and four different electron spin basis states, $|00\rangle_e$, $|01\rangle_e$, $|10\rangle_e$, and $|11\rangle_e$ where $|0\rangle_e$ means the electron spin is up. The final reduced electron density matrix is defined as the composite density matrix

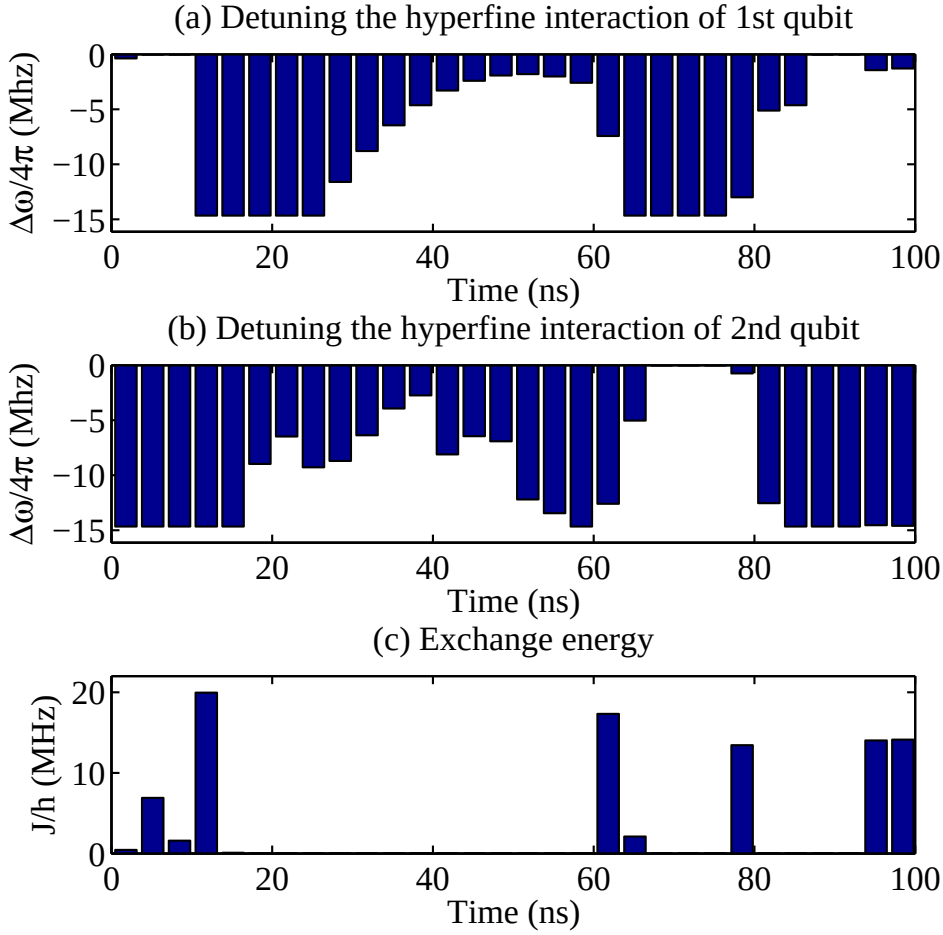

Figure 5.2 Time optimal CNOT gate control sequence with 30 steps during $100ns$. In the reduced Hamiltonian, we assume that the nuclei will always be in spin-up state and the dynamics of nuclei have been frozen out. In (a) and (b), the maximum energy difference of σ_z term from detuning the hyperfine interaction is $(1/2)\hbar\Delta\omega/h = \Delta\omega/4\pi = -14.7$MHz. In (c), the maximum electron-electron exchange energy is $J/h = 20$MHz.

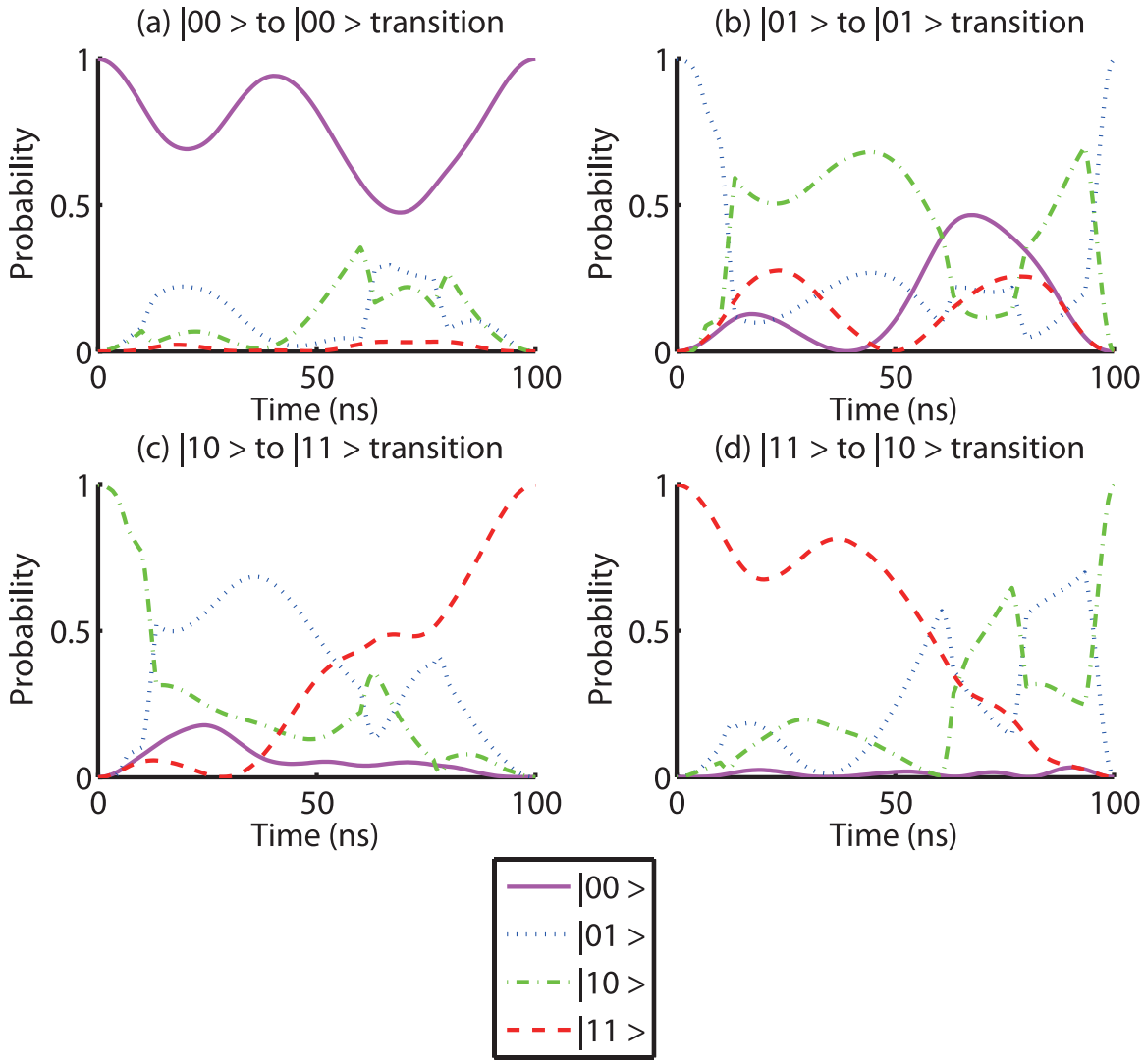

Figure 5.3 Numerical simulation of CNOT gate using full Hamiltonian in the rotating frame with the different initial conditions where all the nucleus are initially spin-up. During each time, we partial trace the density matrix over the nuclear states and obtain reduce density matrix for the electrons.

traced over all the nuclear spin states.

$$\rho^f_{|ij\rangle_e} = \mathrm{Tr}_n \left(\rho^f_{|ij\rangle_e \otimes |00\rangle_n} \right) , \tag{5.4}$$

where the subscripts i and j in the equation stand for the spin basis states. Define four projective measurement observable operators as

$$M_{00} = |00\rangle \langle 00| , \tag{5.5a}$$

$$M_{01} = |01\rangle \langle 01| , \tag{5.5b}$$

$$M_{10} = |10\rangle \langle 10| , \text{ and} \tag{5.5c}$$

$$M_{11} = |11\rangle \langle 11| , \tag{5.5d}$$

where the sum of the four measurement operators equals the identity operator. The probability of obtaining the result M_{ij} of electrons with spin input state $|kl\rangle_e \otimes |00\rangle_n$ is

$$P(|kl\rangle \rightarrow |ij\rangle) = \mathrm{Tr}(M_{ij} \rho^f_{|kl\rangle_e}) . \tag{5.6}$$

When the CNOT gate operation is finished, the errors of transition with the four input states to their correspondingly expected output electron spin states are shown in Table 5.1, where errors are defined as $1 - P$. The evolutions of the states of the CNOT gate are show in Fig. 5.3.

We find the errors of about 10^{-6} which is below the error threshold required for fault-tolerant (10^{-4}) quantum computation. Most of these errors result from the accuracy of the second-order approximation and the assumption that the eigenstates are the computational basis states. The CNOT gate operation time of 100ns is 3 times faster than the globally controlled electron spin scheme [16] of 297ns (in the paper [16] the indicated CNOT time is 148ns which is due to a factor of 2 missing in the denominator of the σ_z term of their Hamiltonian).

Table 5.1 Summary of the CNOT gate fidelities and the probabilities that the nuclear spins may flip after the CNOT gate operation.

Input State, $\|kj\rangle \otimes \|00\rangle_n$	Expected output state, $\|ij\rangle \otimes \|00\rangle_n$	The fidelity of CNOT gate[a]	The probability that nuclear spin may flip[b]
$\|00\rangle_e \otimes \|00\rangle_n$	$\|00\rangle_e \otimes \|00\rangle_n$	$1 - 1.80 \times 10^{-8}$	1.57×10^{-7}
$\|01\rangle_e \otimes \|00\rangle_n$	$\|01\rangle_e \otimes \|00\rangle_n$	$1 - 1.80 \times 10^{-7}$	2.00×10^{-7}
$\|10\rangle_e \otimes \|00\rangle_n$	$\|11\rangle_e \otimes \|00\rangle_n$	$1 - 1.92 \times 10^{-6}$	1.93×10^{-6}
$\|11\rangle_e \otimes \|00\rangle_n$	$\|10\rangle_e \otimes \|00\rangle_n$	$1 - 1.20 \times 10^{-6}$	1.56×10^{-6}

[a] The fidelity of CNOT gate operation corresponding to the four electron spin basis input states, $|kl\rangle$, is defined by $P(|kl\rangle \rightarrow |ij\rangle)$ in Eq. (5.8), where $|ij\rangle$ is the excepted output state. Note that in Eq. (5.6), the output state is traced over all the nuclear states first, and we obtain a reduce density matrix for the electron spin states. Finally, we use the measurement operator to compute the fidelities.

[b] The probability that nuclear spin may flip after the CNOT gate operation is computed by using Eq. (5.8). Here, we trace the total output density matrix over the electron spin states, and use the reduced density matrix for the nuclear spin states to compute the flipping probability.

5.3 Reinitialize the nuclear spin

We have been utilizing the assumption that the eigenstates are the computational basis state and therefore the nuclear spins will be always in spin-up states. However, this is not the case in real operation – the nuclear spin actually changes its state as time goes by. If the polarization of the nuclear spins are not perfectly spin-up state, it will increase the error of the gate operations. Once the tolerance of error goes above threshold often a series of gate operations, we may have to reinitialize the nuclear spins. Define the final reduced nuclear density matrix as the composite density matrix traced over all the electron spin states,

$$\rho^f_{|\phi\rangle_n} = \mathrm{Tr}_e \left(\rho^f_{|ij\rangle_e \otimes |\phi\rangle_n} \right) \ , \tag{5.7}$$

where the subscript $|\phi\rangle_n$ in the equation stands for the input nuclear state, which is expected to be $|00\rangle_n$. We can calculate the probability of the nuclear spin still being in the spin-up state by

$$P(|00\rangle_n) = \mathrm{Tr} \left(M^{(n)}_{00} \rho^f_{|\phi\rangle_n} \right) \ , \tag{5.8}$$

where $M^{(n)}_{00}$ is the nuclear spin measurement operator. With the initialized nuclear state - $|\phi\rangle_n = |00\rangle_n$, the error probabilities that nuclear spins may flip after the CNOT gate operation are shown in Table 5.1. The probabilities corresponding to the four different input electron states are around around 10^{-6}.

We investigate below how many gate operations can be allowed by the fault-tolerant error threshold before we have to reinitialize the nuclear spin states. With the initialized nuclear state and the four input computational basis states, i.e. $|ij\rangle_e \otimes |00\rangle_n$, we first perform a CNOT gate operation, then trace over the electron spin states, input the same electron state $|ij\rangle_e$, and then perform a CNOT gate again. If we repeat the process for N times by simply inputing the same pure electron state $|ij\rangle$ but not

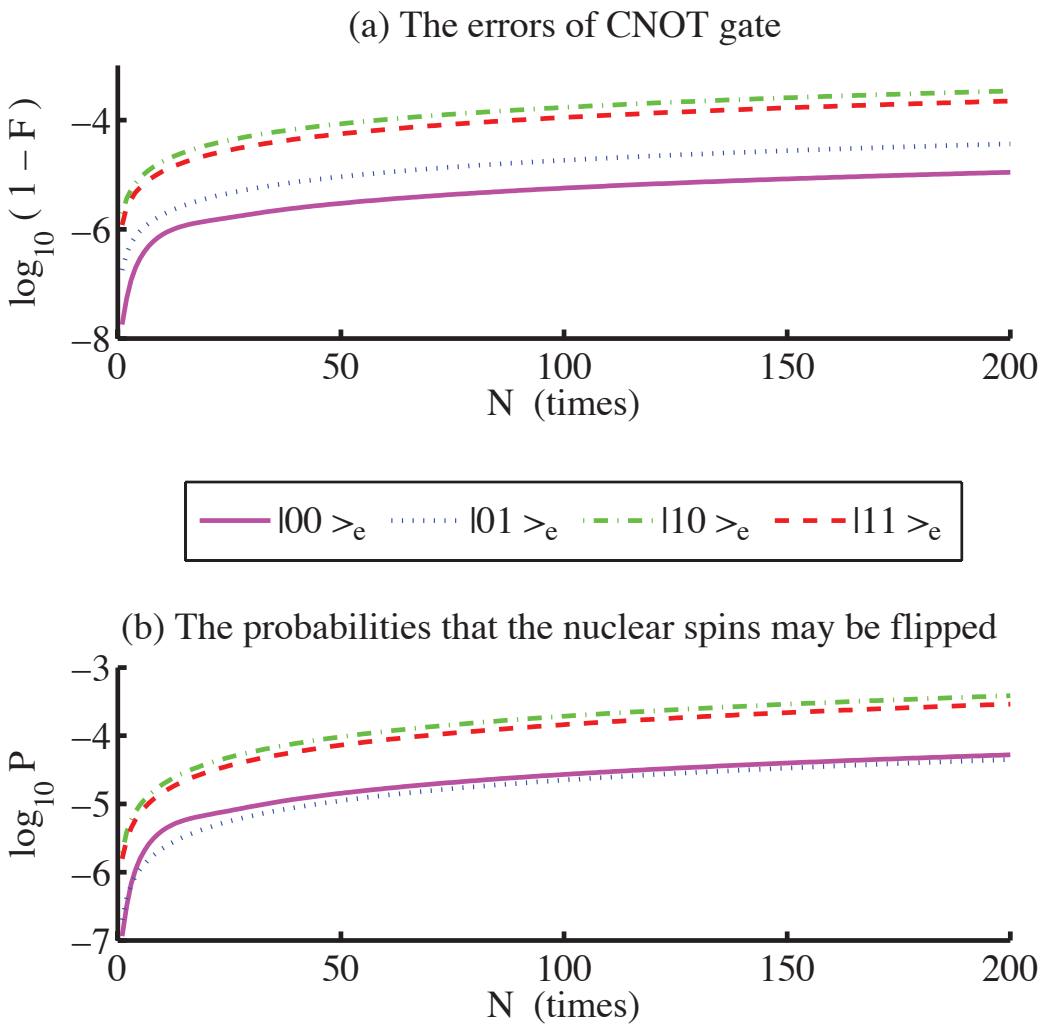

Figure 5.4 (a) gives the errors and (b) gives the probabilities that nuclei may be flipped due to the imperfect polarized nuclear spins caused by N times of CNOT operation without reinitializing the nuclear spins.

reinitializing the nuclear state, the errors of CNOT gate operation will accumulate. The numerical result shown in Fig. 5.4 indicate that in the worst case - $|10\rangle_e$, after around 60 times of operations, the error sums up to 1.03×10^{-4} , above the fault-tolerant threshold. The corresponding probability that nuclei may be flipped is up to 1.16×10^{-4}. Therefore, in order to maintain high fidelity, one has to reinitialize the nuclear spin right before about 60 times of operations.

5.4 The robust control over the AWGN channel

Since We apply voltage on the A and J gates to control the strength of hyperfine interaction and exchange interaction, there might be some noise induced from the thermal vibrations of atoms in the control circuits or the device's limitation in accurate control of the applied voltages. These uncertainties of the control parameters will decrease the fidelity of a specific operation. In order to analyze the decreasing of fidelity due to these uncertainties, we model the noise as additive white Gaussian noise (AWGN). The AWGN channel model is a random function which can simulate the noise due to thermal fluctuation in a circuit. It is a Gaussian white noise model that has a constant spectral density (expressed as watts per hertz of bandwidth) and a Gaussian distribution of amplitudes.

In engineering, signal-to-noise ratio (SNR) is often used to describe the amplitude ratio between a signal (our control sequence) and the background noise :

$$\mathrm{SNR(dB)} = 20 \log_{10} \left(\frac{\mathcal{A}_{signal}}{\mathcal{A}_{noise}} \right) , \tag{5.9}$$

where $\mathcal{A}_{noise}$ is defined as the root mean square (RMS) of the noise amplitudes in the control sequence, and $\mathcal{A}_{single}$ is defined as the maximum value of the signal amplitudes in the control sequence. That is to say, for A gate, the $\mathcal{A}_{signal} = A_0 = 1.21 \times 10^{-7}$eV, and for J gate, $\mathcal{A}_{signal} = 8.3 \times 10^{-8}$eV. The simulation results are shown in Fig. 5.5.

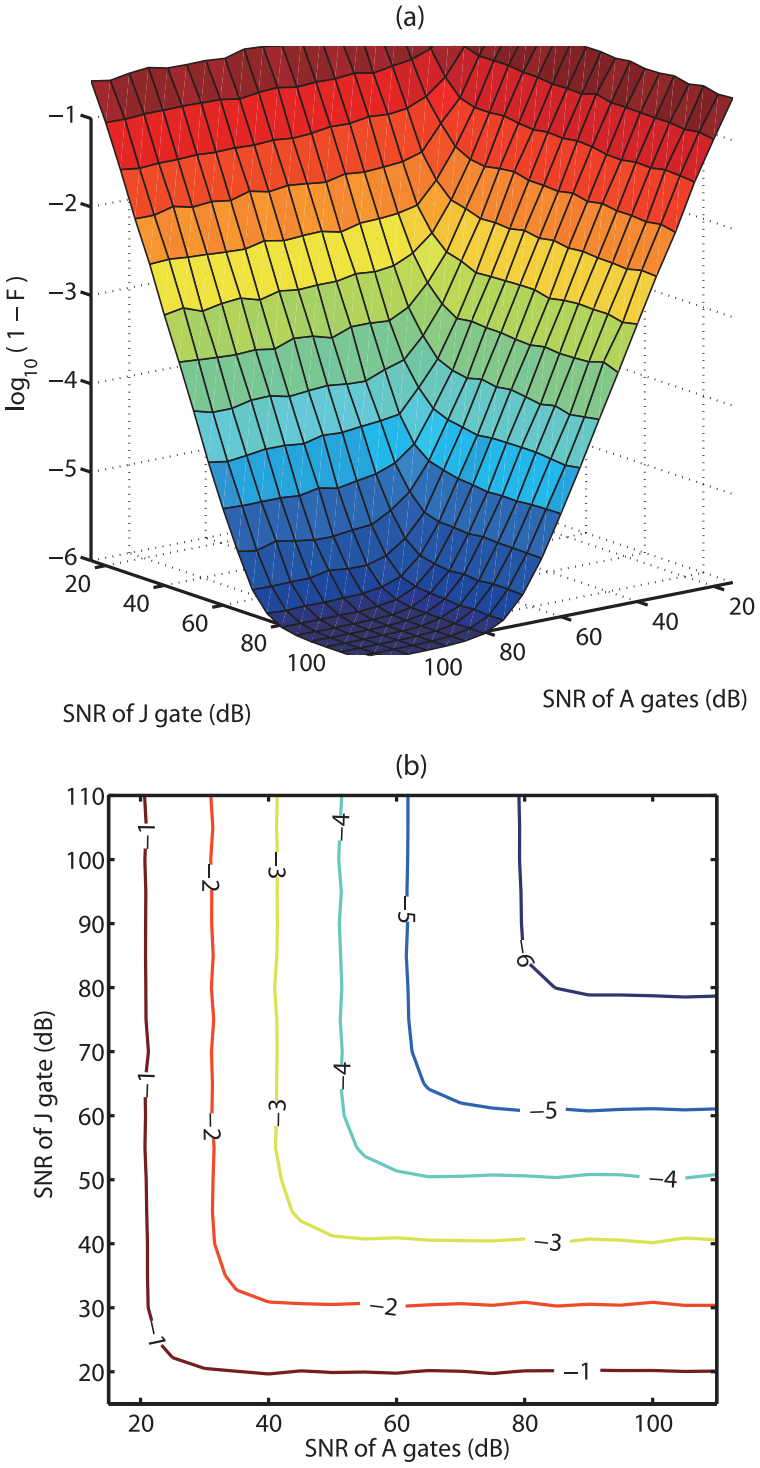

Figure 5.5 (a) The error simulated with Gaussian noise for different values of noises applied on A and J gates by using the full Hamiltonian. (b) The contour plots of error defined as $\log_{10}(1 - F)$.

The fault-tolerant error correction theory requires that the probability of introducing an error in each gate to be below 1×10^{-4}. To satisfy the error threshold, both the noises of A and J gates have to be larger than 55dB, meaning that $\mathcal{A}_{noise}/\mathcal{A}_{signal}$ has to be smaller than 1.78×10^{-3}. Therefore, the variations of the control parameters are around $\pm 0.2\%$ - achievable with modern electronic devices.

Chapter 6

Conclusions

We have investigated how pulse-sequences and operation times of elementary quantum gates can be optimized for silicon-based donor electron spin quantum computer architecture [8, 16], complementary to the original Kane's nuclear spin proposal [7, 13, 14]. This gate-sequence-optimal or time-optimal quantum gate control in a quantum circuit is in addition to the more conventional concept of optimality in terms of the number of elementary gates needed in a quantum transformation. The optimal control method we use is the so-called gradient ascent pulse engineering (GRAPE) scheme [11, 18]. We focus on the high fidelity controlled-NOT (CNOT) gate and explicitly find the digitized control sequences by optimizing the effective, reduced donor electron spin Hamiltonian, with external controls over the hyperfine A and exchange J interactions.

With operation time $t = 100ns$ and stopping criteria of error set to 10^{-16}, we find that the near time-optimal, high-fidelity CNOT gate control sequence has an error of 1.11×10^{-16}. We then simulate the control sequences of the CNOT gate, obtained from reduced Hamiltonian simulations, with the full spin Hamiltonian. We find the error of about 10^{-6} which is below the error threshold required for fault-

tolerant (10^{-4}) quantum computation. The CNOT gate operation time of 100ns is 3 times faster than the globally controlled electron spin scheme [16] of 297ns (in the paper [16] the indicated CNOT time is 148ns which is due to a factor of 2 missing in the denominator of the σ_z term of their Hamiltonian). One of the great advantages of this near optimal-time high fidelity CNOT gate is that the exchange interaction is not required to be strong (the maximum value is $J/h = 20$MHz compared to the typical value of 10.2GHz in [7, 13, 14, 16]). This relaxes significantly the stringent distance constraint of two neighboring donor atoms of about 10nm as reported in the original Kane's proposal [7] to be about 30nm which is within the reach of the current fabrication technology.

We have also studied the time-optimal unitary operation using the variational principle approach for a model Hamiltonian for quantum computation. Specifically, we construct the controlled-Z gate using this approach and canonical gate decomposition method, respectively. We find that the variational approach can give the optimal gate control sequence and operation time. Although the canonical decomposition method in this simple example can combine two steps of operations into a single step, the amplitudes of the controlled parameters in this approach do not have any relation. If we set the amplitudes of all the control parameters to be the same as given in the variational principle approach, we then find the canonical decomposition method could in this case also give the same results.

Unlike traditional decomposition method that decomposes gate operation into several single qubit operation and some interaction operation between qubits, this concept of the time-optimal control approaches (such as GRAPE and the variational principle approaches) is in a sense more like parallel computing. As a result, the more complex gate operation it is applied, the more time one may save, especially for those that could not be simply decomposed using the tradition method. So the

time-optimal control approaches maybe proven useful in implementing quantum gate operations in real quantum computing experiments in the future.

Bibliography

[1] P. W. Shor, "Algorithms for quantum computation: discrete logarithms and factoring," In 35th Annual Symposium on Foundations of Computer Science. IEEE Press, Los Alamos (1994).

[2] P. W. Shor, "Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer," SIAM J. Computing 26 (1997).

[3] L. K. Grover, "A fast quantum mechanical algorithm for database search," In *STOC '96: Proceedings of the twenty-eighth annual ACM symposium on Theory of computing*, pp. 212–219 (ACM, New York, NY, USA, 1996).

[4] L. K. Grover, "Quantum Mechanics Helps in Searching for a Needle in a Haystack," Phys. Rev. Lett. **79,** 325–328 (1997).

[5] L. K. Grover, "Quantum Computers Can Search Rapidly by Using Almost Any Transformation," Phys. Rev. Lett. **80,** 4329–4332 (1998).

[6] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information*, 2 ed. (Cambridge University Press, 2001).

[7] B. E. Kane, "A silicon-based nuclear spin quantum computer," Nature **393,** 133–137 (1998).

[8] L. C. L. Hollenberg, A. D. Greentree, A. G. Fowler, and C. J. Wellard, "Two-dimensional architectures for donor-based quantum computing," Phys. Rev. B **74,** 045311 (2006).

[9] A. Carlini, A. Hosoya, T. Koike, and Y. Okudaira, "Time-Optimal Quantum Evolution," Phys. Rev. Lett. **96,** 060503 (2006).

[10] A. Carlini, A. Hosoya, T. Koike, and Y. Okudaira, "Time-optimal unitary operations," Phys. Rev. A **75,** 042308 (2007).

[11] N. Khaneja, T. Reiss, C. Kehlet, T. Schulte-Herbrüggen, and S. J. Glaser, "Optimal control of coupled spin dynamics: design of NMR pulse sequences by gradient ascent algorithms," J. Magn. Reson. **172** (2005).

[12] A. G. Fowler, C. J. Wellard, and L. C. L. Hollenberg, "Error rate of the Kane quantum computer controlled-NOT gate in the presence of dephasing," Phys. Rev. A **67,** 012301 (2003).

[13] C. D. Hill and H.-S. Goan, "Fast nonadiabatic two-qubit gates for the Kane quantum computer," Phys. Rev. A **68,** 012321 (2003).

[14] C. D. Hill and H.-S. Goan, "Gates for the Kane quantum computer in the presence of dephasing," Phys. Rev. A **70,** 022310 (2004).

[15] A. M. Tyryshkin, S. A. Lyon, A. V. Astashkin, and A. M. Raitsimring, "Electron spin relaxation times of phosphorus donors in silicon," Phys. Rev. B **68,** 193207 (2003).

[16] C. D. Hill, L. C. L. Hollenberg, A. G. Fowler, C. J. Wellard, A. D. Greentree, and H.-S. Goan, "Global control and fast solid-state donor electron spin quantum computing," Phys. Rev. B **72,** 045350 (2005).

[17] T. Yamamoto, Y. A. Pashkin, O. Astafiev, Y. Nakamura, and J. S. Tsai, "Demonstration of conditional gate operation using superconducting charge qubits," Nature **425,** 941–944 (2003).

[18] A. Spörl, T. Schulte-Herbrüggen, S. J. Glaser, V. Bergholm, M. J. Storcz, J. Ferber, and F. K. Wilhelm, "Optimal control of coupled Josephson qubits," Phys. Rev. A **75,** 012302 (2007).

[19] L. Kettle, H.-S. Goan, S. C. Smith, L. C. L. Hollenberg, C. I. Pakers, and C. Wellard, (unpublished).

[20] C. D. Hill, Ph.D thesis (University of Queensland, Brisbane, Australia, 2006).

[21] C. Herring and M. Flicker, "Asymptotic Exchange Coupling of Two Hydrogen Atoms," Phys. Rev. **134,** A362–A366 (1964).

[22] M. J. Testolin, C. D. Hill, C. J. Wellard, and L. C. L. Hollenberg, "Robust controlled-NOT gate in the presence of large fabrication-induced variations of the exchange interaction strength," Phys. Rev. A **76,** 012302 (2007).

[23] R. B. Sidje, "Expokit. A Software Package for Computing Matrix Exponentials," ACM Trans. Math. Softw. 24 (1998).

[24] C. Moler and C. V. Loan, "Nineteen Dubious Ways to Compute the Exponential of a Matrix, Twenty-Five Years Later," Society for Industrial and Applied Mathematics 45 (2003).

[25] E. G. Birgin, J. M. Martinez, and M. Raydan, "Nonmonotone Spectral Projected Gradient Methods on Convex Sets," SIAM J. on Optim. **10** (2000).

[26] B. Kraus and J. I. Cirac, "Optimal creation of entanglement using a two-qubit gate," Phys. Rev. A **63,** 062309 (2001).

[27] G. Vidal, K. Hammerer, and J. I. Cirac, "Interaction Cost of Nonlocal Gates," Phys. Rev. Lett. **88,** 237902 (2002).

[28] M. J. Bremner, C. M. Dawson, J. L. Dodd, A. Gilchrist, A. W. Harrow, D. Mortimer, M. A. Nielsen, and T. J. Osborne, "Practical Scheme for Quantum Computation with Any Two-Qubit Entangling Gate," Phys. Rev. Lett. **89,** 247902 (2002).

[29] M. L. Liou, "A novel method of evaluating transient response," Proc. IEEEE 54 (1966).

[30] R. B. Sidje, *Parallel Algorithms for Large Sparse Matrix Exponentials: application to numerical transient analysis of Markov processes.* (Ph. D. thesis, Univ. of Rennes 1., 1994).

Appendix A

Computing Matrix Exponentials

A.1 Introduction

Evaluating the exponential of a matrix is an important problem that arises in physics, mathematics and engineering. For example, in quantum theory, a central problem consists in solving the ODE of Schrödinger equation,

$$\frac{\partial}{\partial t}\left|\psi\right\rangle = -\frac{i}{\hbar}H\left|\psi\right\rangle \ , \tag{A.1}$$

where H is a hermitian matrix and $|\psi\rangle$ is a complex vector. If the Hamiltonian, H is time independent, the solution for Eq. (A.1) is given by

$$\left|\psi_f\right\rangle = \exp\left(-\frac{i}{\hbar}Ht\right)\left|\psi_i\right\rangle \ , \tag{A.2}$$

where $|\psi_f\rangle$ is the final state, and $|\psi_i\rangle$ is the initial state.

There are several methods to compute matrix exponential. Methods involving approximation theory, differential equations, the matrix eigenvalues, and the matrix characteristic polynomial have been proposed.

A.2 Computed by Taylor Series Expansion

A direct way to define the matrix exponential $\exp(At)$ where t is a real number is undoubtedly through the exponential power series expansion,

$$e^{At} = I + \sum_{n=1}^{\infty} \frac{(At)^n}{n!} \,, \tag{A.3}$$

whose convergence is guaranteed for any square matrix A. If we momentarily ignore the efficiency, we can simply sum up the series until adding another term which does not alter the numbers stored in computer. That is, if

$$T_k(At) = \sum_{n=0}^{k} \frac{(At)^n}{n!} \,, \tag{A.4}$$

, then we can find k such that

$$\|T_k(At) - T_{k+1}(At)\| \leqslant \delta \,, \tag{A.5}$$

where δ is some prescribed error tolerance. Concern over where to truncate the series is of importance if efficiency is being considered. Unfortunately, such an algorithm is know to be inefficiency even in the scalar case, and there are lots of papers [24, 29] concerning the truncation error of Taylor series. As a result, directly summing up all Taylor series seems not work very well.

A.3 Computed by Diagonalization of the Matrix

For most physicists, another instinctive method is to diagonalize the matrix A, such that

$$A = VDV^{-1} \,, \tag{A.6}$$

then using the power series definition of e^{tA} implies

$$e^{tA} = Ve^{tD}V^{-1} \,, \tag{A.7}$$

where V is the matrix whose columns are eigenvectors of A, $D = \text{diag}(\lambda_1, ..., \lambda_n)$, and λ_n are the eigenvalues of A. The exponential of tD is very easy to compute and we have a satisfactory build-in routine in lots of computer languages for computing the exponential of a scalar,

$$e^{tD} = \text{diag}(e^{\lambda_1 t}, ..., e^{\lambda_n t}) \ . \tag{A.8}$$

The difficulty with this approach is not find the eigenvalues of the matrix A in itself, but occurs when A does not have a complete set of linearly independent eigenvectors. In this case, A is defective and there is no invertible matrix of V, therefore, the algorithm will break down. In real computing world, the difficulties occur even when A is nearly defective. Define the condition number as $\text{cond}(V) = \|V\|\|V^{-1}\|$. While A is nearly defective, then $\text{cond}(V)$ will be very large. The errors of computing e^{tA}, including the roundoff errors from the eigenvalues computation, may be magnified by $\text{cond}(V)$. As a result, the computed exponential of a matrix will most likely be inaccurate when $\text{cond}(V)$ is very large. An example had been demonstrated in this paper [24].

A.4 Computed by Padé Approximation

The most easy method for computing e^{At} numerically might be Taylor series, but we have discussed that this approach is inefficiency and not accurate. A Padé approximation often gives better result of a given function - e^At here than truncating its Taylor series, and it may still work where the Taylor series does not converge well. Since it is defined as a rational function which is a ratio of polynomial series, it can be calculated numerically easily. Therefore, Padé approximations are used extensively in computer calculation. A Padé approximation approximates a function in only one variable, an

approximation of a function in two variables is called a Chisholm approximation, and in multiple variables is called a Canterbury approximation.

A Padé rational approximation to $f(x)$ is the quotient of two polynomials $N_{p/q}(x)$ and $D_{p/q}(x)$ of degrees p and q respectively. This is so-called (p,q)-degree type Padé approximation. We use the notation $R_{p/q}(x)$ to denote this quotient:

$$R_{p/q}(x) = \frac{N_{p/q}(x)}{D_{p/q}(x)} \ , \tag{A.9}$$

where x is a scalar. If x is a matrix - A, the quotient is defined as

$$R_{p/q}(A) = \left[D_{p/q}(A)\right]^{-1} N_{p/q}(A) \ . \tag{A.10}$$

The polynomials used in Eq. (A.10) are

$$N_{p/q}(A) = n_0 \mathrm{I} + n_1 A + n_2 A^2 + ... + n_p A^p \ , \tag{A.11}$$

and

$$D_{p/q}(A) = \mathrm{I} + d_1 A + d_2 A^2 + ... + d_q A^q \ . \tag{A.12}$$

In the case $q = 0$, the approximation will reduce to the Taylor (Maclaurin) expansion for $f(A)$. There are $p+1$ unknown coefficients in $N_{p/q}(A)$, and q unknown coefficients in $D_{p/q}(A)$, hence the rational function $R_{p/q}(A)$ has $p + q + 1$ unknown coefficients. Assume that $f(A)$ is analytic and has the Maclaurin expansion,

$$f(A) = a_0 \mathrm{I} + a_1 A + a_2 A^2 + ... + a_k A^k + ... \ , \tag{A.13}$$

then $R_{p/q}(A)$ is said to be a Padé approximation to the series $f(A)$. Since the highest possible oder of nonzero derivative of $R_{p/q}(A)$ is $p + q$, that is

$$\frac{\partial^k}{\partial A^k} R_{p/q}(A) = 0 \ , \qquad \forall k \geqslant p + q + 1 \ , \tag{A.14}$$

as a result, the first $p+q$ derivatives of $f(A)$ and $R_{p/q}(A)$ are to agree at $A=0$,

$$\left.\frac{\partial^k}{\partial A^k} R_{p/q}(A)\right|_{A=0} = \left.\frac{\partial^k}{\partial A^k} f(A)\right|_{A=0}, \qquad k = 0, 1, ..., p+q\ . \tag{A.15}$$

Eq. (A.15) implies that,

$$R_{p/q}(A) - f(A) = \sum_{k=p+q+1}^{\infty} c'_k x^k. \tag{A.16}$$

Multiply $D_{p/q}$ on Eq. (A.16) giving

$$\begin{aligned} D_{p/q}(A)f(A) - N_{p/q}(A) &= D_{p/q} \sum_{k=p+q+1}^{\infty} c'_k x^k \\ &= \sum_{k=p+q+1}^{\infty} c_k x^k\ , \end{aligned} \tag{A.17}$$

and we obtain

$$\sum_{i=0}^{p} n_i A^i - \left(I + \sum_{i=1}^{q} d_i Ai\right)\left(\sum_{i=0}^{\infty} a_i A^i\right) = \sum_{k=p+q+1}^{\infty} c_k x^k\ . \tag{A.18}$$

When the left side of Eq. (A.18) is multiplied out out and the coefficients of the power of A^i are set equal to zero for $i = 0, 1, ..., p+q$, the result is a system of $p+q+1$ linear equations:

$$\begin{aligned} a_0 - n_0 &= 0 \\ d_1 a_0 + a_1 - n_1 &= 0 \\ d_2 a_0 + d_1 a_1 + a_2 - n_2 &= 0 \\ \vdots \qquad\quad \vdots \qquad\quad \vdots & \\ d_q a_{p-q} + d_{q-1} a_{p-q+1} + \cdots + a_p - n_p &= 0\ , \end{aligned} \tag{A.19}$$

and

$$\begin{aligned} d_q a_{p-q+1} + d_{q-1} a_{p-q+2} + \cdots + d_1 a_p + a_{p+1} &= 0 \\ d_q a_{p-q+2} + d_{q-1} a_{p-q+3} + \cdots + d_1 a_{p+1} + a_{p+2} &= 0 \\ \vdots \qquad\qquad \vdots \qquad\qquad \vdots & \\ d_q a_p + d_{q-1} a_{p+1} + \cdots + d_1 a_{p+q-1} + a_{p+q} &= 0\ . \end{aligned} \tag{A.20}$$

The q equations in Eq. (A.20) involve only the unknowns $d_1, d_2, \cdots, d_q$, and have to be solved first. Then the equations in Eq. (A.19) are used successively to find $n_0, n_1, \cdots, n_p$.

Go back to our original problem, and setting $f(A)$ as e^A gives us,

$$a_n = \frac{1}{n!} \, . \tag{A.21}$$

Solving Eq. (A.19) and Eq. (A.20) together with Eq. (A.24) gives us,

$$N_{p/q} = \sum_{i=0}^{p} \frac{(p+q-i)!p!}{(p+q)!i!(p-i)!} A^i \, , \tag{A.22}$$

and

$$D_{p/q} = \sum_{i=0}^{q} \frac{(p+q-i)!q!}{(p+q)!i!(q-i)!} (-A)^i \, . \tag{A.23}$$

It has been discussed that there are several reasons [24] why the diagonal approximants ($p = q$) are preferred over the off diagonal approximants ($p \neq q$) for stability and economy of computation. For $p = q$, we have $n_0 = 1$, $n_i = n_{i-1} \frac{p+1-i}{(2p+1-i)i}$, and $d_i = (-1)^i n_i$. As noted in Sidje's thesis [30] , Eq. (A.10) can be written as the following irreducible form for economical computing reason,

$$R_{p/p} = \begin{cases} 1 + 2 \left(\sum_{i=0}^{p/2} n_{2i} A^{2i} - A \sum_{i=0}^{p/2-1} n_{2i+1} A^{2i} \right)^{-1} \left(A \sum_{i=0}^{p/2-1} n_{2i+1} A^{2i} \right) \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{if } p \text{ is even} \, , \\ -1 - 2 \left(A \sum_{i=0}^{(p-1)/2} n_{2i+1} A^{2i} - \sum_{i=0}^{(p-1)/2} n_{2i} A^{2i} \right)^{-1} \left(\sum_{i=0}^{(p-1)/2} n_{2i} A^{2i} \right) \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{if } p \text{ is odd} \, . \end{cases} \tag{A.24}$$

Since Padé approximation is only accurate near the origin so that the approximation of $exp(A)$ is not valid, when $\|A\|$ is too large. Fortunately, this problem will be solved when we introduce so-called scaling and squaring' technology [23, 24]. We make use of the exponential property

$$exp(A) = \left[\exp\left(2^{-s} A\right) \right]^{2^s} \approx \left[\mathrm{R_{p/p}}\left(2^{-s} A\right) \right]^{2^s} \, , \tag{A.25}$$

where s is chosen such that $\|2^{-s}A\| \leqslant 1/2$. The idea is to choose s to be a nature number for which $M = \exp(2^{-s}A)$ can be reliably and efficiently computed, and then to compute the result $\exp(A) = M^{2^s}$ by repeated squaring. Because the result is evaluated by repeated squaring the exponential of scaled matrix, the drawback of this algorithm may come from the fact that if $s \gg 1$, then the roundoff errors may be large. The error analysis discussed in these papers [23, 24] has shown that if $\|2^{-s}A\| \leqslant 1/2$ then

$$\left[\mathrm{R}_{\mathrm{p/p}}\left(2^{-s}A\right)\right]^{2^s} = \exp(A+E)\ , \tag{A.26}$$

where

$$\frac{\|E\|}{\|A\|} \leqslant \frac{(p!)^2}{((2p)!(2p+1)!}\left(\frac{1}{2}\right)^{2p-3} \approx \begin{cases} 0.34\times 10^{-15} & (p=6) \\ 0.11\times 10^{-18} & (p=7) \\ 0.27\times 10^{-22} & (p=8) \end{cases} . \tag{A.27}$$

Therefore, a value of $p = 6$ is generally satisfactory for computer computing while using double precision.

A.5 Matrix Exponential Source Code

In this section, we will show the source code of matrix exponential in Listing A.1 which is implemented based on irreducible (p, p)-degree Padé approximation described in the previous section and Fortran source code of Expokit software package [23]. Our source code is written by C++, and we take the advantage of Boost C++ Libraries which have a lots of useful Basic Linear Algebra routines (uBLAS). Our source code is released under GPLv2, its later version, or Boost Software License Version 1.0.

Listing A.1 Matrix Exponential C++ Source Code

```
/*  expm.hpp
*   Implement matrix exponential using pade approximation.
*
*   Copyright (c) 2007
*   Tsai, Dung-Bang
*
*   Department of Physics,
*   National Taiwan University.
*   Version : v0.5
*
*   expm_pad computes the matrix exponential exp(H) for
*   general matrixs, including complex and real matrixs
*   using the irreducible (p,p) degree rational Pade
*   approximation to the exponential
*   exp(z) = r(z) = (+/-)( I+2*(Q(z)/P(z))).
*
*   Usage :
*     U = expm_pad(H)
*     U = expm_pad(H, t),
*     U = expm_pad(H, t, p),
*   where t is a real number which is default set to 1.0
*   such that U=exp(t*H), and p is internally set to 6
*   (recommended and gererally satisfactory).
*
* Licenses : GPLv2, its later version,
```

```
*          or Boost Software License Version 1.0.
*/

#ifndef _BOOST_UBLAS_EXPM_
#define _BOOST_UBLAS_EXPM_
#include <boost/numeric/ublas/vector.hpp>
#include <boost/numeric/ublas/matrix.hpp>
#include <boost/numeric/ublas/lu.hpp>
#include <boost/numeric/ublas/traits.hpp>

namespace boost { namespace numeric { namespace ublas {

template<typename MATRIX> MATRIX expm_pad(
                  const MATRIX &H, typename type_traits<
                  typename MATRIX::value_type>
                          ::real_type t = 1.0,
        const int p = 6)
{
  typedef typename MATRIX::value_type value_type;
  typedef typename MATRIX::size_type size_type;
  typedef typename type_traits<value_type>
    ::real_type real_value_type;
  assert(H.size1() == H.size2());
  assert(p >= 1);
  const size_type n = H.size1();
  const identity_matrix<value_type> I(n);
```

```
  matrix<value_type> U(n,n),H2(n,n),P(n,n),Q(n,n);
  real_value_type norm = 0.0;
// Calcuate Pade coefficients
  vector<real_value_type> c(p+1);
  c(0)=1;
  for(size_type i = 0; i < p; ++i)
    c(i+1) = c(i) * ((p - i)/((i + 1.0) * (2.0 * p - i)));
// Calcuate the infinty norm of H, which is defined
// as the largest row sum of a matrix
 for(size_type i=0; i<n; ++i) {
   real_value_type temp = 0.0;
   for(size_type j = 0; j < n; j++)
     temp += std::abs(H(i, j));
   norm = t * std::max<real_value_type>(norm, temp);
  }
// If norm = 0, and all H elements are not NaN or
// infinity but zero, then U should be identity.
  if (norm == 0.0) {
    bool all_H_are_zero = true;
    for(size_type i = 0; i < n; i++)
       for(size_type j = 0; j < n; j++)
          if( H(i,j) != value_type(0.0) )
            all_H_are_zero = false;
    if( all_H_are_zero == true ) return I;
// Some error happens, H has elements
// which are NaN or infinity.
```

```
  std::cerr<<"Null input error in the template expm_pad.\n";
  std::cout << "Null INPUT : " << H <<"\n";
  exit(0);
  }
// Scaling, seek s such that || H*2^(-s) || < 1/2,
// and set scale = 2^(-s)
  int s = 0;
  real_value_type scale = 1.0;
  if(norm > 0.5) {
    s = std::max<int>(0, static_cast<int>
          ((log(norm) / log(2.0) + 2.0)));
    scale /= real_value_type(std::pow(2.0, s));
    U.assign((scale * t) * H);
// Here U is used as temp value due to that H is const
  }
  else
     U.assign(H);

// Horner evaluation of the irreducible fraction,
// Initialise P (numerator) and Q (denominator)
  H2.assign( prod(U, U) );
  Q.assign( c(p)*I );
  P.assign( c(p-1)*I );
  size_type odd = 1;
  for( size_type k = p - 1; k > 0; --k) {
     ( odd == 1 ) ?
```

```
            ( Q = ( prod(Q, H2) + c(k-1) * I ) ) :
            ( P = ( prod(P, H2) + c(k-1) * I ) ) ;
        odd = 1 - odd;
    }
    ( odd == 1 ) ? ( Q = prod(Q, U) ) : ( P = prod(P, U) );
      Q -= P;
// Since in ublas, there is no matrix inversion template,
// I simply use the build-in LU decompostion package
// in ublas, and back substitute by myself.

// Implement Matrix Inversion
    permutation_matrix<size_type> pm(n);
    int res = lu_factorize(Q, pm);
    if( res != 0) {
      std::cerr << "Matrix inversion error
                    in the template expm_pad.\n";
      exit(0);
    }
// H2 is not needed anymore, so it is
// temporary used as identity matrix for substituting.
    H2.assign(I);
    lu_substitute(Q, pm, H2);
    (odd == 1) ?
      ( U.assign( -(I + real_value_type(2.0) * prod(H2, P))) ):
      ( U.assign(    I + real_value_type(2.0) * prod(H2, P) ) );
// Squaring
```

```
  for(size_type i = 0; i < s; ++i)
    U = (prod(U,U));
  return U;
}
}}}
#endif
```