MODELING, ANALYSIS AND CONTROL OF QUANTUM ELECTRONIC DEVICES

A Dissertation

by

ZHIGANG ZHANG

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

August 2006

Major Subject: Mathematics

MODELING, ANALYSIS AND CONTROL OF QUANTUM ELECTRONIC DEVICES

A Dissertation

by

ZHIGANG ZHANG

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Approved by:

Chair of Committee,	Goong Chen
Committee Members,	Philip Hemmer
	J. Maurice Rojas
	Joe Pasciak
Head of Department,	Al Boggess

August 2006

Major Subject: Mathematics

ABSTRACT

Modeling, Analysis and Control
of Quantum Electronic Devices. (August 2006)
Zhigang Zhang, B.S., University of Science and Technology of China;
M.S., University of Science and Technology of China
Chair of Advisory Committee: Dr. Goong Chen

This dissertation focuses on two connected areas: quantum computation and quantum control. Two proposals to construct a quantum computer, using nuclear magnetic resonance (NMR) and superconductivity, are introduced. We give details about the modeling, qubit realization, one and two qubit gates and measurement in the language that mathematicians can understand and fill gaps in the original literatures. Two experimental examples using liquid NMR are also presented. Then we proceed to investigate an example of quantum control, that of a magnetometer using quantum feedback. Previous research has shown that feedback makes the measurement robust to an unknown parameter, the number of atoms involved, with the assumption that the feedback is noise free. To evaluate the effect of the feedback noise, we extend the original model by an input noise term. We then compute the steady state performance of the Kalman filter for both the closed-loop and open-loop cases and retrieve the estimation error variances. The results are compared and criteria for evaluating the effects of input noise are obtained. Computations and simulations show that the level of input noise affects the measurement by changing the region where closed loop feedback is beneficial.

To my wife and daughter

ACKNOWLEDGMENTS

I would like to thank my committee chair, Dr. Goong Chen, and my committee members, Dr. Hemmer, Dr. Rojas, and Dr. Pasciak, for their guidance and support through the course of this research.

Thanks also to my friends and colleagues and the department faculty and staff for making my time at Texas A&M University a great experience. I am also grateful to Dr. Zijian Diao for stimulating discussions.

Finally, thanks to my wife and daughter, for their patience and love, and my parents for their encouragement and understanding.

TABLE OF CONTENTS

CHAPTER Page

LIST OF FIGURES

FIGURE Page

FIGURE Page

FIGURE Page

FIGURE Page

CHAPTER I

INTRODUCTION

In the past two decades, the research on quantum computer and quantum information has undergone rapid development and attracted great attention. One reason is the expected end of the famous Moore's law which has been driving the IT industry for a long time. When more and more transistors are integrated in one chip, the size of a transistor becomes smaller and smaller and will finally reach the limit of hundreds of atoms. The quantum effect then appears and naturally raises the question what we can do in such a dimension of several atoms. Another reason is the development of *quantum algorithms* [1, 2, 3, 4]. It provides us with a scenario that a quantum computer has the potential to solve problems that a classical computer can not, making it superior to its classical counterpart.

A suitable quantum system for the purpose of quantum computation must satisfy several conditions. It normally consists of many two-level subsystems and each subsystem constitutes a *qubit* (quantum bit), an elementary information unit analogous to a bit in a classical computer. First, the system is initialized to a certain state before computation. During the computation, the system needs to be isolated from its environment to reduce unwanted decoherence. Evolution of every subsystem is controlled by its time varying Hamiltonian, and coupling between qubits is necessary. In an ideal situation, the coupling can be switched on or off. Finally, after the computation, experimental methods are needed to retrieve the result, which means to couple the system with its environment and apply a measurement. Many quantum computing devices have been developed, based on AMO (atomic, molecular

---

This dissertation follows the style of Phys. Rev. A.

and optical) or semiconductor physics and technologies. Experiments on different systems with different algorithms also have been reported. Currently, liquid NMR is still the most successful through which a 7-qubit Shor's algorithm has been tested in laboratory conditions [4, 5]. A great obstacle facing scientists and engineers today is decoherence. Most literature of quantum computer is written by physicists and is not easy to read for mathematicians. In the first part of my dissertation, I will give the models and description in more mathematical terms for two types of quantum computers: NMR and superconductivity.

A closely related area to quantum computation is *quantum control* which is essential in the realization of a quantum computer [8, 9, 10, 11, 12, 13, 14]. Inspired by the goal to build a fast and realistic quantum computer, many control problems have arisen. Some of those problems have been widely studied already in classical control theory, but others are new in need of further investigation. For example, bilinear model appears in the pulse design of the operation of qubits and optimal control theory is used to find the shortest time path. To deal with decoherence and uncertainties, robust designs of pulses are desired in order to achieve high fidelity.

Although quantum control has potential applications and is essential to quantum computation, it is still a long-term goal and needs much work. A short term application is in a field called "quantum metrology" [15, 16, 17], which has a similar setting and theoretical background to quantum computation. In the second part of my dissertation, I will introduce a setup of magnetometry by a Caltech group to test the effect of closed-loop feedback. Initial results show that the feedback makes the measurement robust versus the uncertainty in the number of spins involved. We extend the model to include input noise, and we will attempt to find conditions under which the input noise will overwhelm the benefit brought on by feedback control.

CHAPTER II

NMR QUANTUM COMPUTING

One can never overstate the importance of computers in today's information society. Computers come in various types, ranging from the antique gigantic ENIAC to the modern portable laptop, and supercomputers such as "Deep Blue". But there is a universal model that applies to all known physical computers, the Turing machine [18].

The Turing machine is initially defined to address Hilbert's Entscheidungsproblem, or to model a formalist mathematical reasoning. The whole procedure is simplified by Alan Turing in several steps: representing the symbols with sequence of 1s and 0s; using a 1-dimensional tape as a writing pad; inventing a read/write head which moves back and forth over the tape; and allowing the head to exist in different states which define where to read or write. The computation begins with the "program" and the initial data written on the tape. The head is put in a state to read the program. The program is read and interpreted, i.e., the head uses this instruction from the program to define its movement, either moving backward/foreward, or reading/writing on the tape. It turns out that the Turing machine is universal. The Church-Turing thesis states: *All algorithmic process can be simulated efficiently using a Turing machine.*

The Turing machine originally designed by Turing is deterministic. The movement of the head and its states are totally defined by the program and data initially saved on the tape. That is how the electronic computer works where the memory serves as a tape and the CPU interprets the saved program and decides where to read and where and what to write. In contrast, a quantum system is indeterminate and for many years quantum effects had never been considered in the theory of computation until the early 1980's. Benioff [19] first coined the term of quantum

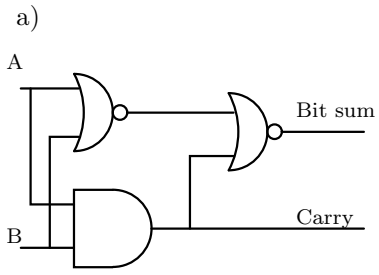
Turing machine (QTM). Motivated by the problem that classical computers can not simulate quantum systems efficiently, Feynman [20] posed the quantum computer as a solution. Now we know that, in terms of computability, quantum computers and classical computers possess exactly the same computational power. But in terms of computational complexity, which measures the efficiency of computation, there are many exciting examples confirming that quantum computers do solve certain problems faster. The two most significant ones are Shor's factorization algorithm [4] and Grover's search algorithm [3], among others such as the Deutsch-Jozsa problem [1], the Bernstein-Vazirani problem [21], and Simon's problem [22].

In contrast to a classical system, a quantum system can exist in different states at the same time, an interesting phenomenon called *superposition*. Superposition enables quantum computers to encode many inputs for a problem simultaneously and process all the data at the same time, called *quantum parallelism*. That is why a quantum computer can solve certain problems faster than a classical computer.

In this dissertation, we will give a detail explanation of quantum computation and examples to physically realize a quantum computer using the nuclear magnetic resonance (NMR) and superconductivity (SQUID). This, Chapter II, is divided into several parts: in Section A, we will give a short introduction of quantum computation; in Section B, we will show some basic quantum mechanics tools, including the Schröding equation and the evolution of a two level quantum system; in Section C, the NMR realization will be explained in detail; in Section D and Section E, experiments of Shor's algorithm and lattice gas algorithm are introduced, respectively.

A. Quantum computation

In a classical electronic computer, the information is encoded in multi-bits binary states (0 or 1) and saved in registers. The computation is accomplished by processing the information with logic gates in concatenation. The transformation after the processing, and thus that after every gate, can be presented as a table that maps every input state to a designed output state, called a "truth table." Fig. 2.1 illustrates the logic circuit of half-adder and its truth table. It has two input binary bits and two output binary bits. It is composed of two NOR gates and one AND gate and in turn it can be used in more complicated logic circuit as a brick. A logic gate in classical computer may not be revertable. Assume it has m input binary bits and n output binary bits, it can be viewed as a mapping from group 2^m to 2^n.



b)

A	B	Bit sum	Carry
0	0	0	0
0	1	1	0
1	0	1	0
1	1	0	1

Fig. 2.1. a) . The logic circuit of a half adder which has two input binary digits, called two bits, and two output binary bits. It is composed of two NOR gates and one AND gate. The output "bit sum" is in fact the exclusion OR of input A and input B. Thus it is zero whenever A and B are the same, otherwise it is set to 1. The "carry" output is set to 1 only when both A and B are 1. b). The truth table of the half-adder. It is not invertable since there are two states are mapped into the same output state.

Current physical realization of quantum computers follows the quantum circuit model [23], much like the logic circuit in Fig. 2.1, instead of the QTM model. Quan-

tum circuit model is another fundamental model of computation, which is equivalent to the QTM model [24], but easier to implement. This model shares many common features of the classical computers. In a quantum computer, information is represented by the quantum states of the qubits, and manipulated by various quantum control mechanisms. Those control mechanisms trigger quantum operation to process information in a way resembling the gates in a classical computer. Such quantum operations are called *quantum gates* and a series of quantum gates in concatenation constitute a *quantum circuit* [25]. However, because of the special effects of quantum mechanics, major distinctions exist.

In contrast to a classical system, a quantum system can exist in different states at the same time, an interesting phenomenon called *superposition*. Superposition enables quantum computers to process data in parallel. That is why a quantum computer can solve certain problems faster than a classical computer. From now on, we will use the Dirac bra-ket notation. In this notation a pure one-qubit quantum state can be written as $|\phi\rangle = a|0\rangle + b|1\rangle$. Here $|0\rangle$ and $|1\rangle$ are the two basis states of the qubit, e.g., in NMR, the spin-up and spin-down states, and $a, b \in \mathbf{C}$ with $|a|^2 + |b|^2 = 1$. When we make a measurement of a qubit, the result might be either $|0\rangle$ or $|1\rangle$, with probabilities $|a|^2$ and $|b|^2$ respectively. More generally, a string of n qubits can exist in any state of the form $|\psi\rangle = \sum_{x=00\ldots0}^{11\ldots1} \psi_x|x\rangle$, where $\psi_x \in \mathbf{C}$ and $\sum |\psi_x|^2 = 1$. When we make a measurement on $|\psi\rangle$, it collapses to $|x\rangle$, one of the 2^n basis states, with probability $|\psi_x|^2$. This indeterministic nature makes the design of efficient quantum algorithms highly non-trivial.

Another distinctive feature of the quantum circuit is that the operations performed by quantum gates must be unitary ($U^\dagger U = I$). It is the natural consequence of the unobserved quantum systems evolving according to the Schrödinger equation. A quantum gate may operate on any number of qubits. Here are some examples (see

Fig. 2.2 for the circuit diagrams) [18]:

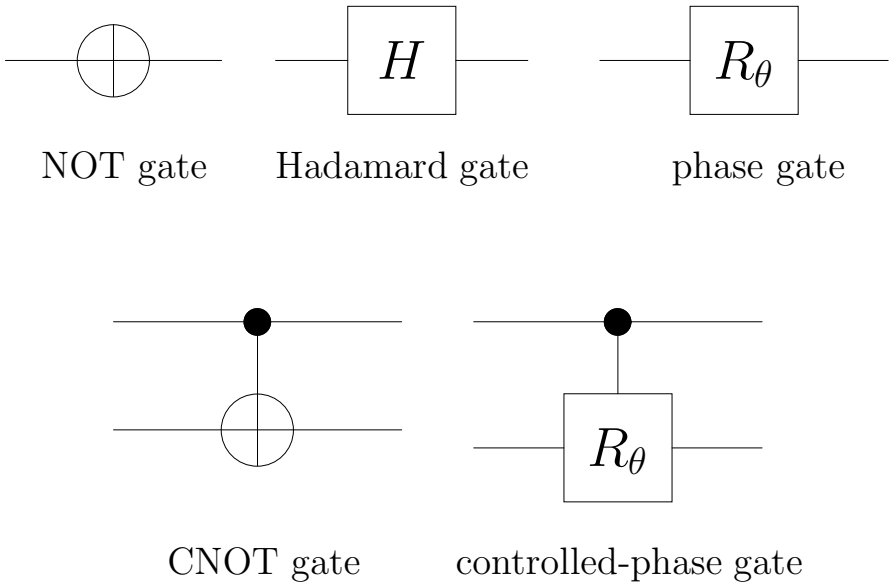


Fig. 2.2. Circuit diagrams of the NOT/Hadamard/phase/CNOT/controlled-phase gate.

1. NOT gate Λ_0: $\Lambda_0|0\rangle = |1\rangle$, $\Lambda_0|1\rangle = |0\rangle$, or $\Lambda_0 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$.

2. The Hadamard gate H: $H|0\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$, $H|1\rangle = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$, or

$$H = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}.$$

3. One-qubit phase gate R_θ: $R_\theta|0\rangle = |0\rangle$, $R_\theta|1\rangle = e^{i\theta}|1\rangle$, or

$$R_\theta = \begin{bmatrix} 1 & 0 \\ 0 & e^{i\theta} \end{bmatrix}.$$

4. Two-qubit controlled-NOT (CNOT) gate Λ_1: $\Lambda_1|00\rangle = |00\rangle$, $\Lambda_1|01\rangle = |01\rangle$, $\Lambda_1|10\rangle = |11\rangle$, $\Lambda_1|11\rangle = |10\rangle$, or,

$$\Lambda_1 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}.$$

5. Two-bit controlled-phase gate $\Lambda_1(R_\theta)$, where R_θ is the one-bit phase gate: $\Lambda_1(R_\theta)|00\rangle = |00\rangle$, $\Lambda_1(R_\theta)|01\rangle = |01\rangle$, $\Lambda_1(R_\theta)|10\rangle = |10\rangle$, $\Lambda_1(R_\theta)|11\rangle = e^{i\theta}|11\rangle$, or,

$$\Lambda_1(R_\theta) = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & e^{i\theta} \end{bmatrix}.$$

The one-qubit and two-qubit quantum gates are of particular importance to the construction of a quantum computer, because of the following universality result.

Theorem A.1. *(D. DiVincenzo [26, 27]) The collection of all the one-qubit gates and the two-qubit CNOT gate suffice to generate any unitary operations on any number of qubits.*

Fig. 2.3 illustrates, as an example, how to generate the two-qubit controlled-phase gate using 2 CNOT gates and 3 one-qubit phase gates. The controlled-phase gate is an important building block for the quantum Fourier transform.

B. A short introduction to quantum mechanics

Quantum mechanics is the greatest discovery in the 20th century. When Einstein first envisioned that the light propagates like particles in his paper about the pho-

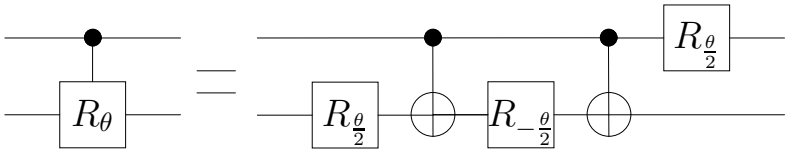


Fig. 2.3. Construction of the controlled-phase gate with CNOT gates and phase gates.

toelectric effect after Planck's description of black body radiation, the perception and understanding of the physical world was changed. Then de Broglie extended the wavefunction of light to particles, such as electrons. Erwin Schrödinger took the de Broglie "matt waves" further and postulated the famous Schrödinger equation. Assuming that the wavefunction is $\psi(\boldsymbol{r}, t)$, where $\boldsymbol{r}$ is the coordinate vector and t is time, he argued that the partial derivative operator of time corresponds to "quantized" energy, $i\hbar\dfrac{\partial}{\partial t} = \hat{E}$, and the gradient operator corresponds to momentum: $-i\hbar\dfrac{\partial}{\partial \boldsymbol{r}} = \hat{p}$. Here we use hat to mean it is an operator. Recall that the energy of a particle consists of its kinetic energy and potential energy, $E = \dfrac{p^2}{2m} + V$. By replacing the energy and momentum with the corresponding operators, he derived the following:

$$\begin{aligned} i\hbar\frac{\partial\psi(\boldsymbol{r},t)}{\partial t} &= \hat{E}\psi(\boldsymbol{r},t) \\ &= -\frac{\hbar^2}{2m}\Delta^2\psi(\boldsymbol{r},t) + V(\boldsymbol{r},t)\psi(\boldsymbol{r},t). \end{aligned} \tag{2.1}$$

Interpretation of $\psi(\boldsymbol{r}, t)$ was given late and not satisfied by many physicists. A concrete interpretation, which is what we will use, is that $|\psi(\boldsymbol{r}, t)|^2$ is the probability density that one can find a particle such as an electron at position $\boldsymbol{r}$ and time t. This means that $\psi(\boldsymbol{r}, t)$ may be complex and it lives in $\mathcal{L}^2$ since $\int |\psi(\boldsymbol{r}, t)|^2 dr = 1$. In general, wavefunctions are not necessary a function of $\boldsymbol{r}$ and in the simple form like that of an electron, but they still satisfies the Schrödinger equation:

$$i\hbar\frac{\partial}{\partial t}|\psi(t)\rangle = H|\psi(t)\rangle, \tag{2.2}$$

where H is an operator called Hamiltonian, corresponding to the system energy. We put $\psi(t)$ in a half bracket like $|\rangle$ in the Dirac notation and the symbol $|\rangle$ is called a "ket". If H does not depend on time, the Schrödinger equation can be formally represented as an "exponential" time evolution operator solution:

$$|\psi(t)\rangle = e^{-iHt/\hbar}|\psi(0)\rangle. \tag{2.3}$$

If we regard $|\psi\rangle$ as a vector in Hilbert space $\mathcal{L}^2$, we can denote its dual as $\langle\psi|$. If $|\psi\rangle$ and $|\phi\rangle$ are two *states*(say, respectively, two wavefunctions), then we denote their inner product as $\langle\phi|\psi\rangle$. In the case of an electron where $|\psi\rangle$ is a function of $\boldsymbol{r}$,

$$\langle\phi|\psi\rangle = \int \phi(\boldsymbol{r})^*\psi(\boldsymbol{r})dr. \tag{2.4}$$

Thus the unitarity of $|\psi\rangle$ can now be simply expressed as $\langle\psi|\psi\rangle = 1$, and we know all the possible wavefunctions form a unit sphere in $\mathcal{L}^2$. Let $\{|\psi_i\rangle\}_{i\in N}$ be an orthonormal bais of $\mathcal{L}^2$, then the following identity holds:

$$I = \sum_{i\in N} |\psi_i\rangle\langle\psi_i|, \tag{2.5}$$

where I is the identity operator. A direct result of (2.5) is the matrix representation of the system Halmtonian:

$$\begin{aligned} H &= \textstyle\sum_{i\in N} |\psi_i\rangle\langle\psi_i| H \sum_{j\in N} |\psi_j\rangle\langle\psi_j| \\ &= \textstyle\sum_{i,j\in N} \langle\psi_i|H|\psi_j\rangle|\psi_i\rangle\langle\psi_j|, \end{aligned} \tag{2.6}$$

and $\langle\psi_i|H|\psi_j\rangle$ is called the *matrix coefficent* of the Hamiltonian. Although in most cases, we prefer to choose the eigenstates of the system Hamiltonian as the basis, which corresponds to the energy levels with real eigenvalues, other basis is sometimes more convenient.

1. Two level system and Bloch sphere

In practice, the possible states of a quantum system form an infinite dimensional space. But, in some cases, we can constrain it to a finite dimensional subspace. One such example is when the system is in the low temperature, where only the states of several lowest energy levels are active, and transition probabilities to other higher levels are too small to have any observable effect, such as superconductivity. Another situation is when there is a proper subspace which has minimum probability for the system to jump out the subspace, or all the other states are kept off resonance. An ideal qubit has two orthonormal eigenstates as basis, which simplify the space into a two dimensional linear space $\mathbf{C}^2$, and the quantum state corresponds to a unit vector in $\mathbf{C}^2$. Taking the two orthonormal states as $|0\rangle$ and $|1\rangle$, we can write a wavefunction $|\phi\rangle$ as the linear combination of $|0\rangle$ and $|1\rangle$: $|\phi\rangle = a|0\rangle + b|1\rangle$. With the normalization condition, $|a|^2 + |b|^2 = 1$. We can also assume that a is real since only the relative phase is important. Thus this state can be represented using two angles θ and ψ:

$$|\phi\rangle = \cos\frac{\theta}{2}|0\rangle + e^{i\psi}\sin\frac{\theta}{2}|1\rangle, \tag{2.7}$$

where $\theta \in [0, \pi]$ and $\psi \in [0, 2\pi)$.

For the study of NMR spectroscopy with many nuclei, density matrices are preferred and are often written as the linear combination of *product operators* [28]:

$$\begin{aligned}\rho &= |\phi\rangle\langle\phi| \\ &= \begin{bmatrix} \cos^2\frac{\theta}{2} & e^{-i\psi}\frac{\sin\theta}{2} \\ e^{i\psi}\frac{\sin\theta}{2} & \sin^2\frac{\theta}{2} \end{bmatrix} \\ &= I_0 + \sin\theta\cos\psi I_x + \sin\theta\sin\psi I_y + \cos\theta I_z,\end{aligned} \tag{2.8}$$

where the product operators are defined as

$$I_0 = \frac{1}{2}\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad I_x = \frac{1}{2}\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad I_y = \frac{1}{2}\begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \quad I_z = \frac{1}{2}\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}. \tag{2.9}$$

They are different from the Pauli matrices[1] only by a constant factor and share the similar commutative law. Upon collecting all the coefficients of I_x, I_y, and I_z together, we obtain a vector

$$\boldsymbol{v} = [\sin\theta\cos\psi \quad \sin\theta\sin\psi \quad \cos\theta]^T,$$

which is called a *Bloch vector*. In essence, we have defined a mapping from the set of unit vectors $|\phi\rangle \in \mathbf{C}^2$ to the set of unit vectors $\boldsymbol{v} \in \mathbf{R}^3$. We have good reasons to ignore the coefficient of I_0, since it has no effect on the spectroscopy and remains unchanged under any unitary transformation. Each Bloch vector determines a point on the unit sphere, called the *Bloch sphere*, which is displayed in Fig. 2.4 [29, 28]. Bloch vectors have proven to be a very good tool for NMR quantum operations.

The mapping defined above is surjective, because every point on the Bloch sphere gives rise to a unit vector $\boldsymbol{v} = [\sin\theta\cos\psi \quad \sin\theta\sin\psi \quad \cos\theta]^T$ for some pair of (θ, ψ). Conversely, if $\boldsymbol{v}(\theta', \psi') = \boldsymbol{v}(\theta, \psi)$, we get

$$\begin{cases} \cos\theta & = & \cos\theta', \\ \sin\theta\cos\psi & = & \sin\theta'\cos\psi', \\ \sin\theta\sin\psi & = & \sin\theta'\cos\psi', \end{cases} \tag{2.10}$$

which can be used to show that the mapping is also injective if we identify all pairs of $(0, \psi)$ with one point and all pairs of (π, ψ) with another point. In fact, these two sets correspond to two states $|0\rangle$ and $|1\rangle$, respectively.

[1]The Pauli matrices are $\sigma_x = \left[\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right]$, $\sigma_y = \left[\begin{smallmatrix} 0 & -i \\ i & 0 \end{smallmatrix}\right]$, and $\sigma_z = \left[\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right]$.

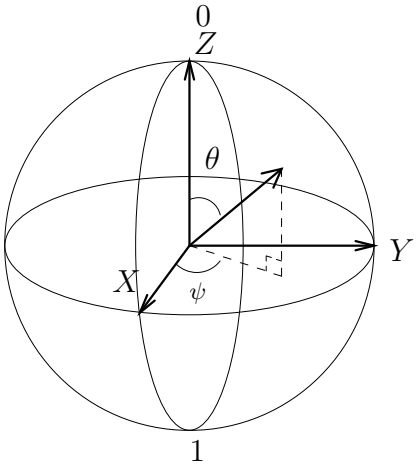


Fig. 2.4. The Bloch sphere representation of a quantum state.

2. Transformation of quantum states: $\mathbf{SU}(2)$ and $\mathbf{SO}(3)$

When a quantum operation is applied to a quantum system, it may change the quantum state of the system from one to another. The representation of the operation depends on how the quantum state is represented. For example, (2.7) leads to an operator or matrix U which connects the new and old states of a single spin quantum system:

$$|\phi'\rangle = U|\phi\rangle,$$

where $|\phi'\rangle$ and $|\phi\rangle$ are the quantum state after and before the operation, respectively. The fact that both states are unit vectors implies that U is a 2×2 unimodular complex matrix. More than that, U is also unitary, i.e., $U \in \mathbf{SU}(2)$[2], a Lie group endowed with a certain topology.

If the quantum state is represented by a three-dimensional Bloch vector, the effect

[2] $\mathbf{SU}(n)$ is the special unitary group of $n\times n$ matrices. An $n\times n$ matrix $A \in \mathbf{SU}(n)$ if and only if A is unitary, i.e., $A \cdot A^\dagger = I_n$, where $A^\dagger$ is the Hermitian adjoint of A, and $\det A = 1$.

of a unitary operation can be viewed as that of a rotation which rotates the Bloch sphere, and the operator is represented by a 3×3 real matrix S. If the quantum system has states $\boldsymbol{v}$ and $\boldsymbol{v}'$ in Bloch vector form before and after the operation, respectively, then

$$\boldsymbol{v}' = S\boldsymbol{v}.$$

The matrix S is a proper rotation matrix, i.e., $S \in \mathbf{SO}(3)$[3]. It is isometric and preserves the three-fold product.

If both S and U represent the same physical operation, such as a transformation induced by a series of pulses in NMR, there must be a connection between them. One can show that there is a mapping R from $\mathbf{SU}(2)$ to $\mathbf{SO}(3)$ such that $S = R(U)$, for any $U \in \mathbf{SU}(2)$ and its corresponding Bloch-sphere representation S [30]. Simple computation shows that the entry of matrix $S = R(U)$ at the k^{th} row and i^{th} column is given as

$$S_{ki} = Tr(\sigma_k\, U\, I_i\, U^\dagger), \tag{2.11}$$

where σ_k are the Pauli matrices, and Tr is the trace operator. It can also be shown that R is a two-to-one homomorphism between $\mathbf{SU}(2)$ and $\mathbf{SO}(3)$ with kernel $ker(R) = \{I, -I\}$. It coincides with the fact that U and $-U$ in $\mathbf{SU}(2)$ represent the same operation because only the relative phase matters. This mapping is also surjective, so it defines an isomorphism from the quotient group $\mathbf{SU}(2)/ker(R)$ to $\mathbf{SO}(3)$. We provide a more detailed discussion about this isomorphism in the Appendix.

It is known that any $U \in \mathbf{SU}(2)$ can be written into an exponential form param-

[3] $\mathbf{SO}(3)$ denotes the special orthogonal group of 3×3 matrices. An $n \times n$ matrix $A \in \mathbf{SO}(n)$ if and only if A is real, $AA^T = I_n$ and $\det A = 1$.

eterized by a angle $\theta \in [0, 2\pi)$ and a unit vector $\boldsymbol{n}$ such that

$$\begin{aligned} U(\theta, \boldsymbol{n}) &= e^{-i\frac{\theta}{2}\boldsymbol{n}\cdot\boldsymbol{\sigma}} \\ &= \begin{bmatrix} \cos\frac{\theta}{2} - in_3 \sin\frac{\theta}{2} & -\sin\frac{\theta}{2}(n_2 + in_1) \\ \sin\frac{\theta}{2}(n_2 - in_1) & \cos\frac{\theta}{2} + in_3 \sin\frac{\theta}{2} \end{bmatrix} \\ &= \cos\frac{\theta}{2} I - i \sin\frac{\theta}{2} \boldsymbol{n} \cdot \boldsymbol{\sigma}, \end{aligned} \tag{2.12}$$

where $\boldsymbol{\sigma} = [\sigma_x, \sigma_y, \sigma_z]$. With this parameterization of $\mathbf{SU}(2)$, entries of $S = R(U)$ can be computed using (2.11) as

$$S_{ij} = R(U)_{ij} = \cos\theta\, \delta_{ij} + (1 - \cos\theta) n_i n_j + \sum_{k=1}^{3} \sin\theta\, \epsilon_{ikj} n_k. \tag{2.13}$$

It should be noted now that S coincides with a rotation about the axis along $\boldsymbol{n}$ with an angle θ in the three dimensional Euclidean space after comparing S_{ij} with the standard formula of a rotation matrix. This interpretation is important in understanding the terminologies used in NMR. For example, the rotations around x, y, and z axes ($x/y/z$-rotations) with an arbitrary angle θ define the following three unitary operators in $\mathbf{SU}(2)$, respectively:

$$X_\theta = e^{-i\theta\sigma_x/2} = \begin{bmatrix} \cos\frac{\theta}{2} & -i\sin\frac{\theta}{2} \\ -i\sin\frac{\theta}{2} & \cos\frac{\theta}{2} \end{bmatrix}, \tag{2.14}$$

$$Y_\theta = e^{-i\theta\sigma_y/2} = \begin{bmatrix} \cos\frac{\theta}{2} & -\sin\frac{\theta}{2} \\ \sin\frac{\theta}{2} & \cos\frac{\theta}{2} \end{bmatrix}, \tag{2.15}$$

$$Z_\theta = e^{-i\theta\sigma_z/2} = \begin{bmatrix} e^{-i\theta/2} & 0 \\ 0 & e^{i\theta/2} \end{bmatrix}. \tag{2.16}$$

C. NMR quantum computer

NMR is an important tool in chemistry which has been in use for the determination of molecular structure and composition of solids, liquid and gases since the mid 1940s, by research groups in Stanford and MIT independently, led by F. Bloch and E.M. Purcell, both of whom shared the Nobel prize in physics in 1952 for the discovery.

There are many excellent monographs on NMR [31, 32, 33]. There are also many other nice internet website resources offering concise but highly useful information about NMR; e.g., [34, 35, 36]. Let us briefly explain the physics of NMR by following Edwards [34]. The NMR phenomenon is based on the fact that the spin of nuclei of atoms have magnetic properties that can be utilized to yield chemical, physical, and biological information. Through the famous Stern-Gerlach experiment in the earlier development of quantum mechanics, it is known that subatomic particles (protons, neutrons and electrons) have spins. Nuclei with spins behave like a bar magnet in a magnetic field. In some atoms, e.g., ^{12}C(carbon-12), ^{16}O (oxygen-16), ^{32}S(sulphur-32), these spins are paired and cancel each other out so that the nucleus of the atom has no overall spin. However, in many atoms (^{1}H,^{13}C, ^{31}P, ^{15}N, ^{19}F etc.) the nucleus does possess an overall spin. To determine the spin of a given nucleus one can use the following rules:

1. If the number of neutrons and the number of protons are both even, the nucleus has no spin.

2. If the number of neutrons plus the number of protons is odd, then the nucleus has a half-integer spin (i.e., 1/2, 3/2, 5/2).

3. If the number of neutrons and the number of protons are both odd, then the nucleus has an integer spin (i.e., 1, 2, 3).

In quantum mechanical terms, the nuclear magnetic moment of a nucleus can align with an externally applied magnetic field of strength B_0 in only $2I + 1$ ways, either with or against the applied field B_0, where I is the nuclear spin given in (i), (ii) and (iii) above. For example, for a single nucleus with $I = 1/2$, only one transition is possible between the two energy levels. The energetically preferred orientation has the magnetic moment aligned parallel with the applied field (spin $m = +1/2$) and is often denoted as α, whereas the higher energy anti-parallel orientation (spin $m = -1/2$) is denoted as β. See Fig. 2.5. In NMR quantum computing, these spin-up and spin-down quantum states resemble the two binary states 0 and 1 in a classical computer. Such a nuclear spin can serve as a quantum bit, or *qubit*. The rotational axis of the spinning nucleus cannot be orientated exactly parallel (or anti-parallel) with the direction of the applied field B_0 (aligned along the z axis) but must precess (motion similar to a gyroscope) about this field at an angle, with an angular velocity, ω_0, given by the expression $\omega_0 = \gamma B_0$. The precession rate ω_0 is called the Larmor frequency, see Fig. 2.6. See more discussion of ω_0 below. The constant γ is called the magnetogyric ratio. This precession process generates an magnetic field with frequency ω_0. If we irradiate the sample with radio waves (MHz), then the proton can absorb the energy and be promoted to the higher energy state. This absorption is called resonance because the frequencies of the applied radiation and the precession coincide at that frequency, leading to resonance.

1. More about the Hamiltonian of NMR

A classical way to explain NMR is to regard it as a rotating charged particle that acts like a current circulating in a loop ([37, 31]), which creates a magnet with magnetic moment μ, $\mu = qvr/2$, where q is the electronic charge. The particle is rotating at $v/2\pi r$ revolutions per second.

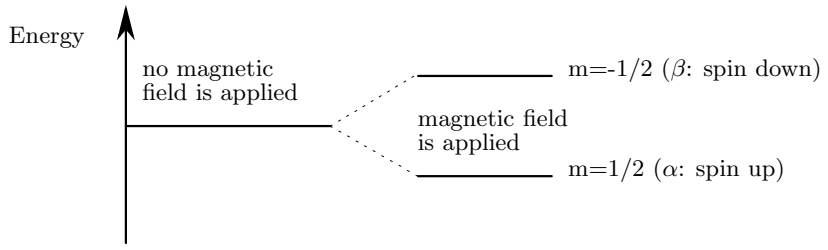


Fig. 2.5. Splitting of energy levels of a nucleus with spin quantum number 1/2.

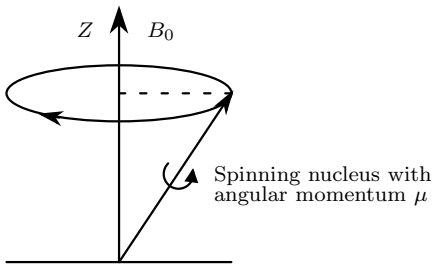


Fig. 2.6. A magnetic field B_0 is applied along the z-axis, causing the spinning nucleus to precess around the applied magnetic field.

Converting μ to electromagnetic units by dividing it by the velocity of light, and using angular momentum of the particle rather than the velocity of the particle, we obtain

$$\boldsymbol{\mu} = (q/2Mc)\boldsymbol{p},$$

where $\boldsymbol{p}$ is the angular momentum oriented along the rotating axis. The ratio μ/p is called the magnetogyric ratio, denoted by γ. A static magnetic field with strength B will apply a torque, which is equal to $\boldsymbol{\mu} \times \boldsymbol{B}$, on this particle. Newton's law states that the angular momentum will change according to a differential equation

$$\frac{d\boldsymbol{p}}{dt} = \boldsymbol{\mu} \times \boldsymbol{B} = \frac{q}{2Mc}\,\boldsymbol{p} \times \boldsymbol{B}.$$

Computation shows that $\boldsymbol{p}$ will rotate around the direction of $\boldsymbol{B}$ with frequency ω_0 defined by

$$\omega_0 = \frac{q}{2Mc}B.$$

The above is called the *Larmor equation*, and the frequency ω_0 is called the Larmor frequency, the precession frequency, or the resonance frequency as mentioned previously in Fig. 2.6.

The above classical considerations are now modified by *quantization* to incorporate the quantum-mechanical behaviors of the nuclear spin. The vector variable $\boldsymbol{p}$ is quantized with quantum number $(I(I+1))^{1/2}$, and its projection to z axis (the direction of the magnetic field) is $m\hbar$. In total, there are $2I+1$ valid values of m evenly distributed from $-I$ to I, i.e., $m = -I, -I+1, \cdots, I-1, I$. A factor g is introduced to include both the spin and orbital motion in the total angular momentum, called the Landé or spectroscopic splitting factor. For a free electron and proton, the magnetic momenta can be given as

$$\mu_e = \frac{g_e}{2}\left(\frac{he}{4\pi M_e c}\right) = \frac{g_e\beta}{2},$$
$$\mu_n = g_n I\left(\frac{he}{4\pi M_N c}\right) = g_n\, I\, \beta_N,$$

where $g_e = 2.0023$, $g_n = 5.58490$. Numbers β and β_N are called, respectively, the Bohr and the nucleus magneton where $\beta = 9.27 \times 10^{-21} erg\ gauss^{-1}$ and $\beta_N = 5.09 \times 10^{-24} erg\ gauss^{-1}$. These values vary for different particles. In NMR, it is convenient to use the resonance frequency ω_0:

$$\hbar\omega_0 = g_e\beta B_0,$$
$$\hbar\omega_0 = g_N I\, \beta_N B_0.$$

Now we can write the Hamiltonian of a free nucleus as

$$H = -\boldsymbol{\mu} \cdot \boldsymbol{B} = -\hbar\gamma \boldsymbol{I} \cdot \boldsymbol{B}, \tag{2.17}$$

where γ is the magnetogyric ratio defined by $\gamma = \dfrac{\mu}{I\hbar}$ just as in the classical case. It is a characteristic constant for every type of nuclei; different nuclei have different magnetogyric ratios. Vector $\boldsymbol{I}$ after quantization, becomes the operator of angular momentum. The eigenvalues of this system, or the energy levels are

$$E = \gamma\hbar m B, \quad m = -I, -I+1, \cdots, I-1, I. \tag{2.18}$$

The difference between two neighboring energy levels is $\gamma\hbar B$, which defines the resonance frequency depending on the magnetic field B and the particle.

There are other factors to be considered. The resonance frequency changes with the chemical environment of the nucleus. An example is the fluorine resonance spectrum of perfluorioisopropyl iodide. Two resonance lines of fluorine are observed in the spectrum, and the intensities ratio 6:1 agrees with the population ratio of the two groups of fluorine atoms. This phenomenon, called the *chemical shift*, is proportional to the strength of the magnetic field applied. This effect comes up because electrons close to the nucleus change the magnetic field around it; in other words, they create a diamagnetic shielding surrounding the nucleus. If the static field applied is B_0, then the electrons precessing around the magnetic field direction produce an induced magnetic field opposing B_0. The total effective magnetic field around the nucleus is then

$$\boldsymbol{B} = \boldsymbol{B}_0 - \boldsymbol{B}' = (1-\sigma)\boldsymbol{B}_0,$$

where the parameter σ is called *shielding coefficient*. In some cases σ is dependent on the temperature.

High resolution NMR spectroscopy has found that the chemical shifted peaks are also composed of several lines, a result of the spin-spin coupling, which is the second term in the NMR Hamiltonian:

$$H_{II} = \sum_{i>j} \boldsymbol{I_i} \cdot J_{ij} \cdot \boldsymbol{I_j}.$$

2. Realization of a qubit

As mentioned previously, NMR quantum computing is accomplished by using the spin-up and spin-down states of a spin-$\frac{1}{2}$ nucleus. A molecule with several nuclear spins may work as a quantum computer where each spin constitutes a *qubit*. In fact, NMR has a long history in information science. Back in the 1950s, nuclear spins were already used for information storage in computers.

Liquid NMR receives more interest due to its mature technology and readiness for application. For now, spin-$\frac{1}{2}$ nuclei such as proton and ^{13}C are preferred because they naturally represent a qubit, but multi-level qubits formed by spin-n nuclei, $n = 1, 2, \cdots$, may provide more freedom in the future. Through careful design, the potential qubits or nuclei are configured with different resonance frequencies and can be distinguished from each other. In a low viscosity liquid, dipolar coupling between nuclei is averaged away by the random motion of the molecules. The J-coupling (scalar coupling) dominates the spin-spin interaction, which is an indirect through-bond electronic interaction. Previously, a very difficult part of the system operation was to set the quantum system to a special state (or to initialize it). Now a very complicated technology has been developed to solve this problem [18].

Fig. 2.7 shows the structure of a trichloroethylene (TCE) molecule and a chloroform molecule used in NMR quantum computers. The hydrogen nucleus (proton) and two ^{13}C nuclei in a TCE molecule form three qubits which can be manipulated,

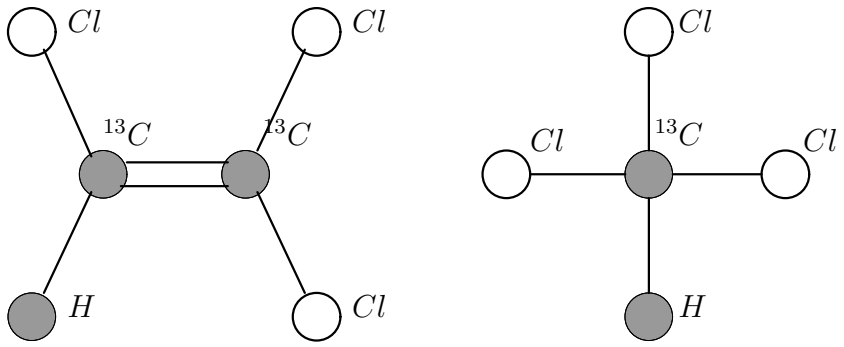


Fig. 2.7. The molecule structure of a candidate 3-qubit quantum system, trichloroethylene (left), and a candidate 2-qubit quantum system, chloroform. The trichloroethylene molecule has two labelled ^{13}C and a proton, all having one-half-spin nuclei. By considering the static magnetic field and spin-spin interaction, its Hamiltonian can be written as $H = -\sum_{i=1}^{3} g_{ni}\beta_{ni}\boldsymbol{I_i} \cdot \boldsymbol{B} + \sum_{i=1}^{2}\sum_{j=i+1}^{3} \boldsymbol{I_i} \cdot J_{i,j} \cdot \boldsymbol{I_j}$. The chloroform has one labelled ^{13}C and one proton.

while the chloroform molecule provides two qubits. The sample used by an NMR quantum computer has a large number ($\sim 10^{23}$) of such molecules. This is also called a *bulk* quantum computer. Although most molecules are in a totally random state at room temperature, there are still a small amount of spins standing out and serving our purpose. Theoretically, we use a statistical spin state called a pseudo-pure state, which has the same transformation property as that of a pure quantum state.

3. Construction of quantum gates

From Theorem A.1, we know that the collection of all the one-qubit gates and the two-qubit CNOT gate are universal. In addition, the following fact [38, p. 175] holds for one-qubit quantum gates:

Theorem C.1. *Suppose U is a unitary operation on a single qubit. Then there exist*

real numbers α, β, γ, and δ such that

$$U = e^{i\alpha} Z_\beta Y_\gamma Z_\delta.$$

For example, the Hadamard gate H can be decomposed as $H = e^{i\frac{\pi}{2}} Y_{\pi/2} Z_\pi$. Clearly, the $x/y/z$ rotation gates provide building blocks sufficient to construct any one qubit unitary gate. In this subsection, we will show how to realize these one-qubit rotation gates and the two-qubit CNOT gate using NMR. We will also show how to decouple the interaction between two spins, a process called *refocusing* [18, 38].

a. One-qubit gates

A single spin system has Hamiltonian $H = -\boldsymbol{\mu} \cdot \boldsymbol{B}$, where $\boldsymbol{\mu}$ is the magnetic moment, and

$$\boldsymbol{B} = B_0 \boldsymbol{e}_z + B_1(\boldsymbol{e}_x \cos(\omega t) + \boldsymbol{e}_y \sin(\omega t)) \tag{2.19}$$

is the magnetic field applied. B_0, a large constant, is the amplitude of the static magnetic field, and B_1 is the amplitude of the oscillating magnetic field in the x-y plane. When $B_1 = 0$, the Hamiltonian and Schrödinger equation can be obtained as ([38])

$$H = \frac{\omega_0}{2} \sigma_z \tag{2.20}$$

and

$$i\partial_t |\psi(t)\rangle = H|\psi(t)\rangle, \tag{2.21}$$

respectively, where $\hbar$ has been divided from both sides in the second equation and we take $\hbar$ away from H in the first one just for simplicity. The Larmor frequency $\omega_0 = -B_0\gamma$ is defined by the nuclei and the magnetic field, see (2.18). Assume that the initial state is $|\psi_0\rangle = a_0|0\rangle + b_0|1\rangle$. Then the evolution of the quantum state of

the spin and the density matrix can be solved directly and given as

$$\begin{aligned}
|\psi(t)\rangle &= e^{-i\omega_0\sigma_z t/2}|\psi_0\rangle \\
&= \begin{bmatrix} e^{-i\omega_0 t/2} & 0 \\ 0 & e^{i\omega_0 t/2} \end{bmatrix} \begin{bmatrix} a_0 \\ b_0 \end{bmatrix} \\
&= e^{-i\omega_0 t/2} \begin{bmatrix} 1 & 0 \\ 0 & e^{i\omega_0 t} \end{bmatrix} |\psi_0\rangle, \\
\rho(t) &= e^{-itH}\rho(0)e^{itH}.
\end{aligned}$$

This evolution is also called a *chemical shift evolution*, resembling the precessing of a magnet in a static field. Recall the Bloch vector on the Bloch sphere. It is exactly Z_θ, the rotation operator around the z axis with $\theta = \omega_0 t$.

To achieve an x-rotation operator, we need a small magnetic field transverse to the z direction to control the evolution of the quantum state. The Hamiltonian is given as in (2.19) by choosing B_1 different from zero:

$$H = -\boldsymbol{\mu} \cdot \boldsymbol{B} = \frac{\omega_0}{2}\sigma_z + \frac{\omega_1}{2}\left(\sigma_x \cos(\omega t) + \sigma_y \sin(\omega t)\right),$$

where ω_1 depends on the x-y plane component B_1 of the magnetic field, $\omega_1 = -B_1\gamma$. To solve the Schrödinger equation, we put $|\psi(t)\rangle$ in a "frame" rotating with the magnetic field around the z axis at frequency ω, $|\phi(t)\rangle = e^{i\omega t\sigma_z/2}|\psi(t)\rangle$. With this substitution, the Schrödinger equation (2.21) becomes

$$i\partial_t|\phi(t)\rangle = (e^{i\omega\sigma_z t/2}He^{-i\omega\sigma_z t/2} - \frac{\omega}{2}\sigma_z)|\phi(t)\rangle. \tag{2.22}$$

Using properties

$$\begin{aligned} e^{i\omega\sigma_z t/2}\sigma_z e^{-i\omega\sigma_z t/2} &= \sigma_z, \\ e^{i\omega\sigma_z t/2}\sigma_x e^{-i\omega\sigma_z t/2} &= \sigma_x\cos(\omega t) - \sigma_y\sin(\omega t), \\ e^{i\omega\sigma_z t/2}\sigma_y e^{-i\omega\sigma_z t/2} &= \sigma_x\sin(\omega t) + \sigma_y\cos(\omega t), \end{aligned} \tag{2.23}$$

we obtain

$$\begin{aligned} i\partial_t|\phi(t)\rangle &= \left(\frac{\omega_0-\omega}{2}\sigma_z + \frac{\omega_1}{2}\sigma_x\right)|\phi(t)\rangle, \\ |\phi(t)\rangle &= e^{-i((\omega_0-\omega)\sigma_z/2+\omega_1\sigma_x/2)t}|\phi(0)\rangle. \end{aligned} \tag{2.24}$$

We know from (2.12) that this is a rotation around the axis

$$\boldsymbol{n} = \frac{1}{\sqrt{1+(\frac{\omega_1}{\omega_0-\omega})^2}}\left(\boldsymbol{z} + \frac{\omega_1}{\omega_0-\omega}\boldsymbol{x}\right). \tag{2.25}$$

An important case is $\omega_0 = \omega$, also called the *resonance* case where its name came from the zero denomination in (2.25). By (2.24), we see that a relatively weak transverse magnetic field causes a rotation around the x axis:

$$|\psi(t)\rangle = e^{-i\omega_0\sigma_z t/2}|\phi(t)\rangle = e^{-i\omega_0 t\sigma_z/2}e^{-i\omega_1 t\sigma_x/2}|\phi(0)\rangle = Z_\theta X_\beta|\psi(0)\rangle, \tag{2.26}$$

where $X_\beta = e^{-i\omega_1 t\sigma_x/2}$, $\beta = \omega_1 t$. By applying another $Z_{-\theta}$, we obtain a rotation X_β as desired. Since the frequency of the precession is in radio frequency band, the field applied is called an *RF pulse*.

When $|\omega_0 - \omega| \gg \omega_1$, the rotation axis direction is almost along $\boldsymbol{z}$ and the RF pulse has no effect on it:

$$|\psi(t)\rangle = e^{-i\omega\sigma_z t/2}|\phi(t)\rangle \approx e^{-i\omega_0 t\sigma_z/2}|\psi(0)\rangle = Z_{\omega_0 t}|\psi(0)\rangle,$$

thus we can tell one qubit from another because their resonance frequencies are de-

signed to be different. There are still cases where the difference of resonance frequencies between spins is not large enough. The RF pulse may cause similar rotations on all those spins. To avoid or at least minimize it, a *soft* pulse is applied instead of the so called *hard* pulse. It is a pulse with longer time span and weaker magnetic field, in other word, a smaller ω_1. This strategy makes these "close" qubits fall into the $|\omega_0 - \omega| \gg \omega_1$ case.

If we change the magnetic field to

$$\boldsymbol{B} = B_0 \boldsymbol{e}_z + B_1(\boldsymbol{e}_x \cos(\omega_0 t + \alpha) + \boldsymbol{e}_y \sin(\omega_0 t + \alpha)), \tag{2.27}$$

the Hamiltonian will become

$$H = \frac{\omega_0}{2}\sigma_z + \frac{\omega_1}{2}(\sigma_x \cos(\omega_0 t + \alpha) + \sigma_y \sin(\omega_0 + \alpha)) \tag{2.28}$$

where ω_1 is defined as before. The RF field is almost the same as (2.19) in the resonance case except a phase shift. Using the same rotation frame as before with $\omega = \omega_0$, we obtain

$$i\partial_t|\phi(t)\rangle = \frac{\omega_1}{2}(\sigma_x \cos(\alpha) + \sigma_y \sin(\alpha))|\phi(t)\rangle, \tag{2.29}$$

after simplification. After time duration t, the new system state is given as

$$|\phi(t)\rangle = e^{-i\frac{\omega_1}{2}(\sigma_x \cos(\alpha)+\sigma_y \sin(\alpha))t}|\phi(0)\rangle, \tag{2.30}$$

and the evolution operator can be computed using (2.12) as

$$\begin{aligned} U_{\theta/2,\alpha} &= e^{-i\frac{\omega_1}{2}(\sigma_x \cos(\alpha)+\sigma_y \sin(\alpha))t} \\ &= \begin{bmatrix} \cos(\frac{\theta}{2}) & -i\sin(\frac{\theta}{2})e^{-i\alpha} \\ -i\sin(\frac{\theta}{2})e^{i\alpha} & \cos(\frac{\theta}{2}) \end{bmatrix}, \end{aligned} \tag{2.31}$$

where $\theta = \omega_1 t$. This is a one-qubit rotation operator, and sometimes is called a Rabi

rotation gate. When $\alpha = \pi/2$,

$$\begin{aligned} U_{\theta/2,\pi/2} &= \begin{bmatrix} \cos(\frac{\theta}{2}) & -\sin(\frac{\theta}{2}) \\ \sin(\frac{\theta}{2}) & \cos(\frac{\theta}{2}) \end{bmatrix} \\ &= Y_\theta. \end{aligned} \tag{2.32}$$

We have achieved a y-rotation operator just by adding a phase shift to the RF field.

b. Two-qubit gates

The construction of a two-qubit gate requires the coupling of two spins. In a liquid sample of NMR, *J-coupling* is the dominating coupling between spins. Under the assumption that the resonance frequency difference between the coupled spins is much larger than the strength of the coupling (a so-called weak coupling regime), the total Hamiltonian of a two spin system without transverse field may be given as

$$H = \frac{1}{2}\omega_1\sigma_z^1 + \frac{1}{2}\omega_2\sigma_z^2 + \frac{1}{2}J\sigma_z^1\sigma_z^2, \tag{2.33}$$

where ω_i is the frequency corresponding to spin i, σ_z^i is the z projection operator of spin i, for $i = 1, 2$, and J is the coupling coefficient. Take the chloroform in Fig. 2.7 for example [39, 29]. In a 11.7T magnetic field, the precession frequency of ^{13}C is about $2\pi \times 500$MHz and the precession frequency of proton is about $2\pi \times 125$MHz. The coupling constant J is about $2\pi \times 100$Hz. Here we set $B_1 = 0$, which means no transverse magnetic field is applied and those terms such as σ_x, σ_y do not appear. The remaining terms in the Hamiltonian only contains operators σ_z^1 or σ_z^2, which are commutative. Thus, we can obtain the eigenstates and eigenvalues of this two-spin system and we map the set of eigenstates to the standard basis of $\mathbf{C}^4$, as follows:

$$|00\rangle = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix},\ |01\rangle = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix},\ |10\rangle = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix},\ |00\rangle = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}; \tag{2.34}$$

$$\begin{aligned} H|00\rangle &= k_{00}|00\rangle, \quad k_{00} = \frac{1}{2}\omega_1 + \frac{1}{2}\omega_2 + \frac{1}{2}J; \\ H|01\rangle &= k_{01}|01\rangle, \quad k_{01} = \frac{1}{2}\omega_1 - \frac{1}{2}\omega_2 - \frac{1}{2}J; \\ H|10\rangle &= k_{10}|10\rangle, \quad k_{10} = -\frac{1}{2}\omega_1 + \frac{1}{2}\omega_2 - \frac{1}{2}J; \\ H|11\rangle &= k_{11}|11\rangle, \quad k_{11} = -\frac{1}{2}\omega_1 - \frac{1}{2}\omega_2 + \frac{1}{2}J. \end{aligned} \tag{2.35}$$

Since the matrix is diagonal, the evolution of this two-spin system can be easily derived as

$$|\psi(t)\rangle = e^{-iHt}|\psi(0)\rangle = \begin{bmatrix} e^{-ik_{00}t} & & & \\ & e^{-ik_{01}t} & & \\ & & e^{-ik_{10}t} & \\ & & & e^{-ik_{11}t} \end{bmatrix} |\psi(0)\rangle. \tag{2.36}$$

We can also rewrite the one qubit rotation operators for this two $\frac{1}{2}$-spin system in matrix form with respect to the same basis:

$$Z^1_{\pi/2} = \begin{bmatrix} e^{-i\pi/4} & & & \\ & e^{-i\pi/4} & & \\ & & e^{i\pi/4} & \\ & & & e^{i\pi/4} \end{bmatrix}, \tag{2.37}$$

$$Z^2_{-\pi/2} = \begin{bmatrix} e^{i\pi/4} & & & \\ & e^{-i\pi/4} & & \\ & & e^{i\pi/4} & \\ & & & e^{-i\pi/4} \end{bmatrix}, \tag{2.38}$$

$$Y^2_{\pi/2} = \frac{\sqrt{2}}{2} \begin{bmatrix} 1 & -1 & & \\ 1 & 1 & & \\ & & 1 & -1 \\ & & 1 & 1 \end{bmatrix}, \tag{2.39}$$

$$Y^2_{-\pi/2} = \frac{\sqrt{2}}{2} \begin{bmatrix} 1 & 1 & & \\ -1 & 1 & & \\ & & 1 & 1 \\ & & -1 & 1 \end{bmatrix}, \tag{2.40}$$

where Z^i_θ is the rotation operator for spin i with angle θ around the z axis while keeping another spin unchanged, and all Y^i_θ are similarly defined operators about the y axis; see (2.15). A careful reader may raise issues about the one-qubit gate we have obtained in subsection a because the coupling between two qubits always exists and has not been considered. We need to turn off the coupling when we only want to operate one spin but the coupling is non-negligible. This is in fact one of the major characteristic difficulties associated with the NMR quantum computing technology. A special technology called *refocusing* is useful. It works as follows. We apply a soft π pulse on the spare spin that we don't want to change at the middle point of the operation time duration while we are working on the target spin. The effect is that the coupling before the pulse cancels the one after the pulse, so the result of no-coupling is achieved. Another π pulse will be needed to turn the spin back. All pulses are soft.

This technology is so important that we now state it here as a theorem.

Theorem C.2. *Let* $H = \frac{\omega_1}{2}\sigma_z^1 + \frac{J}{2}\sigma_z^1\sigma_z^2 + A$ *be a given Hamiltonian, where* A *is a Hamiltonian that does not act on spin 1 and commutes with* σ_z^2*. Then the evolution operators of* A *and* H *satisfy*

$$e^{-iAt} = -X_\pi^1 e^{-iHt/2} X_\pi^1 e^{-iHt/2}, \tag{2.41}$$

i.e., the collective evolution of the quantum system with Hamiltonian H *and additional two* X_π^1*-pulses at the middle and the end of the time duration, equals that of a system with Hamiltonian* A *(up to a global phase shift* π*, or a factor* -1*).*

Proof. Assume that the time duration is t and denote U for

$$U = X_\pi^1 e^{-iHt/2} X_\pi^1 e^{-iHt/2}. \tag{2.42}$$

Note that $X_\pi^1 = e^{-i\frac{\pi}{2}\sigma_x^1}$ and it commutes with A which contains no operators acting on spin 1, thus

$$U = X_\pi^1 e^{-i(\frac{\omega_1}{2}\sigma_z^1 + \frac{J}{2}\sigma_z^1\sigma_z^2)\frac{t}{2}} X_\pi^1 e^{-i(\frac{\omega_1}{2}\sigma_z^1 + \frac{J}{2}\sigma_z^1\sigma_z^2)\frac{t}{2}} e^{-iAt}. \tag{2.43}$$

It suffices to prove that the part before e^{-iAt} satisfies

$$B = X_\pi^1 e^{-i(\frac{\omega_1}{2}\sigma_z^1 + \frac{J}{2}\sigma_z^1\sigma_z^2)\frac{t}{2}} X_\pi^1 e^{-i(\frac{\omega_1}{2}\sigma_z^1 + \frac{J}{2}\sigma_z^1\sigma_z^2)\frac{t}{2}} = -I. \tag{2.44}$$

We first check the effect of B on the four basis vector. We have

$$\begin{aligned} B|11\rangle &= X_\pi^1 e^{-i(\frac{\omega_1}{2}\sigma_z^1+\frac{J}{2}\sigma_z^1\sigma_z^2)\frac{t}{2}} X_\pi^1 e^{-i(\frac{\omega_1}{2}\sigma_z^1+\frac{J}{2}\sigma_z^1\sigma_z^2)\frac{t}{2}}|11\rangle \\ &= e^{-i\frac{-\omega_1+J}{4}t}(-i)X_\pi^1 e^{-i(\frac{\omega_1}{2}\sigma_z^1+\frac{J}{2}\sigma_z^1\sigma_z^2)\frac{t}{2}}|01\rangle \\ &= (-i)e^{-i\frac{-\omega_1+J}{4}t}X_\pi^1 e^{-i\frac{\omega_1-J}{4}t}|01\rangle \\ &= (-i)^2|11\rangle \\ &= -|11\rangle, \end{aligned} \tag{2.45}$$

$$\begin{aligned} B|01\rangle &= X_\pi^1 e^{-i(\frac{\omega_1}{2}\sigma_z^1+\frac{J}{2}\sigma_z^1\sigma_z^2)\frac{t}{2}} X_\pi^1 e^{-i(\frac{\omega_1}{2}\sigma_z^1+\frac{J}{2}\sigma_z^1\sigma_z^2)\frac{t}{2}}|01\rangle \\ &= e^{-i\frac{\omega_1-J}{4}t}(-i)X_\pi^1 e^{-i(\frac{\omega_1}{2}\sigma_z^1+\frac{J}{2}\sigma_z^1\sigma_z^2)\frac{t}{2}}|11\rangle \\ &= (-i)e^{-i\frac{-\omega_1+J}{4}t}X_\pi^1 e^{-i\frac{-\omega_1+J}{4}t}|11\rangle \\ &= (-i)^2|01\rangle \\ &= -|01\rangle, \end{aligned} \tag{2.46}$$

and similarly,

$$\begin{aligned} B|10\rangle &= -|10\rangle, \\ B|00\rangle &= -|00\rangle. \end{aligned} \tag{2.47}$$

In the computation above, we have used the fact that X_π^1 has no effect on the second spin and the four basis vectors $|00\rangle$, $|01\rangle$, $|10\rangle$ and $|11\rangle$ are the eigenstates of the operator $\frac{\omega_1}{2}\sigma_z^1 + \frac{J}{2}\sigma_z^2\sigma_z^1$. The result shows that $B = -I$, and we are done. □ □

When the Hamiltonian is given in the form as (2.33), the above theorem tells us that both the chemical shift evolution (precession) and the J-coupling effect on spin 1 are removed and only the term $\frac{\omega_2}{2}\sigma_z^2$ remains. We obtain a z-rotation of spin 2 while freezing spin 1. By combining it with several hard pulses, we can also achieve any arbitrary rotation on spin 2 with the motion of spin 1 frozen [40]. Similar computation shows that a hard π pulse applied at the middle point of the time duration cancels

the chemical shift evolution of both spins. This can be seen by checking the identity

$$e^{-iHt/2}X_\pi^1 X_\pi^2 e^{-iHt/2} = \begin{bmatrix} & & & e^{-iJt/2} \\ & & e^{iJt/2} & \\ & e^{iJt/2} & & \\ e^{-iJt/2} & & & \end{bmatrix}. \tag{2.48}$$

Another hard π pulse can rotate two spins back, so we have achieved an evolution which has only the J-coupling effect, denoted by $\mathbb{Z}_\theta$:

$$\mathbb{Z}_\theta = \begin{bmatrix} e^{-i\theta/2} & & & \\ & e^{i\theta/2} & & \\ & & e^{i\theta/2} & \\ & & & e^{-i\theta/2} \end{bmatrix},$$

and when $\theta = \pi/2$,

$$\mathbb{Z}_{\pi/2} = \begin{bmatrix} e^{-i\pi/4} & & & \\ & e^{i\pi/4} & & \\ & & e^{i\pi/4} & \\ & & & e^{-i\pi/4} \end{bmatrix}. \tag{2.49}$$

Although we give only an example of the 2-qubit system in the above, the reader should note that a general method is available to reserve only the couplings wanted while keeping all the others cancelled for multi-qubit systems [40, 41, 42]. Combining operators in (2.37) through (2.40) and (2.49), we can now construct a CNOT gate as in Fig. 2.8 which includes four one-qubit $\pi/2$ rotations around y or z axes and one

two-qubit $\pi/2$ rotation. The total operator, denoted by CN, can be computed as

$$CN = Z^1_{\pi/2} Y^2_{-\pi/2} Z^2_{-\pi/2} \mathbb{Z}_{\pi/2} Y^2_{\pi/2} = e^{-\frac{\pi}{4}i} \begin{bmatrix} 1 & & & \\ & 1 & & \\ & & 0 & 1 \\ & & 1 & 0 \end{bmatrix}, \tag{2.50}$$

which is a CNOT gate up to a phase of $-\pi/4$ [29].

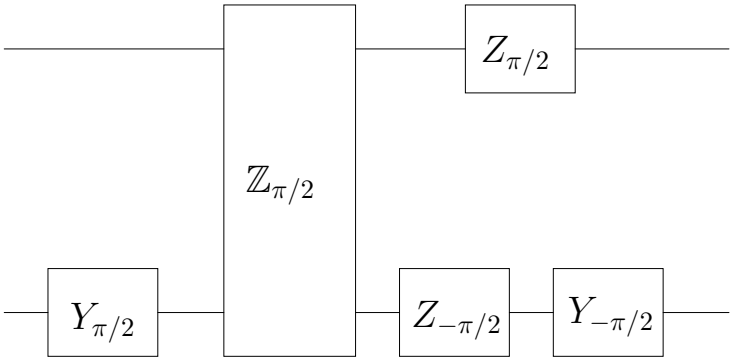


Fig. 2.8. The quantum circuit used to realize a quantum controlled-not gate.

We have shown how to construct one-qubit gates and the two-qubit CNOT gate using the NMR technology. The simple pulse design works fine in ideal situations. In practice, errors arise from various factors. Decoherence causes the lost of quantum information with time. Thus, all operations should be completed within a short time, roughly constrained by the energy relaxation time T_1 and the phase randomization time T_2. Again, take the chloroform for an example. For protons, $T_1 \approx$ 7sec and $T_2 \approx$ 2sec; for carbons, $T_1 \approx$ 16sec and $T_2 \approx$ 0.2sec [29, 39]. The pulses have to be short enough so that all the pulses can be jammed in the time window. Ideally, a pulse can be completed quite fast, but this may incur undesirable rotations in

other qubits because the frequency band width is inversely proportional to the time length of the pulse. A shorter and stronger pulse will have a wider frequency band that may cover the resonance frequency of another spin, called *cross-talking*. It should also be noted that both T_1 and T_2 are defined and measured in simplified situation, and they can only be used as an approximation of the decoherence rate for the quantum computation. Coupling is also a problem which makes the pulse design much more complicated. Finally, any experimental facility is not perfect, which may introduce more errors. Typical error resources include *inhomogeneities in the static and RF field*, *pulse length calibration errors*, *frequency offsets*, and *pulse timing/phase imperfections*.

If the quantum circuit can be simplified and the number of gates needed is reduced, the requirements on the pulses can be alleviated. Mathematicians are looking for methods to find time-optimal pulse sequences [43, 44, 45, 46], with the goal of finding the shortest path between the identity and a point in the space of $\mathbf{SU}(n)$ allowed by the system and the control Hamiltonians. Besides that, NMR spectroscopists have already developed advanced pulse techniques to deal with system errors such as cross-talking and coupling. They turn out to work well and are now widely used in NMR quantum computation. Such techniques include composite pulses [47, 48, 49, 50, 51, 53] and pulse shaping. The latter consists mainly of two methods: phase profiles [54] and amplitude profiles [55, 56].

c. Initialization

An NMR sample eventually will go into its equilibrium state when no RF pulse is applied for a long time. Then the density matrix is proportional to $e^{-H/kT}$, according to the Boltzmann distribution, where $k = 1.381 \times 10^{-23}$J/K and T is the absolute temperature. Normally, the environment temperature is far larger than the energy

difference between the up and down states of the spin, and H/kT is very small, about 10^{-4}. We also make the assumption that the coupling terms are small enough compared with the resonant frequency, thus we can make a reasonable approximation of the equilibrium state density matrix of a system with n spins:

$$\rho_{eq} = \frac{e^{-H/kT}}{tr(e^{-H/kT})} \approx I - \frac{1}{kT}(\epsilon_1\sigma_z^1 + \epsilon_2\sigma_z^2 + \cdots + \epsilon_n\sigma_z^n). \tag{2.51}$$

In the four operators appearing in the density matrix (2.8), only those with zero traces can be observed in NMR. The operator I_0 is invisible, and moreover, it remains invariant under any unitary similarity transformation. Therefor, we only need to take care of the zero-trace part of the initial density matrix, noting that only that part (called deviation) is effective. Most algorithms prefer an initial state such as

$$\rho_0 = \frac{1-\epsilon}{2^n} I + \epsilon|00\cdots0\rangle\langle0\cdots00|,$$

which is an example of the so called *pseudo-pure* states, corresponding to the pure state $|00\cdots0\rangle$.

To initialize the system to a pseudo-pure state as above, we may use a scheme called *averaging*. Let us explain this for a 2-spin system. Suppose we have three 2-spin subsystems with density matrices

$$\rho_1 = \begin{bmatrix} a & 0 & 0 & 0 \\ 0 & b & 0 & 0 \\ 0 & 0 & c & 0 \\ 0 & 0 & 0 & d \end{bmatrix}, \quad \rho_2 = \begin{bmatrix} a & 0 & 0 & 0 \\ 0 & c & 0 & 0 \\ 0 & 0 & d & 0 \\ 0 & 0 & 0 & b \end{bmatrix}, \quad \rho_3 = \begin{bmatrix} a & 0 & 0 & 0 \\ 0 & d & 0 & 0 \\ 0 & 0 & b & 0 \\ 0 & 0 & 0 & c \end{bmatrix}, \tag{2.52}$$

respectively, where a, b, c, and d are nonnegative, and $a+b+c+d=1$. These are three diagonal matrices with three of their diagonal elements in cyclic permutation.

Now, we mix these three subsystems together (for n-qubit system, we may have

$2^n - 1$ subsystems) and assume that the three subsystems have the same signal scale. Because the readout is linear with respect to the initial state, we are in fact working on a system with an effective initial density matrix

$$\frac{1}{3}\sum_{i=1}^{3}\rho_i = \frac{1}{3}\begin{bmatrix} 3a & & & \\ & b+c+d & & \\ & & b+c+d & \\ & & & b+c+d \end{bmatrix}$$
$$= \frac{b+c+d}{3}I + \frac{1}{3}\begin{bmatrix} 4a-1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \tag{2.53}$$

which is a pseudo-pure state corresponding to $|00\cdots0\rangle$.

Various methods have been developed to achieve this effect of averaging. Because ρ_1, ρ_2, and ρ_3 differ only by a permutation of the diagonal elements, a sequence of CNOT pulses can be used to transform one to another. In most cases, we only have one sample, the same algorithm can be repeated on the very sample three times but with different initial states ρ_1, ρ_2, and ρ_3, respectively. At last, after all the three outputs are obtained and added together (average), we achieve the same result as what we will get when the algorithm is employed on a system with the expected initial state $|00\cdots0\rangle$. This is called "temporal averaging" [57]. Gradient fields can also be used to divide the sample into different slices in space which are prepared into different initial states, and the averaging is realized spatially, called "spatial averaging" [58]. The number of the experiments and pulses needed grows very large when the number of qubits increases. For example, 9 experiments are combined in order to prepare one pseudo-pure state for a 5-qubit system and 48 pulses are used

to form one pseudo-pure state in a 7-qubit system [59] after modifications such as logical labeling [60, 61] and selective saturation [62].

d. Measurement

An NMR computer differs from other quantum computers in that it works on an ensemble of spins instead of just a single one. It produces an observable macroscopic signal which can be picked up by a set of coils positioned on the x-y plane. The signal measures the change rate of the magnetic field created by a large number of spins in the sample rotating around the z-axis, called *free induction decay* (FID). Due to relaxation, peaks of the Fourier transform of the signal, or spectra, have width. However, we do not need to worry about that since it will not make any substantial difference in our discussion here. One disadvantage is that the readout from NMR is an average of all the possible states, in contrast to most existing quantum algorithms that ask for the occurrence of only a single state. But it is possible for one to modify ordinary quantum algorithms to make NMR results usable [18].

The magnetization detected by the coil is proportional to the trace of the product of the density matrix with $\sigma_+ = \sigma_x + i\sigma_y$:

$$M_x + iM_y = nV\langle \mu_x + i\mu_y \rangle = nV\gamma\hbar Tr(\rho(\sigma_x + i\sigma_y)), \tag{2.54}$$

where γ is the magnetogyric ratio as in (2.18) and ρ is the density matrix. When the external RF magnetic field is removed, the density matrix will change according to the system's Hamiltonian as we discussed earlier. If we decompose the density matrix into a sum of product operators as in (2.8), only I_x and I_y contribute to the readout. We can not "see" the coefficients of I_0 and I_z. Recall (2.23): if a one-spin system begins from density matrix $\rho_0 = I_0 + \sin\theta\cos\psi I_x + \sin\theta\sin\psi I_y + \cos\theta I_z$, the

magnetization will rotate with the resonant frequency as

$$\begin{aligned} M_z + iM_y &= C\,Tr(e^{-iHt}\rho_0 e^{iHt}\sigma_+) \\ &= C\,Tr(e^{-iHt}(I_0 + \sin(\theta)\cos(\psi)I_x + \\ &\quad \sin(\theta)\sin(\psi)I_y + \cos(\theta)I_z)e^{iHt}\sigma_+) \\ &= C\,Tr((\sin\theta\cos\psi(\cos(\omega t)I_x + \sin(\omega t)I_y) + \\ &\quad sin\theta\sin\psi(\cos(\omega t)I_y - \sin(\omega t)I_x))\sigma_+) \\ &= C\ \sin\theta\ e^{i(\omega t+\psi)}, \end{aligned} \tag{2.55}$$

where $C = nV\gamma\hbar$. This rotating magnetization will introduce an oscillating electric potential in the receiver coils, which will be processed by a computer to generate the spectra. Note that the signal is proportional to $\sin\theta$. If an x rotation with angle $\pi/2$ is applied on the spin before the measurement, the magnetization will become

$$M_z + iM_y = \frac{\sqrt{2}}{2}C(\sin\theta\ -\ i\cos\theta)e^{i\omega t}.$$

For simplicity, we have chosen $\psi = 0$. The imaginary part is proportional to the population difference:

$$\cos\theta = \cos^2\frac{\theta}{2} - \sin^2\frac{\theta}{2}.$$

Computation of a two-spin system is complicated, so we will only give some partial results here. The purpose is to point out what methodology is used. We will still use the basis given by (2.34) and the Hamiltonian in (2.33). The system begins from a density matrix as

$$\rho_0 = \begin{bmatrix} \rho_{11} & \rho_{12} & \rho_{13} & \rho_{14} \\ \rho_{21} & \rho_{22} & \rho_{23} & \rho_{24} \\ \rho_{31} & \rho_{32} & \rho_{33} & \rho_{34} \\ \rho_{41} & \rho_{42} & \rho_{43} & \rho_{44} \end{bmatrix}. \tag{2.56}$$

The operator σ_+ is a summation of operators from the two subsystems:

$$\begin{aligned} \sigma_+ &= \sigma_+^1 + \sigma_+^2 \\ &= \begin{bmatrix} 0 & 2 & 2 & 0 \\ 0 & 0 & 0 & 2 \\ 0 & 0 & 0 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix}. \end{aligned} \tag{2.57}$$

The magnetization in the x-y plane is composed of four frequencies:

$$\begin{aligned} M_x + iM_y &= C\, Tr(e^{-iHt}\rho_0 e^{iHt}\sigma_+) \\ &= C\,(\rho_{31}e^{i(\omega_1+J)t} + \rho_{42}e^{i(\omega_1-J)t} + \rho_{43}e^{i(\omega_2-J)t} + \rho_{21}e^{i(\omega_2+J)t}). \end{aligned} \tag{2.58}$$

The spectrum has two pairs of peaks, one pair around the precession frequency ω_1, another pair around ω_2. See Fig. 2.9. The splitting is a result of coupling. If the system have more than two spins, the coupling will split up a peak into up to 2^{n-1} peaks where n is the number of spins. We also combine all the constants in C to make the formula concise. Only four of the elements out of the density matrix appear in this spectrum, so we need to design certain control pulses to move the expected information to these four positions where numbers can be shown via free induction signal. If multi-tests are allowed, theoretically, all the elements of the density matrix can be retrieved [63, 64]. It is also possible to transport the desired information (computational results) to the four positions where the observer can see.

A typical pulse used in reading out is a hard $X_{\pi/2}$ pulse which rotate all the spins about the x-axis with angle $\pi/2$. Let us still use two-spin systems as an example. The operation is the tensor product of two x-rotation operators, i.e., $X_{\pi/2} = X^1_{\pi/2}X^2_{\pi/2}$. The imaginary part of the four effective elements of the density matrix ρ' after the

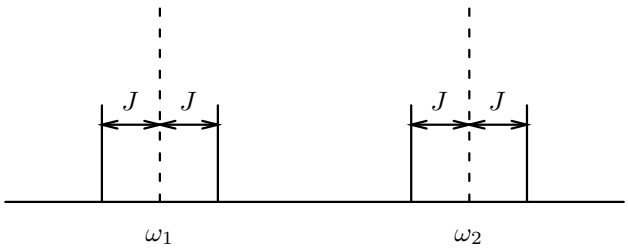


Fig. 2.9. Simplified stick spectra of a two-qubit molecule. The two dotted lines show two peaks at ω_1 and ω_2, respectively, when no coupling is applied ($J = 0$). After coupling, every peak is split into two small peaks with the intensities reduced to half.

operation, utilizing the fact that the density matrix is Hermitian, are

$$\begin{aligned} Im(\rho'_{31}) &= \frac{1}{4}(\rho_{33} + \rho_{44} - \rho_{11} - \rho_{22} - 2Im(\rho_{21}) - 2Im(\rho_{34})), \\ Im(\rho'_{42}) &= \frac{1}{4}(\rho_{33} + \rho_{44} - \rho_{11} - \rho_{22} + 2Im(\rho_{21}) + 2Im(\rho_{34})), \\ Im(\rho'_{43}) &= \frac{1}{4}(\rho_{22} + \rho_{44} - \rho_{11} - \rho_{33} + 2Im(\rho_{31}) + 2Im(\rho_{24})), \\ Im(\rho'_{21}) &= \frac{1}{4}(\rho_{22} + \rho_{44} - \rho_{11} - \rho_{33} - 2Im(\rho_{31}) - 2Im(\rho_{24})). \end{aligned} \tag{2.59}$$

Find the sum of $Im(\rho'_{31})$ and $Im(\rho'_{42})$ and that of $Im(\rho'_{43})$ and $Im(\rho'_{21})$:

$$\begin{aligned} Im(\rho'_{31} + \rho'_{42}) &= -\frac{1}{2}(\rho_{11} + \rho_{22} - \rho_{33} - \rho_{44}), \\ Im(\rho'_{43} + \rho'_{21}) &= -\frac{1}{2}(\rho_{11} - \rho_{22} + \rho_{33} - \rho_{44}). \end{aligned} \tag{2.60}$$

Because what the coils pick up is the change rate of the magnetic field rather than the magnetic field itself, the imaginary part we have listed above is reflected in the real part of the spectra. The computation above shows that the sum of the real parts of each pair of peaks in the spectra is proportional to the population difference between the spin-up and the spin-down states of the corresponding spin.

D. Shor's algorithm and its experimental realization

Through the rest of the paper, we will describe two applications of the NMR quantum computer: the Shor's algorithm and a lattice algorithm [18]. Entangled states are extremely important in quantum computation. Entanglement, together with superposition, gives a quantum computer the power to perform massively parallel computation and thus makes it particularly suitable for computing certain complex problems. Shor's algorithm for the factorization of integers aimed at decryption is a special example of a "killer ap" of quantum computing [2, 4]. Recently, a successful experiment has shown the potential capability of the implementation of Shor's algorithm, although it is still very simple and tentative. In [5], Vandersypen et al. factor 15 into 3 times 5. That work has demonstrated the liquid NMR quantum computer to be the most successful quantum computer so far.

1. Shor's algorithm

It is not difficult to factor a composite integer (i.e., non-prime) into prime numbers when that integer is small, but the computation burden grows rapidly when the number increases. The currently most efficient algorithm, the number field sieve, requires an exponential running time $e^{c(\log n)^{1/3}(\log\log n)^{2/3}}$, where n is the number to be factored and clearly $\log n$ is proportional to the number of the bits needed to store this number. This makes it practically impossible to factor a large number using a classical computer. This difficulty is used to construct several cryptosystems, such as the RSA public key cryptosystem [6]. Peter W. Shor has shown that this problem can be solved in polynomial running time instead of exponential time by using the quantum computer. A more accessible account of Shor's algorithm is given by Lomonaco [7].

Let n be an odd integer to be factored, and choose another random integer x less than n. We require x to be coprime with n; otherwise, we find a factor of n immediately by the Euclidean method. It is then known that function $f(s) = x^s \mod n$ is periodic. The period of f (and also of x) is the smallest integer r such that $x^r = 1 \mod n$. For example, when $n = 15$ and $x = 3$, the moduli of x^s, with s being 1, 2, 3, ..., are 3, 9, 12, 6, 3, 9, 12, 6, ..., and the period is 4.

Now we check the period r. If r is even, $r = 2t$, then $x^{2t} - 1 = (x^t + 1)(x^t - 1) = 0 \mod n$, so either $x^t - 1$ or $x^t + 1$ has a common factor with n. A classical computer can use the Euclidean algorithm to compute the greatest common divisors, denoted as $\gcd(x^t + 1, n)$ and $\gcd(x^t - 1, n)$, in polynomial time. It is possible that we only obtain the trivial factors 1 or n using the x we choose. This happens only when $x^t = -1 \mod n$, since $x^t - 1 = 0 \mod n$ can not happen with r being already the smallest integer such that $x^r = 1 \mod n$. Fortunately it has been proved that the probability to meet such a bad x is at most $1/2^k$, where k is the number of distinct prime factors of n. Since k is at least 2, the probability is still large enough for us to find a good x, which has an even period r and $x^t \neq -1 \mod n$.

A quantum computer can find the period r because of the speedup afforded by quantum Fourier transform (QFT). Let us have two b-qubit registers. We select b large enough such that we can observe many periods. At the beginning, we set the two registers to state $|0\rangle$. Then we randomize the first register to a new state

$$|\psi_1\rangle = \frac{1}{\sqrt{S}} \sum_{k=0}^{S-1} |k\rangle|0\rangle, \tag{2.61}$$

where $S = 2^b$, the number of the total b-qubit states of the first register, with b large enough such that $2n^2 > S > n^2$.

We now design a certain series of pulses to compute $f(k) = x^k \mod n$, and

change the quantum state to

$$|\psi_2\rangle = \frac{1}{\sqrt{S}} \sum_{k=0}^{S-1} |k\rangle |f(k)\rangle. \tag{2.62}$$

Now, apply QFT [65] to the first register in (2.62), which is a unitary transform mapping every $|k\rangle$ to another state:

$$|k\rangle \rightarrow \frac{1}{\sqrt{S}} \sum_{k=0}^{S-1} e^{2\pi iuk/S} |u\rangle. \tag{2.63}$$

Then the quantum state of the system changes to

$$|\psi_3\rangle = \frac{1}{S} \sum_{u=0}^{S-1} |u\rangle \sum_{k=0}^{S-1} e^{2\pi iuk/S} |f(k)\rangle. \tag{2.64}$$

Assume that $f(k)$ has period r, and we write $k = d + jr$ such that $0 \leq d < r$, where d is the remainder of k after it is divided by r and j ranges from 0 to A, the largest integer such that $Ar < S$. This way, we can write $|\psi_3\rangle$ as

$$|\psi_3\rangle = \frac{1}{S} \sum_{u=0}^{S-1} |u\rangle \sum_{d=0}^{r-1} |f(d)\rangle e^{2\pi iud/S} \sum_{j=0}^{A} e^{2\pi iurj/S} I_{(d+rj<S)},$$

where $I_{(d+rj<S)} = 1$ when $d + rj < S$, and 0 otherwise. If $S = (A+1)r$, $I_{(d+rj<S)} = 1$ for every d and j. If $S \neq (A+1)r$, it is still reasonable to ignore the difference and let $I_{(d+rj<S)} = 1$ everywhere because we have chosen S large enough. In this case, we let

$$b_u = \frac{1}{S} \sum_{j=0}^{A} e^{2\pi iurj/S} = \frac{1}{S} \left(\frac{1 - e^{2\pi iur(A+1)/S}}{1 - e^{2\pi iur/S}} \right), \tag{2.65}$$

thus our quantum state is now

$$|\psi_3\rangle = \frac{1}{S} \sum_{u=0}^{S-1} \sum_{d=0}^{r-1} b_u e^{2\pi iud/S} |u\rangle |f(d)\rangle.$$

We can now measure the first register, and we want to find such a u, for which

there is an l satisfying

$$\left|\frac{u}{S}-\frac{l}{r}\right| \leq \frac{1}{2S}. \tag{2.66}$$

There are about r such u's, and it has been estimated that the probability to find such a u is at least 0.4 [65]. Because $\frac{1}{2S} < \frac{1}{2n^2}$, and we know that $r < n$, there is at most one fraction $\frac{k}{r}$ satisfying the condition and we can use *continued fraction expansions* to find the fraction. If k and r are coprime, we obtain r as the denominator of the fraction. If not, we only find a factor of r. If r is odd or $x^{r/2}$ does not give us a useful result, choose another x and try again. It may be necessary to try several (of the order $\mathcal{O}(\log\log n)$) times until r is successfully found, but the overall running time is still reasonable.

2. Circuit design for Shor's algorithm

Before we introduce the experiment by Vandersypen, et al. [5], we extend the above discussion a little further to the case when r divides S. Now S/r becomes an integer and (2.65) always holds so that S doesn't have to be very large. Moreover, (2.66) becomes an identity

$$u = \frac{l \cdot S}{r}, \tag{2.67}$$

i.e., r is the denominator of the fraction $\frac{u}{S}$ after cancelling the common factor between u and S if l and r are coprime. The integer 15 falls into this situation. The possible x can be 2, 4, 7, 8, 11, or 13. When we choose x to be 2, 7, 8 or 13, the period r is 4. In other cases, r is 2. The period r divides $S = 2^b$ in both cases. Only 2 qubits at most are required to compute one period of f. In the experiment, 3 qubits are used to obtain more periods.

Vandersypen et al. used liquid NMR to realize Shor's algorithm in factorizing 15. The sample in the experiment is a custom-synthesized material whose molecules have

five ^{19}F and two ^{13}C, so it has seven qubits ready for use. Those seven qubits are divided into two registers, 3 to store the number k (the first register, represented by $|k_2k_1k_0\rangle$) and 4 to store the modular exponentiation y (the second register, represented by $|y_3y_2y_1y_0\rangle$), see Fig. 2.10 and Fig. 2.11. The total Hamiltonian is

$$H = \sum_{i=1}^{7} \frac{1}{2}\omega_i \sigma_z^i + \sum_{i<j} 2\pi J_{ij} \sigma_z^i \sigma_z^j.$$

Each run of the experiment consists of 4 steps. In the first step, the sample is initialized to a certain pseudo-pure state; in the second step, a series of specially designed pulses are applied to realize the computation of modular exponentiation; in the third step, QFT is applied to the first register; finally, the period was obtained through the reading of the spectrum. The system begins from thermal equilibrium, where the density matrix is given by $\rho_0 = e^{-H/kT} \approx I - \frac{H}{kT}$. A suitable initial pseudo-pure state $|\psi_1\rangle = |0000001\rangle$ is obtained by the temporal averaging method.

Although it is difficult to design a general circuit for the modular exponentiation, it is easy to "hard-wire" for this special case in consideration. As the exponent k can be written as $k = k_0 + 2k_1 + 4k_2$, we can change the modular exponential x^k mod 15 into successive operations of modular multiplications by $x^{2^i k_i}$, with $i = 0, 1, 2$, applied to the second register y beginning from $y = 1$.

When $i = 0$, $y \cdot x = x = 1 + (x - 1)$, so the multiplication is actually a controlled-addition with $(x - 1)$ in case $k_0 = 1$. For $x = 7 = (0111)_2$, it is equal to flip the state of y_1 and y_2 ($y = (0001)_2$ before the multiplication). For $x = 11 = (1011)_2$, the same reasoning shows that we only have to flip the state of y_3 and y_1, depending on k_0. Gates A and B in Figs. 2.10 and 2.11 accomplish the modular multiplication x^{k_0}.

The situation is a little more complicated for $i = 1$. We only discuss the situation when $k_1 = 1$, since y will not change when $k_1 = 0$. Different strategies are needed for

$x = 7$ and $x = 11$. When $x = 11$, since $11^2 = 121 = 15 \times 8 + 1$, $y \times 11^2 = y \pmod{15}$. We need to do nothing and the same result holds for the third qubit k_2. When $x = 7$, we can design the circuit by first investigating the following identity

$$\begin{aligned} y \cdot 7^2 &= y \cdot 4 \mod 15 \\ &= (y_0 + 2y_1 + 4y_2 + 8y_3) \cdot 4 \mod 15 \\ &= (4y_0 + 8y_1 + 16y_2 + 32y_3) \mod 15 \\ &= (y_2 + 2y_3 + 4y_0 + 8y_1) \mod 15 \\ &= (y_2 \cdot 2^0 + y_3 \cdot 2^1 + y_0 \cdot 2^2 + y_1 \cdot 2^3) \mod 15. \end{aligned}$$

It shows that the modular multiplication can be achieved by exchanging the first qubit y_0 with the third qubit y_2, and the second qubit y_1 with the fourth qubit y_3. In Fig. 2.10, gates C, D, and E are used to accomplish the former, and gates F, G, and H the latter. Further simplification of the circuit can be made. Since the control bit y_3 is $|0\rangle$ before gate C, that gate can just be omitted. Gates H and E have no effect on the period; they can be omitted, too.

The circuit design for the quantum Fourier transform is just a standard design; see, e.g., [2, Fig. 5]. It has 3 Hadamard gates and 3 controlled-phase gates. Figs. 2.10 and 2.11 show the circuit designs for $y = 7$ and $y = 11$. Totally about 300 pulses are used in the experiment and it takes about 700ms to accomplish all steps in the case of $x = 7$.

3. Experimental result

Readout of the experiment needs a careful interpretation of the data. Because an NMR sample consists of many molecules, the readout is the average value of u from all molecules instead of the reading from a single molecule.

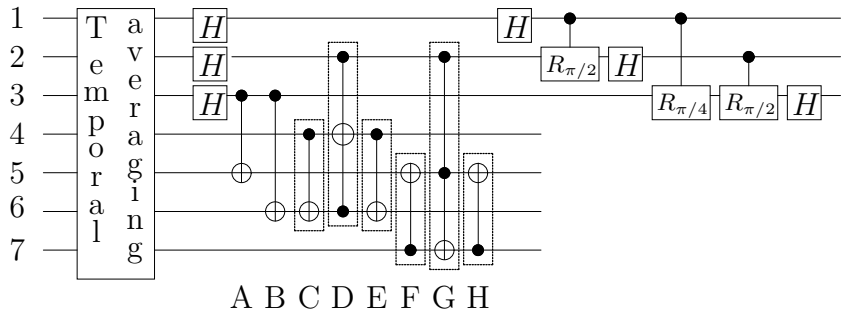


Fig. 2.10. The quantum circuit for the (hard) case for the realization of Shor's Algorithm ($x = 7$). From top to bottom, the qubits are k_2, k_1, k_0, y_3, y_2, y_1 and y_0, respectively, in sequential order.

Both qubits k_0 and k_1 are found to be in state $|0\rangle$ after the extraction of the spectra [5] for the easy case of $x = 11$, while qubit k_2 is in a equally mixture state of $|0\rangle$ and $|1\rangle$. Thus the possible u can be 0 and 4, i.e., 000 and 100 in binary form. From (2.67), r can be obtained as $r = 8/4 = 2$, and the greatest common divisors are computed as $\gcd(11^{2/2} + 1, 15) = 3$ and $\gcd(11^{2/2} - 1, 15) = 5$.

In the case of $x = 7$, the spectra in [5] indicate that only qubit k_0 is in state $|0\rangle$, and both qubits k_1 and k_2 are in equal mixture of states $|0\rangle$ and $|1\rangle$. Thus u is in a mixture of states $|0\rangle$, $|2\rangle$, $|4\rangle$ and $|6\rangle$. We can see that the period of u is 2, thus the period of the modular exponent r is $8/2 = 4$. The factors of 15 can finally be obtained as $\gcd(7^{4/2} - 1, 15) = 3$ and $\gcd(7^{4/2} + 1, 15) = 5$.

E. Quantum algorithm for lattice-gas systems

In the previous sections, we have explained how to construct a quantum computer using liquid NMR and illustrated a successful experiment. We have taken it for granted that the coherence can be maintained long enough and different qubits can be entangled even they are separated far apart in space. Unfortunately, these assump-

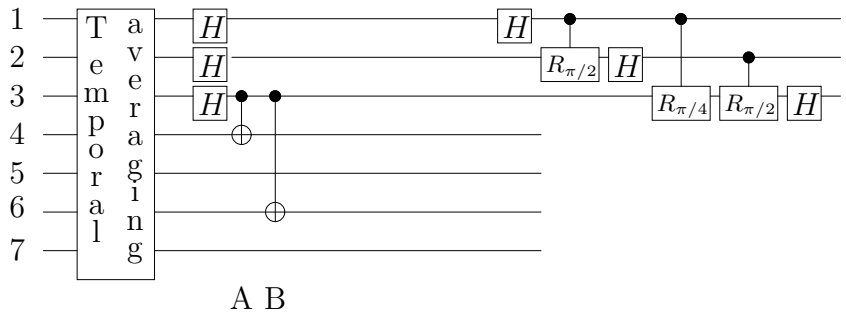


Fig. 2.11. The quantum circuit for the (easy) case for realization of Shor's Algorithm ($x = 11$). From top to bottom, the qubits are k_2, k_1, k_0, y_3, y_2, y_1 and y_0, respectively, in sequential order.

tions are not always practical and in fact they constitute great obstacles to overcome. The problem becomes more serious when more qubits are involved. *Type-II quantum computers* are proposed to alleviate this problem. A type-II quantum computer is composed of a network or array of small quantum computers interconnected by classical communication channels [66]. In stead of the global coherence and entanglement, only local coherence and entanglement within every small quantum computer, called a *node*, are required, and the difficulty faced by the centralized quantum computer is dramatically eased.

The wave function of the whole type-II quantum computer system is a tensor product over all the nodes:

$$|\psi(t)\rangle = |\psi(x_1, t)\rangle \otimes \quad \cdots \quad \otimes |\psi(x_N, t)\rangle, \tag{2.68}$$

where N is the number of nodes. The lattice gas algorithm (LGA) is specially suited for this structure. Every computation cycle can be broken up into three steps with

two intermediate states $|\psi'\rangle$ and $|\psi''\rangle$:

$$\begin{aligned} |\psi'\rangle &= \hat{C}|\psi(t)\rangle, \\ |\psi''\rangle &= \Gamma|\psi'\rangle, \\ |\psi(t+1)\rangle &= T|\psi''\rangle, \end{aligned} \tag{2.69}$$

where $\hat{C}$ is a unitary operator acting locally on every node, while Γ is a projection operator, such as a measurement, and T is the streaming operator which exchanges information among nodes. The type-II quantum computer takes advantage of parallelism in two ways: one classical, all the nodes work simultaneously; the other quantum, quantum entanglement is still kept inside every node. Because measurement is applied and the system is reset at the end of every computation cycle, the coherence only needs to be maintained for a short time.

1. Quantum algorithm for a lattice-gas model

Consider a one-dimensional diffusion equation without boundary condition

$$\frac{\partial \rho}{\partial t} = \frac{\partial^2 \rho}{\partial x^2}, \tag{2.70}$$

where ρ is the mass density or temperature function along the x-axis. Using the finite difference method, we can write a finite difference approximation to solve the above partial differential equation numerically:

$$\frac{\rho(x, t+\tau) - \rho(x,t)}{\tau} = \frac{\rho(x+l,t) - 2\rho(x,t) + \rho(x-l,t)}{l^2}, \tag{2.71}$$

where τ is the time step size and l is the space step size. From physics, we know that the above equation may be studied by a lattice gas algorithm. Without loss of generality, we assume that τ and l are normalized so that the difference equation can

be written as

$$\rho(x_i, k+1) - \rho(x_i, k) = \frac{1}{2}(\rho(x_{i+1}, k) - 2\rho(x_i, k) + \rho(x_{i-1}, k)).$$

Points x_i are evenly distributed along the x-axis, also called *nodes*. To study the above equation, two functions, $f_1(x_i, k)$ and $f_2(x_i, k)$, called *channels*, are defined for each node x_i at time k. The set of values of these two functions are called the *state* of node x_i. Any physical observable, such as the density function $\rho(x_i, k)$, is a function of the state at the node. The evolution of the lattice, or the state of all nodes, consists of two operations: *collision* and *propagation*. A collision is a local operator only defined by the state of the node itself. The propagation operator transfers information from one node to another and the state at one node changes according to the state of other nodes. This is completed by defining a velocity vector for every channel which gives the information flow a direction. In our special example here, information in the two channels flows in opposite directions. After propagation, one channel gets its new value from its left neighbor, while the other from its right neighbor. This LGA is completed with a Type II quantum computer by J. Yepez of the Air Force Research Laboratory and M.A. Pravia, et al. of the Department of Nuclear Engineering at MIT [67, 68, 69, 66]. The actual result is not as good as desired, but improvement is still possible.

To store a floating point number, a classical computer uses a register with 32 or 64 bits, depending on the machine. Quantum computers presently have difficulty to do it the same way as classical computers because there is not yet the technology for 32 or 64 qubits. In this quantum lattice-gas algorithm, a two qubit system is proposed for every node. The two qubits are represented by $|q_1(x_i, k)\rangle$ and $|q_2(x_i, k)\rangle$,

respectively, and

$$|q_1(x_i,k)\rangle = \sqrt{f_1(x_i,k)}|0\rangle + \sqrt{1-f_1(x_i,k)}|1\rangle,$$

$$|q_2(x_i,k)\rangle = \sqrt{f_2(x_i,k)}|0\rangle + \sqrt{1-f_2(x_i,k)}|1\rangle.$$

The state of the whole system $|\psi(x_i,k)\rangle$ at node x_i and time k is a tensor product:

$$\begin{aligned}|\psi(x_i,k)\rangle &= |q_1(x_i,k)\rangle|q_2(x_i,k)\rangle \\ &= \sqrt{f_1(x_i,k)f_2(x_i,k)}|00\rangle + \sqrt{(1-f_1(x_i,k))f_2(x_i,k)}|10\rangle \\ &+\sqrt{f_1(x_i,k)(1-f_2(x_i,k))}|01\rangle \\ &+\sqrt{(1-f_1(x_i,k))(1-f_2(x_i,k))}|11\rangle.\end{aligned}$$

Quantities $f_1(x_i,k)$ and $f_2(x_i,k)$ are the probabilities of occurrence of the state $|0\rangle$ for qubit 1 and 2, respectively, corresponding to the two channels, and $1-f_{1,2}(x_i,k)$ are the occurrence probabilities of the state $|1\rangle$. Since the states are normalized, $0 \le f_{1,2}(x_i,k) \le 1$, and we let $\rho(x_i,k) = f_1(x_i,k) + f_2(x_i,k)$. It is noted that our Type-II quantum computer assigns ρ a continuous value (a function of the occurrence probabilities) rather than a discrete value as a digital computer does. An array of two qubit systems are used in the computation, corresponding to a series of nodes.

The quantum LGA here has three steps in every cycle that complete a step of the finite difference algorithm computation: *collision*, *measurement*, and *re-initialization*. The last two composed are equal to one propagation operation in a normal lattice gas algorithm. Because the propagation needs information exchange among different nodes, measurement and classical communication are needed to accomplish one operation. We map the quantum state to a vector in $\mathbf{C}^4$ as that given in (2.34).

In the collision step, a unitary operator is applied simultaneously to all nodes:

$$|\overline{\psi}(x_i,k)\rangle = U|\psi(x_i,k)\rangle,$$

where

$$U = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \frac{1}{2} - \frac{i}{2} & \frac{1}{2} + \frac{i}{2} & 0 \\ 0 & \frac{1}{2} + \frac{i}{2} & \frac{1}{2} - \frac{i}{2} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \tag{2.72}$$

The new occurrence probabilities of the sate $|0\rangle$ of the two qubits after the operation, $\overline{f}_1$ and $\overline{f}_2$, can be computed using

$$\begin{aligned} \overline{f}_1 &= \langle \overline{\psi} | n_1 | \overline{\psi} \rangle, \qquad n_1 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \\ \overline{f}_2 &= \langle \overline{\psi} | n_2 | \overline{\psi} \rangle, \qquad n_2 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \end{aligned} \tag{2.73}$$

leading to

$$\begin{aligned} \overline{f}_1(x_i, k) &= \frac{1}{2}(f_1(x_i, k) + f_2(x_i, k)), \\ \overline{f}_2(x_i, k) &= \frac{1}{2}(f_1(x_i, k) + f_2(x_i, k)). \end{aligned}$$

The collision operator is actually doing a job of averaging. The state after the collision is also called the local equilibrium.

In the second step, a measurement is applied at every node and $\overline{f}_{1,2}(x_i, k)$ of all the nodes are retrieved for future use.

In the third step, using information from the measurement from the previous

step, the state of all the nodes are re-initialized to a separable state

$$\begin{aligned} |q_1(x_i, k+1)\rangle &= \sqrt{\overline{f}_1(x_{i+1}, k)}|0\rangle + \sqrt{1-\overline{f}_1(x_{i+1}, k)}|1\rangle, \\ |q_2(x_i, k+1)\rangle &= \sqrt{\overline{f}_2(x_{i-1}, k)}|0\rangle + \sqrt{1-\overline{f}_2(x_{i-1}, k)}|1\rangle. \end{aligned} \tag{2.74}$$

It can be seen that the second and third steps have accomplished the propagation operation. At node x_i, the new state of channel one is acquired from the same channel of its right neighbor node x_{i+1}, and channel two acquires its state from its left neighbor. It is complicated here only because the communication between two quantum systems is difficult.

To see how this LGA works, let us begin from a local equilibrium state, $f_1(x_i, k) = f_2(x_i, k) = \rho(x_i, k)/2$, where the states come off a collision operation (step 2). We list the $f_{1,2}$ around position x_i before the third step in two rows

$$\begin{array}{lccccccc} f_1: & \cdots & \dfrac{\rho(x_{i-2},k)}{2} & \dfrac{\rho(x_{i-1},k)}{2} & \dfrac{\rho(x_i,k)}{2} & \dfrac{\rho(x_{i+1},k)}{2} & \cdots \\ f_2: & \cdots & \dfrac{\rho(x_{i-2},k)}{2} & \dfrac{\rho(x_{i-1},k)}{2} & \dfrac{\rho(x_i,k)}{2} & \dfrac{\rho(x_{i+1},k)}{2} & \cdots \end{array}$$

and after the third step

$$\begin{array}{lccccccc} f_1: & \cdots & \dfrac{\rho(x_{i-1},k)}{2} & \dfrac{\rho(x_i,k)}{2} & \dfrac{\rho(x_{i+1},k)}{2} & \dfrac{\rho(x_{i+2},k)}{2} & \cdots \\ f_2: & \cdots & \dfrac{\rho(x_{i-3},k)}{2} & \dfrac{\rho(x_{i-2},k)}{2} & \dfrac{\rho(x_{i-1},k)}{2} & \dfrac{\rho(x_i,k)}{2} & \cdots \end{array}$$

We can see that the row of f_1 (channel one) moves left and the row of f_2 (channel two) moves right. According to our definition, the new value of ρ is the sum of $f_1(x_i, k+1)$ and $f_2(x_i, k+1)$, i.e., $\rho(x_i, k+1) = \dfrac{1}{2}(\rho(x_{i+1}, k) + \rho(x_{i-1}, k))$. It is easy to check that

$$\rho(x_i, k+1) - \rho(x_i, k) = \frac{1}{2}(\rho(x_{i+1}, k) - 2\rho(x_i, k) + \rho(x_{i-1}, k)),$$

as desired.

Applications of the Type-II quantum computer with quantum LGA also have been reported in the simulation of the time-dependent evolution of a many-body

quantum mechanical system [70], solution of a one-dimensional magnetohydrodynamic turbulence [71], representation of solitons [72] and other equations.

2. Physical realization and result

The experiment in Subsection 1 uses a two-qubit molecule, chloroform, whose structure is shown in Fig. 2.7. The hydrogen and carbon nuclei serve as qubit 1 and 2, respectively.

The actual results obtained from the experiment are compared with simulation results [68]. After 12 steps, the error becomes very large. Imperfection in the decoupling sequences is blamed and it is believed that the problem can be mitigated when the technology is improved in the future. Extreme requirement of high accuracy in the control pulse and readout is a disadvantage of this Type-II quantum computer, because it uses a continuous representation (the probability of occurrence) instead of a discrete one. Thus, it is more vulnerable to the inaccuracy in the NMR operation. Small errors in every step accumulate and finally become intolerable. Repeated measurement and re-initialization ease the requirement for coherence time, but place a high requirement on the fidelity at the same time.

F. Conclusion

In this article, we present the basic technology used to construct a quantum computer with liquid state NMR. The successful experiments for many algorithms have shown that liquid state NMR is capable of simulating a quantum computer and forms a test bed for quantum algorithms. It is so far the only technology available to realize a 7 qubit algorithm in laboratory. One reason for its success is the robustness of the spin system which only interacts with the external magnetic field, and it is possible

to maintain the coherence for a long time (from seconds to hours). Besides, over the 60 years history of NMR spectroscopy, analytic tools have been developed for the purpose of chemical and medical applications, and exact description and dedicated coherence control of the dynamics of the quantum spin system is now available to achieve high accuracy in the pulse design and application. In fact, the experimental techniques established in NMR, especially the coherence control technology, can be easily transferred to other quantum systems if they have a similar Hamiltonian, and the research in NMR is therefor helpful for the development of other more complicated and powerful quantum computers.

Liquid NMR has played a pioneering role in the quantum computer technology development. But its lack of scalability has constitute a severe obstacle to its future applicability. However, new technology of *solid state NMR* have the potential to overcome liquid NMR's difficulties. For solid state NMR, under low temperature, the relaxation times of spins are typically very long, and the coupling between qubits is strong so that the control can be fast and easy. The small ratio of the gate time and the decoherence time makes more gates available, and more complicated algorithm can be tested. The nuclei can be cooled down easily and the spin system is highly polarized. The signal is much stronger so that fewer nuclei are needed. Even without the help of gradient field and the silicon technology, as we have mentioned, a quantum computer with 30 to 40 qubits is envisioned with designed molecules similar to that of the liquid state NMR computer except that the ensemble is in a solid crystal state. This is already a quantum system that reaches the limit a classical computer can simulate. Although it is still not scalable and not a standard quantum computer, these small and medium scale quantum computers will help in the building of a scalable and working quantum computer.

CHAPTER III

SUPERCONDUCTING QUANTUM COMPUTING DEVICES

Even though quantum effects are mostly observed in microscopic scales, they also manifest macroscopically. A particular case of such is *superconductivity*. Superconducting devices composed of Josephson junctions (JJ), Cooper-pair boxes and SQUID (superconducting quantum interference devices) have been developed since the 1980s as magnetometers, gradiometers, gyroscopes, sensors, transistors, voltmeters, etc., to perform measurements on small magnetic fields, and to demonstrate the quantum effects of tunneling, resonance and coherence [73, 74, 75, 76, 77, 78]. Many industrial and medical applications have also resulted: maglev trains, superconducting power generator, cables and transformers, MRI and NMR for medical scans,to mention a few. With the advances in solid-state lithography and thin-film technology, superconducting devices have the great advantage of being easily scalable and engineering-designable. For a bulk superconductor, if its size is reduced smaller and smaller, then the quasi-continuous electron conduction band therein turns into discrete energy levels. In principle, such energy levels can be used to constitute a qubit. The first demonstration of quantum-coherent oscillations of a Josephson "charge qubit" in a superposition of eigenstates was made by Nakamura et al. [79] in 1999. Ever since, theoretically and experimentally there has been steady progress. New proposals for qubits based on *charges, flux, phase* and *charge-flux* have been made, with observations of microwave-induced Rabi oscillations of two-level populations in those qubit systems [80, 81, 82, 83].

In this chapter, we will first introduce superconductivity in Section A, the Josephson junction in Section B, and the elementary superconducting circuits in Section C. Superconducting quantum circuits and gates are studied in Sections D and E, and

conclude with measurements in Section F.

A. Superconductivity

Superconductivity was discovered in 1911 by the Dutch physicist Heike Kamerlingh Onnes (1853–1926), who dedicated his career to the exploration of extremely cold refrigeration. In 1908, he successfully liquefied helium by cooling it to -452° F (4 K). In 1911, he began the investigation of the electrical properties of metals in extremely cold temperatures, using liquid helium. He noticed that for solid mercury at cryogenic temperature of 4.2 K, its electric resistivity abruptly disappeared (as if there were a jump discontinuity). This is the discovery of superconductivity, and Onnes was awarded the Nobel Prize of Physics in 1913.

The theory of superconductivity was further advanced in 1957 by three American physicists (then at the University of Illinois), J. Bardeen, L. Cooper, and J. Schrieffer, called the *BCS Theory* [84]. The BCS theory explains superconductivity at temperatures close to absolute zero. Cooper theorized that atomic lattice vibrations were directly responsible for unifying and moderating the entire current. Such vibrations force the electrons to pair up into partners that enable them to pass all of the obstacles which cause resistance in the conductor. These partners of electrons are known as *Cooper pairs*. This electron coupling is viewed as an exchange of *phonons*, with phonons being the quanta of lattice vibration energy. The electron Cooper pairs are coupled over a range of hundreds of nanometers, three orders of magnitude larger than the lattice spacing. The effective net attraction between the normally repulsive electrons produces a pair binding energy on the order of milli-electron volts, enough to keep them paired at extremely low temperatures. Experimental corroboration of an interaction with the lattice was provided by the *isotope effect* on the superconducting

transition temperature T_c. More on Cooper pairs in the next section.

Interested readers may find more information in superconductivity textbooks [85, 86, 87, 88], for example.

B. More on Cooper pairs and Josephson junctions

In the preceding section, we briefly introduced Cooper pairs. For electrons in a metal at low temperature, despite the fact that the electrons Coulomb force repel each other, the lattice of positive ions in the metal can have phonon vibration energy that mediates the coupling or pairing of eletrons to overcome the repelling force. It works as follows [89]: When one of the electrons that make up a Cooper pair and passes close to an ion in the crystal lattice, the attraction between the negative electron and the positive ion cause a vibration (i.e., phonon) to pass from ion to ion until the other electron of the pair absorbs the vibration. The net effect is that the electron has emitted a phonon and the other electron has absorbed the phonon. It is this exchange that keeps the Cooper pairs together. It is important to understand, however, that the pairs are constantly breaking and re-forming. Because electrons are indistinguishable particles, it is easier to think of them as permanently paired. The composite entity, the Cooper pair, thus behaves as a single particle. These coupled electrons can take the character of a *boson* with charge twice that of an electron and zero spin. The first excited state of Cooper pairs has a minimum energy of 2 Δ, where Δ is what we had referred to earlier as the *superconducting gap*. See also Δ in (3.1) and (3.2) below. Cooper pairs carry the current in a superconductor.

Now, consider two superconductors with currents. If they are kept apart and totally isolated from each other, then the phases of their wavefunctions will be independent. Bring them close together but separate by a thin non-conducting oxide

barrier of tens of angströms thickness. Then Cooper pairs begin to tunnel stronger across the barrier as the separation decreases. This current is called the *Josephson current*. The "sandwich-like" arrangement is called the *Josephson junction*, see Fig. 3.1. Both were named after the British physicist B.D. Josephson (Nobel laureate in physics 1973).

The basic equations governing the dynamics of the Josephson tunneling are

$$V(t) = \frac{\hbar}{2e}\frac{\partial\phi(t)}{\partial t}, \qquad I(t) = I_c \sin(\phi(t)), \tag{3.1}$$

where $V(t)$ and $I(t)$ are, respectively, the voltage and current across the JJ, $\phi(t)$ is the phase difference of the superconductors across the JJ, and I_c, a constant, is the *critical current*. In the microscopic theory of superconductivity [88], it is known that

$$I_c = \frac{\pi\Delta}{2eR_N}\tanh\frac{\Delta}{2T}, \tag{3.2}$$

where Δ is the superconducting order parameter *energy gap*, T is the temperature, and R_N is a constant.

A superconducting quantum interference device (SQUID) consists of two superconductors separated by thin insulating layers of JJ. SQUID are usually made of either a lead alloy (with 10% gold or indium) and/or niobium, often consisting of the tunnel barrier sandwiched between a base electrode of niobium and the top electrode of lead alloy.

There are two types of SQUID:

(1) dc-SQUID: It was invented by R. Jaklevic, J. Lambe, A. Silver, and J. Mercereau of Ford Research Labs in 1964. It consists of two JJ placed in parallel such that electrons tunneling through the junctions manifest quantum interference, depending upon the strength of the magnetic field within a loop.

(2) rf-(or ac-) SQUID: It was invented by J. E. Zimmerman and A. Silver at Ford in 1965. It is made up of one Josephson junction, which is mounted on a superconducting ring. An oscillating current is applied to an external circuit, whose voltage changes as an effect of the interaction between it and the ring. The magnetic flux can then be measured.

DC-SQUID are more difficult and expensive to fabricate, but they are much more sensitive. A SQUID can detect a change of energy as much as 100 billion times weaker than the electromagnetic energy that moves a compass needle. We will study dc and rf SQUID in more technical detail in the following sections.

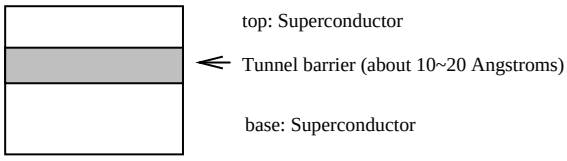


Fig. 3.1. Schematic of a simple Josephson junction. It has a "sandwich" structure. The base is an electrode made of a very thin niobium layer, formed by deposition. The midlayer, the tunnel barrier, is oxidized onto the niobium surface. The top layer, also an electrode, made of lead alloy (with about 10% gold or indium) is then deposited on top of the other two.

C. Superconducting circuits: Classical

There are about a half dozen major proposals for superconducting qubits. We will introduce some of them in this section. First, *classical* superconducting circuits characterized by their *Lagrangians* will be presented. Then we advance to their quantum versions through the *canonical quantization* procedure when only a few electrons are present on such circuits. Our discussions mainly follow the tutorial paper by Wendin and Shumeiko [90].

In superconducting quantum computing applications, four basic types of circuits with JJ are commonly used as *building blocks*:

(1) single current-biased JJ;

(2) single Cooper-pair box (SCB);

(3) rf-SQUID;

(4) dc-SQUID.

We address each of them separately in the following subsection.

1. Current-biased JJ

This is the simplest superconducting circuit, consisting of a tunnel Josephson junction with superconducting electrodes connected to a current source. A schematic is given in Fig. 3.2.

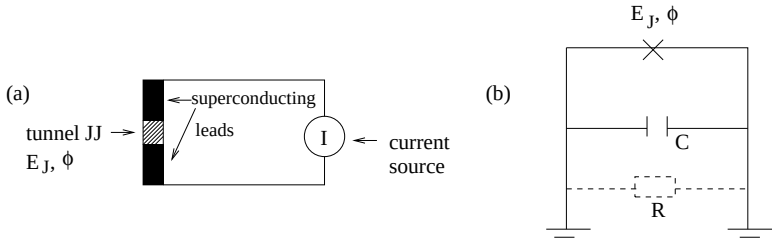


Fig. 3.2. (a) A current biased Josephson junction; (b) An equivalent lumped circuit, where × signifies the barrier of the JJ. (Adapted from [90]).

Let $\phi(t)$ be the phase difference between the wavefunctions in the two superconductors across the junction. Let $V(t)$ denote the voltage difference across the junction. Then by the first equation in (3.1),

$$\phi(t) = \frac{2e}{\hbar} \int_{t_0}^{t} V(\tau) d\tau + \phi_0. \tag{3.3}$$

(The superconducting phase $\phi(t)$ is also related to magnetic flux $\Phi(t)$ as

$$\phi(t) = \frac{2e}{\hbar}\Phi(t) = 2\pi\frac{\Phi(t)}{\Phi_0}, \tag{3.4}$$

where $\Phi_0 = h/(2e)$ is the magnetic flux quantum.) As noted in the second equation of (3.1) in Section B, the JJ current is proportional to the sine of $\phi(t)$ across the insulator:

$$I_J = I_c \sin\phi, \quad I_c \equiv \text{the critical Josephson current, (cf. (3.1))}. \tag{3.5}$$

Differentiating (3.3), we have

$$\dot{\phi}(t) = \frac{2e}{\hbar}V(t). \tag{3.6}$$

Refer to Fig. 3.2 (b). The current-voltage relations for the junction capacitance C and resistance R are given by the standard formulas

$$I_C = C\frac{dV}{dt}, \quad I_R = V/R. \tag{3.7}$$

From (3.6) and (3.7), by the Kirchhoff law of the circuit (see Fig. 3.2 (b)), we now have

$$\frac{\hbar}{2e}C\ddot{\phi} + \frac{\hbar}{2eR}\dot{\phi} + I_c \sin\phi = I_e, \tag{3.8}$$

where I_e is the *bias current*. Eq. (3.8) takes the form of a *damped forced pendulum*.

The damping term $\frac{\hbar}{2eR}\dot{\phi}$ in (3.8) determines the lifetime of the (future superconducting quantum circuit) qubit. Thus, the dissipation must be extremely small. Ideally, we assume that it is zero. So we consider an undamped Eq. (3.8):

$$\frac{\hbar}{2e}C\ddot{\phi} + I_c \sin\phi = I_e. \tag{3.9}$$

Remark C.1. It is necessary to emphasize that dropping the damping term $\hbar\dot{\phi}/(2eR)$ in (3.8) constitutes a reasonable approximation only under the following conditions

of superconductivity:

(i) low temperature, i.e., T is small;

(ii) $|\dot{\phi}|$ is very small;

(iii) $T, \hbar\omega \ll \Delta$, where Δ is the energy gap in (3.2). □

For the undamped Eq. (3.9), Lagrangian and Hamiltonian variational forms can now be obtained by kinetic and potential energies:

$$\text{kinetic energy } K = K(\dot{\phi}) = \left(\frac{\hbar}{2e}\right)^2 \frac{C}{2}\dot{\phi}^2, \tag{3.10}$$

$$\text{potential energy } U = U(\phi) = \frac{\hbar}{2e}\int [I_c \sin\phi - I_e] d\phi$$
$$= \frac{\hbar}{2e} I_c (1 - \cos\phi) - \frac{\hbar}{2e} I_e \phi, \tag{3.11}$$

where the kinetic energy is proportional to the electrostatic energy of the junction capacitor (corresponding to the first term in (3.9)), while the potential energy consists of the energy of the Josephson current and the magnetic energy of the bias current (corresponding to the last two terms in (3.9)).

For future quantum superconducting circuit applications, we introduce several useful constants. The first is the *charging energy* of the junction capacitor charged with a single Cooper pair (of electrons)

$$E_C \equiv \frac{(2e)^2}{2C}. \tag{3.12}$$

The second,

$$E_J \equiv \frac{\hbar}{2e} I_c \tag{3.13}$$

is called the *Josephson energy*. The third constant,

$$\omega_J \equiv \sqrt{\frac{2eI_c}{\hbar C}}, \tag{3.14}$$

is called the *plasma frequency* of the JJ. This is the frequency of the small-amplitude oscillation of the unforced pendulum (i.e., Eq. (3.9) with $I_e = 0$). With (3.12) and (3.13), we can write (3.10) and (3.11) as

$$K = \frac{\hbar^2 \dot{\phi}^2}{4E_C}, \quad U = E_J(1 - \cos\phi) - \frac{\hbar}{2e} I_e \phi.$$

Thus, we obtain the Lagrangian

$$L(\phi, \dot{\phi}) = K - U = \frac{\hbar^2 \dot{\phi}^2}{4E_C} - E_J(1 - \cos\phi) + \frac{\hbar}{2e} I_e \phi,$$

whose Lagrangian variational equation

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{\phi}} - \frac{\partial L}{\partial \phi} = 0$$

is exactly (3.9).

The Hamiltonian H is related to the Lagrangian L through

$$H(p, \phi) = p\dot{\phi} - L, \quad \text{where} \quad p = \frac{\partial L}{\partial \dot{\phi}} = \frac{\hbar^2}{2E_C}\dot{\phi}, \tag{3.15}$$

with p being the canonical momentum operator conjugate to ϕ. Then

$$H(p, \phi) = \frac{E_C}{\hbar^2} p^2 + E_J(1 - \cos\phi) - \frac{\hbar}{2e} I_e \phi, \tag{3.16}$$

and the Hamiltonian equations of motion

$$\dot{\phi} = \frac{\partial H}{\partial p}, \quad \dot{p} = -\frac{\partial H}{\partial \phi} \tag{3.17}$$

are again equivalent to (3.9).

2. Single Cooper-pair box (SCB)

An SCB is driven by an applied voltage Vg through capacitance Cg to induce an offset charge. The circuit consists of a small superconducting "island" connected via

a Josephson tunnel junction to a large superconducting reservoir. See a schematic in Fig. 3.3.

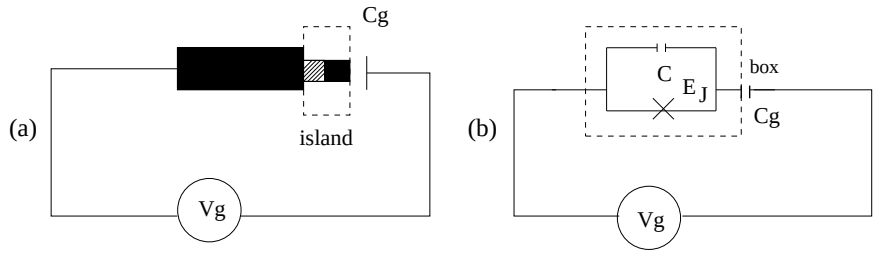


Fig. 3.3. (a) A single Cooper-pair box. (b) An equivalent lumped circuit, where × signifies the barrier of JJ. (Adapted from [90] and [94].)

The electrostatic energy of the SCB is the sum

$$K = \frac{CV^2}{2} + \frac{Cg(Vg - V)^2}{2},$$

which, after using (3.6) and completing the square, gives

$$K = \frac{(C + Cg)}{2}\left(\frac{\hbar}{2e}\dot{\phi} - \frac{Cg}{C + Cg}Vg\right)^2 + \frac{1}{2}\left(Cg - \frac{Cg^2}{C + Cg}\right)Vg^2. \tag{3.18}$$

Dropping the (last) constant term in (3.18) and denoting $C_\Sigma \equiv C + Cg$, we have

$$K = K(\dot{\phi}) = \frac{C_\Sigma}{2}\left(\frac{\hbar}{2e}\dot{\phi} - \frac{Cg}{C_\Sigma}Vg\right)^2.$$

The potential energy U from (3.11) (by dropping the bias current I_e as it is no longer present) is

$$U = U(\phi) = E_J(1 - \cos\phi).$$

Therefore, we obtain the Lagrangian

$$L(\phi, \dot{\phi}) = \frac{C_\Sigma}{2}\left(\frac{\hbar}{2e}\dot{\phi} - \frac{Cg}{C_\Sigma}Vg\right)^2 - E_J(1 - \cos\phi). \tag{3.19}$$

The Hamiltonian, according to (3.15), is

$$H(\phi, p) = \frac{1}{2C_\Sigma}\left(\frac{2e}{\hbar}\right)^2 p^2 + E_J(1 - \cos\phi). \tag{3.20}$$

3. rf- or ac-SQUID

The rf-SQUID, also called an ac-SQUID or a magnetic-flux box, is depicted in Fig. 3.4. It is the magnetic analogue of the (electrostatic) SCB discussed in Subsection 2. It consists of a tunnel JJ inserted in a superconducting loop.

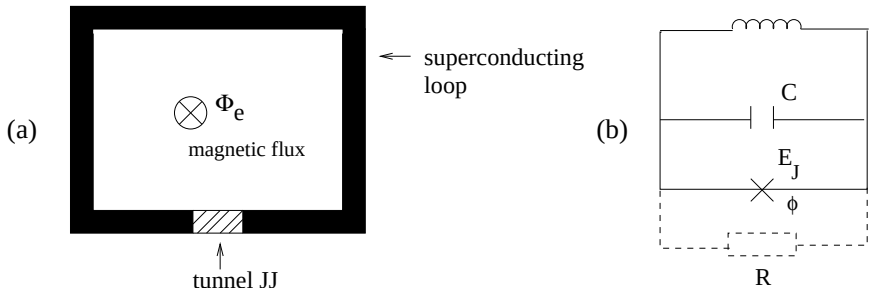


Fig. 3.4. (a) An rf-SQUID. (b) An equivalent lumped circuit. (Adapted from [90, Fig. 8].)

Let I_L denote the current associated with the inductance L of the superconducting leads. Then by (3.4),

$$I_L = \frac{\hbar}{2eL}(\phi - \phi_e), \quad \phi_e = \frac{2e}{\hbar}\Phi_e,$$

where Φ_e is the external magnetic flux piercing the rf-SQUID loop. Using the same arguments as in Subsection 1, by the Kirchhoff circuit law we arrive at

$$\frac{\hbar}{2e}C\ddot{\phi} + \frac{\hbar}{2eR}\dot{\phi} + I_c \sin\phi + \frac{\hbar}{2eL}(\phi - \phi_e) = 0, \tag{3.21}$$

where in (3.8) the bias current I_e is replaced by $-I_L$.

If the damping is very small, then the term containing $\dot{\phi}$ can again be dropped

and the Lagrangian of the rf-SQUID is

$$L(\phi,\dot{\phi}) = \frac{\hbar^2\dot{\phi}^2}{4E_C} - E_J(1-\cos\phi) - E_L\frac{(\phi-\phi_e)^2}{2}, \quad \left(E_L \equiv \frac{\hbar^2}{(2e)^2 L}\right). \tag{3.22}$$

The Hamiltonian is then obtained as

$$H(\phi,p) = \frac{E_C}{\hbar^2}p^2 + E_J(1-\cos\phi) + E_L\frac{(\phi-\phi_e)^2}{2}. \tag{3.23}$$

4. dc-SQUID

A dc-SQUID consists of two JJ in parallel coupling to a current source. It has some similarity to the current-biased single junction (Fig. 3.2), except that there is an additional magnetic flux piercing the SQUID loop, which serves as a control on the effective Josephson energy of the double JJ. See Fig. 3.5 for a schematic of a dc-SQUID.

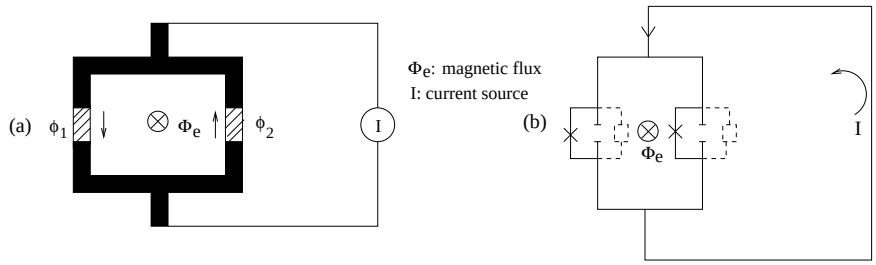


Fig. 3.5. (a) Schematic of a dc-SQUID. (b) An equivalent (nominal) lumped circuit.

Let ϕ_1 and ϕ_2 be superconducting phase differences across the JJ 1 and 2, respectively. Assume that the inductance of the SQUID loop is small so that the magnetic energy of the circulating currents can be neglected. Then the total voltage drop over the two JJ is zero:

$$V_1 + V_2 = 0.$$

From (3.6),

$$\dot{\phi}_1 + \dot{\phi}_2 = 0,$$

and, thus $\phi_1 + \phi_2$ is a constant, and

$$\phi_1 + \phi_2 = \phi_e, \tag{3.24}$$

where ϕ_e is the biasing superconducting phase related to the biasing magnetic flux. Define

$$\phi_\pm = \frac{\phi_1 \pm \phi_2}{2}.$$

Then

$$\phi_+ = \frac{\phi_1 + \phi_2}{2} = \frac{1}{2}\phi_e, \quad \phi_- = \frac{\phi_1 - \phi_2}{2},$$

which leads to

$$\phi_1 = \phi_- + \frac{\phi_e}{2}, \quad \phi_2 = \frac{\phi_e}{2} - \phi_-. \tag{3.25}$$

For the symmetric case, the two JJ have the same E_J, C and I_c. Assume that there is no dissipation, thus we can neglect the $\dot{\phi}$ term. The equation (3.24) can be rewritten using ϕ_+ as

$$2\phi_+ - \phi_e = 0.$$

The Kirchhoff circuit law requires

$$\frac{\hbar}{2e}C\ddot{\phi}_1 + I_c \sin\phi_1 - \frac{\hbar}{2e}C\ddot{\phi}_2 - I_c \sin\phi_2 - I_e = 0,$$

or

$$\frac{\hbar}{e}C\ddot{\phi}_- + 2I_c \cos\phi_+ \sin\phi_- - I_e = 0,$$

by using trigonometric identities. Thus the dynamic equation for the system of ϕ_+

and ϕ_- can be obtained as

$$\begin{cases} \frac{\hbar}{e}C\ddot{\phi}_- + 2I_c\cos\phi_+\sin\phi_- - I_e = 0 \\ \frac{\hbar}{eL}(2\phi_+ - \phi_e) = 0. \end{cases} \tag{3.26}$$

The system has in fact only one degree of freedom since $2\phi_+ = \phi_e$. By substituting ϕ_+ by $\phi_e/2$ and comparing (3.26) with

$$\frac{d}{dt}\frac{\partial L}{\partial\dot{\phi}_-} - \frac{\partial L}{\partial\phi_-} = 0, \tag{3.27}$$

we can obtain the Lagrangian of the dc-SQUID as

$$L = \left(\frac{\hbar}{2e}\right)^2 C\dot{\phi}_-^2 + \frac{\hbar}{2e}2I_c\cos\phi_+\cos\phi_- + \frac{\hbar}{2e}I_e\phi_-,$$

Its Hamiltonian, in turn, is

$$H = \left(\frac{\hbar}{2e}\right)^2 C\dot{\phi}_-^2 - \frac{\hbar}{2e}2I_c\cos\phi_+\cos\phi_- - \frac{\hbar}{2e}I_e\phi_-.$$

The kinetic energy of the dc-SQUID can be obtained from the Lagrangian as

$$K(\phi_-) = \left(\frac{\hbar}{2e}\right)^2 2C\frac{\dot{\phi}_-^2}{2}. \tag{3.28}$$

It has a simple interpretation as the charging energy of the two junction capacitances (see Fig. 3.5 (b)) by looking at identity:

$$\begin{aligned} 2C\frac{\hbar}{2e}\dot{\phi}_- &= C\frac{\hbar}{2e}(\dot{\phi}_1 - \dot{\phi}_2) \\ &= C(V_1 - V_2) \\ &= q. \end{aligned}$$

By setting $E_C \equiv \dfrac{(2e)^2}{2\cdot 2C}$ and define E_J and E_L as before, we can rewrite the Hamiltonian as

$$H = \frac{\hbar^2}{4E_C}\dot{\phi}_-^2 - 2E_J\cos\frac{\phi_e}{2}\cos\phi_- - \frac{\hbar}{2e}I_e\phi_-,$$

and in terms of $p = \dfrac{\hbar^2}{2E_C}\dot{\phi}_-$,

$$H = \frac{E_C}{\hbar^2}p^2 - 2E_J \cos\frac{\phi_e}{2}\cos\phi_- - \frac{\hbar}{2e}I_e\phi_-. \tag{3.29}$$

D. Superconducting circuits: quantum

We know that the quantization of the electromagnetic field gives a simple harmonic oscillator. A classical superconducting circuit may be viewed as an antenna. It can thus radiate electromagnetic waves. From this analogue, we see that superconducting circuits can be quantized as well when the JJ becomes microscopically small, and the continuous electric current becomes discretely charged.

We now formalize the above argument by following the standard approach of *canonical quantization*. From the classical Lagrangian L, and then $p = \partial L/\partial\dot{\phi}$ we have the Hamiltonian H just as in (3.15). Now consider the simplest case of a single junction (Subsection 1, in particular Fig. 3.2). From (3.15),

$$p = \frac{\partial L}{\partial\dot{\phi}} = \frac{\hbar^2}{2E_C}\dot{\phi}, \quad \text{(see (3.12) for } E_C\text{)}. \tag{3.30}$$

From the first equation in (3.1),

$$V = \frac{1}{2e}\frac{h}{2\pi}\dot{\phi} = \frac{\hbar}{2e}\dot{\phi}. \tag{3.31}$$

Thus,

$$\begin{aligned} p &= \frac{\hbar^2}{2E_C}\dot{\phi} = \left(\frac{\hbar}{2e}\right)^2 C\dot{\phi} \\ &= \left(\frac{\hbar}{2e}\right)^2 C\left(\frac{2e}{\hbar}\right)V = \frac{\hbar}{2e}CV \\ &= \frac{\hbar}{2e}q \qquad (q = CV \text{ on the junction capacitor}) \\ &= \hbar\frac{q}{2e} = \hbar n, \end{aligned} \tag{3.32}$$

where $q/(2e)$ is n, the number of Cooper pairs. Therefore, the momentum p has a simple interpretation that it is proportional to the number of Cooper pairs n on the junction capacitor. Substituting (3.32) into (3.16), we obtain the (quantum) Hamiltonian for the current-biased JJ:

$$H = E_C n^2 - E_J \cos\phi - \frac{\hbar}{2e} I_e \phi, \tag{3.33}$$

where the constant E_J in (3.16) has been dropped.

For the SCB, from (3.19) we have the conjugated momentum

$$p = \frac{\partial L}{\partial \dot{\phi}} = \frac{\hbar C_\Sigma}{2e}\left(\frac{\hbar}{2e}\dot{\phi} - \frac{C_g}{C_\Sigma} V g\right), \tag{3.34}$$

and by using (3.32) and (3.34) in (3.20), we have

$$H = \overline{E}_c (n - n_g)^2 - E_J \cos\phi, \tag{3.35}$$

where

$$\overline{E}_c \equiv (2e)^2/(2C_\Sigma), \quad n_g = C_g V_g/(2e), \tag{3.36}$$

and n_g is the number of Cooper pairs on the gate capacitor. This n_g is tunable through different designs of Cg and Vg.

For the dc-SQUID, according to the derivations of (3.29), we obtain the Hamiltonian

$$H = E_C n_-^2 - 2E_J \cos\frac{\phi_e}{2}\cos\phi_- - \frac{\hbar}{2e} I_e \phi_-, \tag{3.37}$$

where $E_C = \dfrac{(2e)^2}{4C}$ and $n_- = 2C\dfrac{\hbar}{(2e)^2}\dot{\phi}_-$.

In quantization, the classical momentum p in (3.30) becomes the *differential operator*

$$\hat{p} = -i\hbar\frac{\partial}{\partial\phi}, \tag{3.38}$$

where using ϕ we mean ϕ_- for the dc-SQUID. From (3.32), we thus also have the

operator of the pair number

$$\hat{n} = -i\frac{\partial}{\partial\phi}, \tag{3.39}$$

and the commutator relation

$$[\phi, \hat{n}] = i. \tag{3.40}$$

The time evolution of the wave function $\psi = \psi(\phi, t)$ satisfies the Schrödinger equation

$$i\hbar\frac{\partial}{\partial t}\psi(\phi, t) = \widehat{H}\psi(\phi, t) = H\left(\phi, \frac{\hbar}{i}\frac{\partial}{\partial\phi}\right)\psi(\phi, t), \tag{3.41}$$

where $H = H(\phi, p) = H(\phi, \hbar n)$ is the Hamiltonian derived in (3.33) through (3.37).

E. Quantum gates

We begin the discussion by using CPB as a major reference model for this section. Recall from (3.36), that the Hamiltonian for a CPB is given by

$$H = E_C(\hat{n} - n_g)^2 - E_J\cos\phi. \tag{3.42}$$

Here we assume that

$$E_C \gg E_J. \tag{3.43}$$

The pair-number operator $\hat{n}$ is defined by

$$\hat{n}|n\rangle = n|n\rangle, \qquad n = \text{an integer}, \tag{3.44}$$

where $|n\rangle$ is called the number state. From (3.39), we see that the wave function $\psi = \psi(\phi)$ of $|n\rangle$ satisfies the differential equation

$$-i\frac{\partial}{\partial\phi}\psi = n\psi. \tag{3.45}$$

To allow only integer n in (3.45) for consideration in solving ψ, a periodic constraint must be imposed:

$$\psi(\phi + 2\pi) = \psi(\phi). \tag{3.46}$$

(Without such a constraint, the number of electrons on the island may be odd, or n could be a real value number. But here the electrode is miniaturized small enough that such cases would not happen as only a finite number of Cooper pairs can exist on the island.) Therefore, from (3.45) and (3.46), we obtain

$$\psi(\phi) = \frac{1}{\sqrt{2\pi}} e^{in\phi}, \quad \text{for} \quad n = 0, \pm 1, \pm 2, \cdots, \tag{3.47}$$

where $1/\sqrt{2\pi}$ is the normalization factor with respect to the $L^2(0, 2\pi)$-norm. From (3.42), we see that for the lowest energy eigenstate $|0\rangle$ and $|1\rangle$ of $\hat{n}$, when (3.43) holds, the states $|0\rangle$ and $|1\rangle$ are nearly degenerate when $n_g = 0.5$:

$$\begin{aligned} H|0\rangle &= [E_C(0 - 0.5)^2 - E_J \cos\phi]|0\rangle \approx \frac{1}{4} E_C |0\rangle, \\ H|1\rangle &= [E_C(1 - 0.5)^2 - E_J \cos\phi]|1\rangle \approx \frac{1}{4} E_C |1\rangle. \end{aligned} \tag{3.48}$$

This is a favorable situation. (Normally, if two states $|0\rangle$ and $|1\rangle$ differ much in energy levels, then even though they discriminate better, the higher lying state $|1\rangle$ is *less* stable, and the system tends to decohere and lie more often in $|0\rangle$ than in $|1\rangle$, an unbalanced situation in quantum computing to be avoided.)

Similarly, if $n_g = n+1/2$, then the two states $|n\rangle$ and $|n+1\rangle$ are nearly degenerate for any integer n. For simplicity, let us just consider $n_g \approx 0.5$.

Theorem E.1. *Assume that (3.43) holds, and that $n_g \approx 0.5$. Let*

$$V = span\{|0\rangle, |1\rangle\}. \tag{3.49}$$

Then the projection of the Hamiltonian H in (3.42) with respect to the ordered basis

in (3.49) satisfies

$$P_H = \begin{bmatrix} E_C[\frac{1}{4} + (n_g - 0.5)] & -\frac{1}{2}E_J \\ -\frac{1}{2}E_J & E_C[\frac{1}{4} - (n_g - 0.5)] \end{bmatrix} + \mathcal{O}(|n_g - 0.5|^2). \tag{3.50}$$

Proof. The projection matrix P_H of H on V is easily evaluated as

$$P_H = \begin{bmatrix} a_0 & b \\ c & a_1 \end{bmatrix}, \tag{3.51}$$

where

$$a_j = \langle j|H|j\rangle \qquad \text{for } j = 0, 1, \tag{3.52}$$

and

$$b = \langle 0|H|1\rangle, c = \langle 1|H|0\rangle. \tag{3.53}$$

Using (3.47) for $|0\rangle$ and $|1\rangle$, we compute, e.g.,

$$\begin{aligned} a_1 &= \langle 1|H|1\rangle \\ &= \int_0^{2\pi} \left(\frac{1}{\sqrt{2\pi}} e^{-i\phi}\right) \left(E_C(-i\frac{\partial}{\partial\phi} - n_g)^2 - E_J\cos\phi\right) \left(\frac{1}{\sqrt{2\pi}} e^{i\phi}\right) d\phi \\ &= \frac{1}{2\pi}\int_0^{2\pi} \{E_C(1 - n_g)^2 - E_J\cos\phi\} d\phi \\ &= \frac{E_C}{2\pi} \cdot 2\pi \left[(1 - 0.5) + (0.5 - n_g)\right]^2 \\ &= E_C\left[0.5^2 + 2(0.5)(0.5 - n_g) + (0.5 - n_g)^2\right] \\ &= E_C\left[\frac{1}{4} - (n_g - 0.5)\right] + \mathcal{O}(|n_g - 0.5|^2). \end{aligned} \tag{3.54}$$

Similarly, the entries a_0, b and c can be computed. We obtain (3.50). □

As

$$P_H = \frac{1}{4}E_C \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} E_C(n_g - 0.5) & -\frac{1}{2}E_J \\ -\frac{1}{2}E_J & -E_C(n_g - 0.5) \end{bmatrix} + \mathcal{O}(|n_g - 0.5|^2), \tag{3.55}$$

we can just use the effective Hamiltonian

$$\bar{P}_H = \begin{bmatrix} E_C(n_g - 0.5) & -\frac{1}{2}E_J \\ -\frac{1}{2}E_J & -E_C(n_g - 0.5) \end{bmatrix} = E_C(n_g - 0.5)\sigma_z - \frac{1}{2}E_J\sigma_x, \tag{3.56}$$

as an approximate Hamiltonian in the subsequent discussion. The state $|0\rangle$ and $|1\rangle$ constitute a charge-qubit system. In addition, a probe gate may be coupled to the box through a junction to perform measurement, shown in Fig. 3.6.

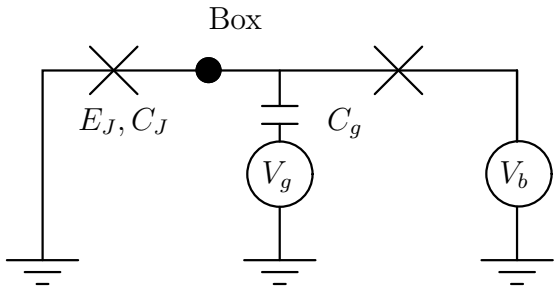


Fig. 3.6. Schematic of a charge qubit constructed with a Cooper pair box. The box is denoted by a black dot and the two Josephson junctions are denoted by two crosses. The pulse gate voltage V_g can change the offset charge of the junction. The other junction is connected to a voltage V_b used for measurement, and the gate is called the probe gate.

1. One qubit operation: charge-qubit

There are various methods to manipulate the information encoded in the CPB system, and the essence is to know how to control the time-varying Hamiltonian. In the constrained linear subspace V (see (3.49)) spanned by number states $|0\rangle$ and $|1\rangle$, the system Hamiltonian has been obtained in Eq. (3.56). We assume that n_g is nearly equal to 0.5 and E_C is far less than the superconducting gap Δ. The evolution matrix of this system in time duration τ can be easily computed using results from NMR:

$$e^{-i\bar{P}_H\tau/\hbar} = e^{-i(E_C(n_g-0.5)\sigma_z - \frac{1}{2}E_J\sigma_x)\tau/\hbar}, \tag{3.57}$$

which is a rotation around the following axis:

$$\frac{1}{\sqrt{E_J^2/4 + E_C^2(n_g-0.5)^2}}\left(-\frac{1}{2}E_J\boldsymbol{e}_x + E_C(n_g-0.5)\boldsymbol{e}_z\right)$$

with angle $\tau\sqrt{E_J^2/4 + E_C^2(n_g-0.5)^2}/\hbar$.

In this section, our main objective is to show that we can derive the Rabi (1-qubit) rotation gate

$$U_{\theta,\alpha} = \begin{bmatrix} \cos(\theta) & -i\sin(\theta)e^{-i\alpha} \\ -i\sin(\theta)e^{i\alpha} & \cos(\theta) \end{bmatrix}, \tag{3.58}$$

by using the evolution matrix (3.57) with different choices of the parameter n_g and time duration τ. Note that the only tunable parameter is n_g. So we signify the dependence of $\bar{P}_H$ on n_g from (3.56) as

$$\bar{P}_H = \bar{P}_H(n_g). \tag{3.59}$$

Lemma E.2. *We have the x-rotation matrix*

$$R_{x,\psi} = e^{-i\bar{P}_H(\bar{n}_g)\tau/\hbar} = \begin{bmatrix} \cos(\psi/2) & -i\sin(\psi/2) \\ -i\sin(\psi/2) & \cos(\psi/2) \end{bmatrix} \tag{3.60}$$

where $\bar{n}_g = 0.5$ *and* $\psi = -E_J\tau/\hbar$. *In particular,*

$$R_{x,\pi} = \begin{bmatrix} 0 & -i \\ -i & 0 \end{bmatrix} = -i\sigma_x. \tag{3.61}$$

Proof. When $\bar{n}_g = 0.5$, we have from (3.56)

$$\bar{P}_H(\bar{n}_g) = \begin{bmatrix} 0 & -\frac{1}{2}E_J \\ -\frac{1}{2}E_J & 0 \end{bmatrix}. \tag{3.62}$$

The rest follow immediately from taking the exponential matrix $e^{-i\bar{P}_H(\bar{n}_g)\tau/\hbar}$. □

Remark E.1. The operation in Lemma E.2 is achieved through several steps. First, the offset charge $n_g = C_gV_g/(2e)$ as controlled by V_g is abruptly switched to the degeneration point $n_g = 0.5$, kept for duration τ, and then abruptly switched back. Time duration τ is in the order of $10^{-10}s$, and the switching must be fast enough to avoid any adiabatic transition. □

Lemma E.3. *Define*

$$R_{+,\theta} = e^{-i\bar{P}_H(\bar{n}_g^1)\tau/\hbar}, \quad R_{-,\phi} = e^{-i\bar{P}_H(\bar{n}_g^2)\tau/\hbar}, \tag{3.63}$$

where $\bar{n}_g^1$ *and* $\bar{n}_g^2$ *satisfy, respectively,*

$$E_C(\bar{n}_g^1 - 0.5) = -\frac{1}{2}E_J \equiv -\delta, \quad E_C(\bar{n}_g^2 - 0.5) = \frac{1}{2}E_J = \delta, \tag{3.64}$$

and

$$\theta = \phi = 2\sqrt{2}\tau\delta/\hbar. \tag{3.65}$$

Then we obtain the y-rotation and z-rotation matrices as

$$R_{y,\theta} = \begin{bmatrix} \cos\frac{\theta}{2} & -\sin\frac{\theta}{2} \\ \sin\frac{\theta}{2} & \cos\frac{\theta}{2} \end{bmatrix} = -R_{-,3\pi/2}R_{+,\theta}R_{-,\pi/2}, \tag{3.66}$$

$$R_{z,\phi} = \begin{bmatrix} e^{-i\phi/2} & 0 \\ 0 & e^{i\phi/2} \end{bmatrix} = -R_{x,\pi/2}R_{y,\phi}R_{x,3\pi/2}. \tag{3.67}$$

Proof. With the choice of $\bar{n}_g^1$, $\bar{n}_g^2$, δ and τ given in (3.64) and (3.65), we have

$$R_{+,\theta} = e^{i\frac{\theta}{2\sqrt{2}}(\sigma_x+\sigma_z)}, \quad R_{-,\phi} = e^{i\frac{\phi}{2\sqrt{2}}(\sigma_x-\sigma_z)}. \tag{3.68}$$

Note that $R_{+,\theta}$ and $R_{-,\theta}$ are rotations with respect to axes $-\frac{1}{\sqrt{2}}(\boldsymbol{e}_x+\boldsymbol{e}_z)$, $-\frac{1}{\sqrt{2}}(-\boldsymbol{e}_x+\boldsymbol{e}_z)$, respectively. According to the properties of matrices in $\mathbf{SU}(2)$, we have

$$\begin{aligned} R_{-,\pi/2}R_{+,-\theta}R_{-,3\pi/2} &= -R_{y,\theta} \\ &= -\begin{bmatrix} \cos\frac{\theta}{2} & -\sin\frac{\theta}{2} \\ \sin\frac{\theta}{2} & \cos\frac{\theta}{2} \end{bmatrix}, \\ R_{x,\pi/2}R_{y,\phi}R_{x,3\pi/2} &= -R_{z,\phi} \\ &= -\begin{bmatrix} e^{-i\phi/2} & 0 \\ 0 & e^{i\phi/2} \end{bmatrix}. \end{aligned} \tag{3.69}$$

The negative sign comes from the fact that $R_{\boldsymbol{n},2\pi} = -I_2$ for any unit vector $\boldsymbol{n}$. □

Corollary E.4. *We have the Rabi rotation gate*

$$\begin{aligned} U_{\theta/2,\alpha} &= e^{-i\frac{\theta}{2}(\cos\alpha\sigma_x+\sin\alpha\sigma_y)} \\ &= -R_{x,\pi/2}R_{y,-\alpha}R_{x,\theta}R_{y,\alpha}R_{x,3\pi/2}, \end{aligned} \tag{3.70}$$

through the cascading of quantum operations $e^{-i\bar{P}_H(n_g)\tau/\hbar}$ by tuning the parameter n_g and time duration τ.

Next, we construct the Rabi rotation gate $U_{\theta,\phi}$ in an alternative approach which

is perhaps easier to implement. From (3.56), if we let the voltage V_g be oscillating (called a phase gate [79]) such that

$$E_C(n_g - 0.5) = \epsilon \cos(\omega t + \alpha), \tag{3.71}$$

where ϵ is the amplitude, then (3.56) gives (an approximate Hamiltonian)

$$H = \epsilon \cos(\omega t + \alpha)\sigma_z - \frac{1}{2}E_J\sigma_x. \tag{3.72}$$

The above Hamiltonian is with reference to the ordered basis $\{|0\rangle, |1\rangle\}$. Now define a new basis

$$|\uparrow\rangle \equiv \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle), \quad |\downarrow\rangle \equiv \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle). \tag{3.73}$$

Then, with respect to the above ordered basis, the Hamiltonian (1) becomes

$$\tilde{H} = \epsilon \cos(\omega t + \alpha)\sigma_x + \frac{1}{2}E_J\sigma_z, \tag{3.74}$$

where we rename E_J to $-E_J$ just for simplicity.

We now utilize a standard procedure in NMR by transforming the system into a rotating frame, namely, for the original wave function $|\chi(t)\rangle$ with Hamiltonian (3.74), let

$$|\psi(t)\rangle = e^{i\omega t\sigma_z/2}|\chi(t)\rangle, \tag{3.75}$$

leading to a revolution matrix of $|\psi(t)\rangle$ as

$$\begin{aligned} U_{\theta/2,\alpha} &= e^{-i\frac{\epsilon t}{2\hbar}(\cos\alpha\sigma_x + \sin\alpha\sigma_y)} \\ &= \begin{bmatrix} \cos(\frac{\theta}{2}) & -i\sin(\frac{\theta}{2})e^{-i\alpha} \\ -i\sin(\frac{\theta}{2})e^{i\alpha} & \cos(\frac{\theta}{2}) \end{bmatrix}, \end{aligned} \tag{3.76}$$

where $\theta = \epsilon t/\hbar$. This is a Rabi rotation with respect to the ordered basis $\{|\uparrow\rangle, |\downarrow\rangle\}$. A Rabi rotation with respect to the ordered basis $\{|0\rangle, |1\rangle\}$ can be obtained by using

a similarity transformation using the Walsh-Hadamard gate.

The density matrix of the system, according to the Boltzmann distribution, is given by

$$e^{\frac{-H}{k_B T}},$$

where $k_B = 1.381 \times 10^{-23}$J/K is the Boltzman constant and T is the absolute temperature. When $k_B T \ll E_C$, and $n_g \neq 0.5$, the Coulomb energy dominates the Hamiltonian and the system is initialized to its ground state, and this initializes the system.

2. Flux-qubit, charge-flux qubit and phase qubit

In this subsection, we briefly describe three other ways of setting up qubits in a superconducting circuit.

In an rf-SQUID, the magnetic flux Φ through the loop is quantized and must satisfy

$$(\Phi_0/2\pi)\phi + \Phi_{ext} + \Phi_{ind} = m\Phi_0, \tag{3.77}$$

where $\Phi_0 = 2.07 \times 10^{-15} Wb$; as before, m is an integer, Φ_{ext} is the external magnetic field and Φ_{ind} is induced by a current through the loop as in Fig. 3.7. That surface current through the loop is induced to compensate Φ_{ext} and its direction can be either clockwise or counterclockwise. If we denote the two surface current states as $|\uparrow\rangle$ and $|\downarrow\rangle$, they form a basis and the qubit is called a *flux qubit*. The main references for this qubit setup are [92, 93, 91]. When Φ_{ext} is near one half of Φ_0, the current can be either clockwise or counterclockwise and the system behaves like a Cooper pair box when n_g is near 0.5. Recall the Hamiltonian of an rf-SQUID in (3.23). When the self-inductance L is large enough such that $\beta_0 = E_J 4\pi^2 L/\Phi_0^2 = E_J/E_L > 1$ and Φ_{ext} is near $\Phi_0/2$ (this means $\phi_{ext} = 2\pi\Phi_{ext}/\Phi_0$ is near π), the Hamiltonian has a

shape of a double-well near $\Phi = \Phi_0/2$ ($\phi = 2\pi\Phi/\Phi_0 = \pi$), see Fig. 3.8. The two lowest states at the bottom of each well are well separated from other excited levels in low temperature and suitable for quantum computation. When $\Phi_{ext} = \Phi_0/2$, the two states are degenerate and they are maximum superpositions of $|\uparrow\rangle$ and $|\downarrow\rangle$ [91]. When Φ_{ext} is away from $\Phi_0/2$, they approach $|\uparrow\rangle$ and $|\downarrow\rangle$. The Hamiltonian of this two level system has a simplified form as

$$H = -\frac{1}{2}B_z\sigma_z - \frac{1}{2}B_x\sigma_x,$$

where B_z can be tuned by Φ_{ext} and B_x is a function of E_J which is also tunable if the junction is replaced by a dc-SQUID. Thus, any 1-qubit operation can be realized through combinations of different choices of Φ_{ext} and E_J. When $\Phi_{ext} = \Phi_0/2$, $B_z = 0$.

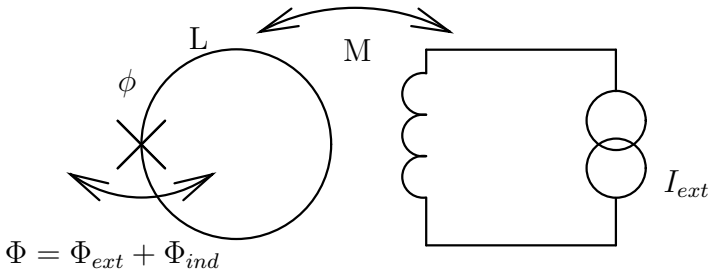


Fig. 3.7. Schematic of a flux qubit constructed with a Josephson junction in a loop.

A shortcoming of the simple rf-SQUID design is that its size is large in order to obtain high self-inductance and that makes it very susceptible to external noise. A better design uses more junctions in the loop and makes the size smaller. A three junction flux-qubit is shown in Fig. 3.9. Two of the junctions are designed to be the same while the third is different. The quantum constraint (3.77) applies and $\phi_1 + \phi_2 + \phi_3 = \phi_{ext} = 2\pi\Phi_{ext}/\Phi_0$. By neglecting the magnetic energy and the

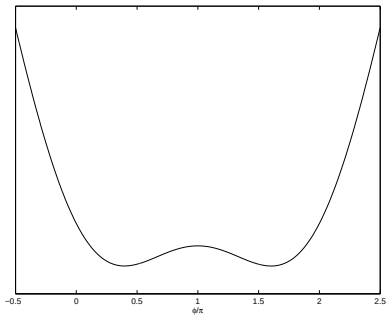


Fig. 3.8. The double-well shape potential of a flux qubit with Hamiltonian (3.23). We take $\Phi_{ext} = \Phi_0/2$ and plot the potential curve near $\phi = \pi$.

Coulomb term, we obtain the potential of the Hamiltonian which is dominated by the Josephson terms:

$$U(\phi_1, \phi_2) = -E_J \cos\phi_1 - E_J \cos\phi_2 - \alpha E_J \cos(\phi_{ext} - \phi_1 - \phi_2).$$

A similar two-well potential appears in the 2-D plane of ϕ_1 and ϕ_2 when $\alpha > 0.5$. The setup has been used to observe the transitions between the two states when irradiated by an rf-photon field, and demonstrate superpositions of $|\uparrow\rangle$ and $|\downarrow\rangle$ in spectroscopic experiments.

For a flux qubit, E_J is much larger than E_C. When E_J is almost equal to E_C, both the Coulomb and JJ terms are important, and the qubit is called the *charge-flux qubit* [95, 96]. Neither ϕ nor n is a good quantum number and the lowest energy states are superpositions of several charge states. A typical design is shown in Fig. 3.10, which is developed from that of a Cooper pair box with a dc-SQUID. A larger junction is inserted in the loop for measurement, which is shunted by capacitors to reduce phase fluctuations. An external flux Φ_{ext} is also imposed as in the dc-SQUID case. Normally, the qubit works near $n_g = 1/2$, and the two lowest eigenstates are superpositions of number states $|0\rangle$ and $|1\rangle$. Denoted by $|+\rangle$ and $|-\rangle$, the two

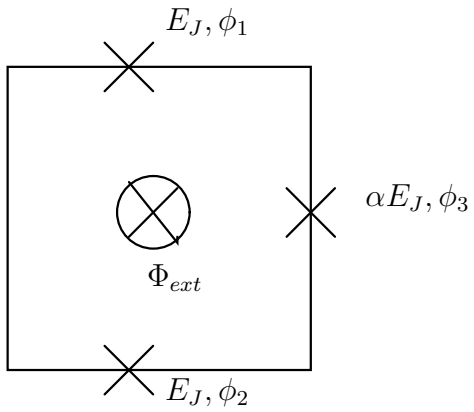


Fig. 3.9. A three junction superconducting loop serving as a flux qubit. Compared with the simple design of rf-SQUID in Fig. 3.7, it has a smaller size and better coherence performance.

states have an energy difference E_J and the system Hamiltonian can be written as $H = \frac{1}{2}E_J\sigma_z$ when $n_g = 0.5$ exactly. Control signal with resonant frequency can be applied on the gate to manipulate the system. After putting the system in a "rotating frame" as before, the system Hamiltonian changes to

$$H = h\nu(\sigma_x \cos\alpha + \sigma_y \sin\alpha),$$

when the control signal is $\Delta n_g \cos(\omega t + \alpha)$, while $\nu = 2E_C \Delta n_g \langle + | \hat{n} | - \rangle / h$. The system behaves like an NMR spin and all technologies, such as composite pulses can be used to increase the accuracy and robustness of the operation [95]. Charge-flux qubit shows better decoherence than charge- or flux- qubit in experiments.

Readout of the charge-flux qubit is realized through the current in the loop instead of the charge on the island. When a biased current I_b slightly below the critical current I_c of the large junction is applied, the large junction is switched into a finite voltage state depending on the qubit state. In theory, the measurement efficiency $p_+ - p_- = 0.95$ holds, where p_i is the probability to obtain a voltage in the

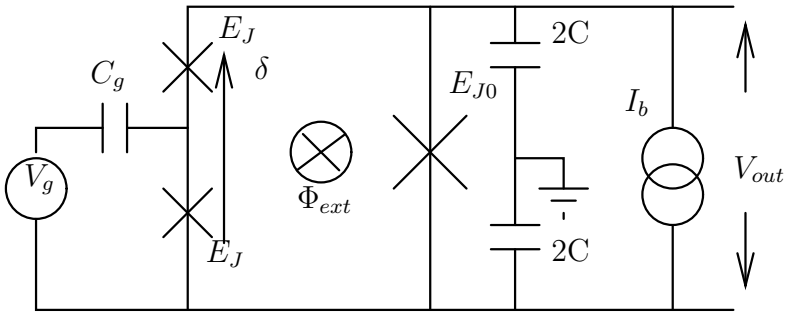


Fig. 3.10. Circuit of a charge-flux qubit, with two small junctions and one large junction in a loop. An external flux Φ_{ext} penetrates the loop and a voltage V_g is applied through capacitor C_g to control the bias charge n_g. A bias current I_b is used for measurement.

read out when the qubit is in state $|i\rangle$.

Lastly, we address the *phase-qubit* setup, which is a current-biased Josephson junction. Its special feature is that the junction energy E_J is much larger than the Coulomb energy E_C. See Fig. 3.2. Here, our references are [97, 98, 99]. For such, the Coulomb term is neglected, so its Hamiltonian can be obtained from (3.33) as

$$H = -E_J \cos\phi - \frac{\hbar}{2e} I_e \phi,$$

and the potential is a periodic function of ϕ offset by $I_e\phi$, with shape appearing like a "washboard", see Fig. 3.11. Normally, the JJ is undamped and we choose I_e not too large so that there are a series wells on the potential curve. In every well formed by $\cos\phi$, it is well-known that the energy is quantized and has different levels. Besides the lowest two states serving as qubit states $|0\rangle$ and $|1\rangle$, sometimes there are one or more other states in the well. The extra level or levels may be used for measurement. Transitions between $|0\rangle$ and $|1\rangle$, in the form of Rabi rotation, is realized by applying a resonant electromagnetic field with $\omega = E_{10}/\hbar$, where E_{10} is

the energy difference between $|0\rangle$ and $|1\rangle$. Measurement is accomplished by inspecting the tunneling probability of states through the well. There are two methods. One is to use a microwave field resonant with E_{21} (i.e. the energy difference between $|1\rangle$ and $|2\rangle$) to pump $|1\rangle$ to the second excited state $|2\rangle$, which has a higher tunneling probability. The other is to tilt the washboard by increasing I_e so that $|1\rangle$ can tunnel through the barrier with high probability.

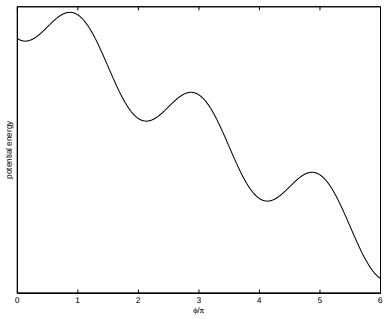


Fig. 3.11. A "washboard" shape potential energy curve of a phase qubit. It is obtained by tilting the cosine function of ϕ by $-\frac{\hbar}{2e}I_e\phi$. When $I_e > \frac{2e}{\hbar}E_J$, there will be no well on the curve.

3. Two qubit operations

Various proposals have been suggested to couple two qubits for different kinds of superconducting qubits. Capacitors, for example, can be used to couple two charge qubits. Experiments have shown two-qubit oscillations using this scheme [100], and a conditional gate operation has also been demonstrated using the same device [52]. One disadvantage of the capacitor coupling is that it is not switchable, which makes the pulse design inflexible. It is also difficult to couple two qubits far away from each other because only the neighboring qubit coupling is convenient.

Inductance, instead, seems more promising. The simplest design is to construct

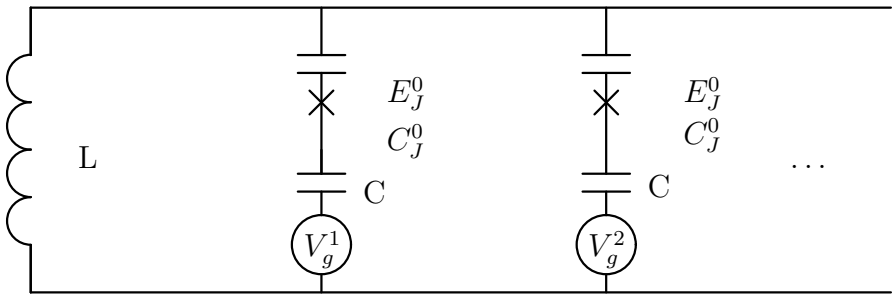


Fig. 3.12. A simple design to couple charge qubits with inductance. The inductance and the effective capacitance of the charge qubits configured in parallel form a weak coupling among the qubits.

a weak coupling between the qubits through the CL (capacitance-inductance) oscillation, see Fig 3.12 [101]. But the coupling is still not switchable and thus lacks engineering flexibility. An improved design embeds a dc-SQUID into the qubit circuit with the advantage that the Josephson energy can be controlled [102]. The junction in Fig. 3.12 is replaced by a dc-SQUID. See Fig. 3.13. An external magnetic field Φ_e^i penetrates the SQUID and changes the term of the Josephson Hamiltonian to $-2E_J^0 \cos(\pi\Phi_e/\Phi_0)\cos\phi$, where the effective phase difference ϕ equals half of the difference of the two phase drops at the two junctions and $\dfrac{2\pi\Phi_e}{\Phi_0} = \phi_e$, cf. Section 4. This means that we replace E_J in equation (3.42) by a tunable $E_J(\Phi_e)$:

$$E_J(\Phi_e) = 2E_J^0 \cos(\pi\Phi_e/\Phi_0).$$

In this configuration, the additional effective interaction Hamiltonian induced by the oscillation in the LC-circuit can be given in the form of Pauli-matrices as

$$H_{int} = -\sum_{i<j} \frac{E_J(\Phi_e^i)E_J(\Phi_e^j)}{E_L}\sigma_y^i\sigma_y^j,$$

where $E_L = [\Phi_0^2/(\pi^2 L)](C_J/C_{qb})^2$, while C_{qb} is the capacitor of the qubit defined by

$C_{qb}^{-1} = C_J^{-1} + C^{-1}$.

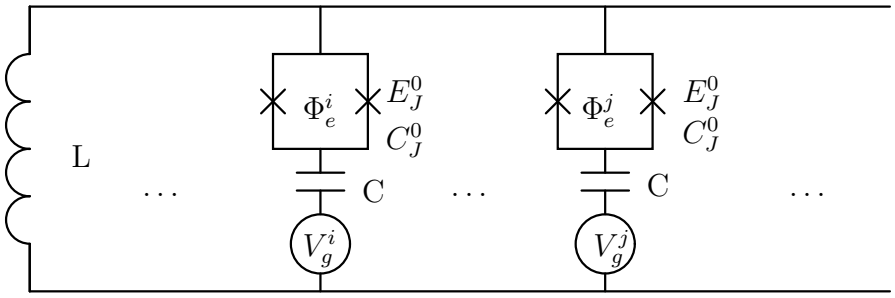


Fig. 3.13. A design for coupling charge qubits with inductance where the junctions in the charge qubits are replaced by a dc-SQUID. All qubits are coupled through an inductor L, and an external field Φ_e^i penetrates every dc-SQUID. This changes the effective Josephson term in the Hamiltonian to $-2E_J^0 \cos(\pi\Phi_e/\Phi_0)\cos\phi$ and makes E_J tunable by Φ_e^i.

Assume that we can still constrain every qubit in the projected subspace spanned by $|0\rangle$ and $|1\rangle$, see V in (3.49), and note that the whole Hamiltonian of the n-qubit system can be written as

$$H = \sum_{i=1}^{n} (\epsilon(V_g^i)\sigma_z^i - \frac{1}{2}E_J(\Phi_e^i)\sigma_x^i) - \sum_{i<j} \frac{E_J(\Phi_e^i)E_J(\Phi_e^j)}{E_L}\sigma_y^i\sigma_y^j, \tag{3.78}$$

where we collect all parameters before σ_z in $\epsilon(V_g^i)$ for simplicity. If we let all $\Phi_e^j = \Phi_0/2$ and $n_g^j = 0.5$ when $j \neq i$, the whole system Hamiltonian changes to

$$H = \epsilon(V_g^i)\sigma_z^i - \frac{1}{2}E_J(\Phi_e^i)\sigma_x^i,$$

and all the other terms are turned off. We can perform any single qubit operation through the approximation offered by (3.56).

Similarly, a two qubit operation between qubits i and j can be performed by turning off all $E_J^k(\Phi_e^k)$ and n_g^k except qubits i and j. By doing this, now the Hamiltonian

becomes

$$H = \epsilon(V_g^i)\sigma_z^i + \epsilon(V_g^j)\sigma_z^j - \frac{1}{2}E_J(\Phi_e^i)\sigma_x^i - \frac{1}{2}E_J(\Phi_e^j)\sigma_x^j + \Pi_{ij}\sigma_y^i\sigma_y^j.$$

If we also move the two qubits to their degenerate state, i.e., $n_g^i = n_g^j = 0.5$, the Hamiltonian is simplified to

$$H = -\frac{1}{2}E_J(\Phi_e^i)\sigma_x^i - \frac{1}{2}E_J(\Phi_e^j)\sigma_x^j + \Pi_{ij}\sigma_y^i\sigma_y^j.$$

Because σ_x does not commute with σ_y, the computation of the evolution matrix is tedious and the design of the CNOT gate and conditional phase change gate is complicated. Although it provides a mechanism to realize any qubit gates in combination with one qubit gates, more simplification is helpful.

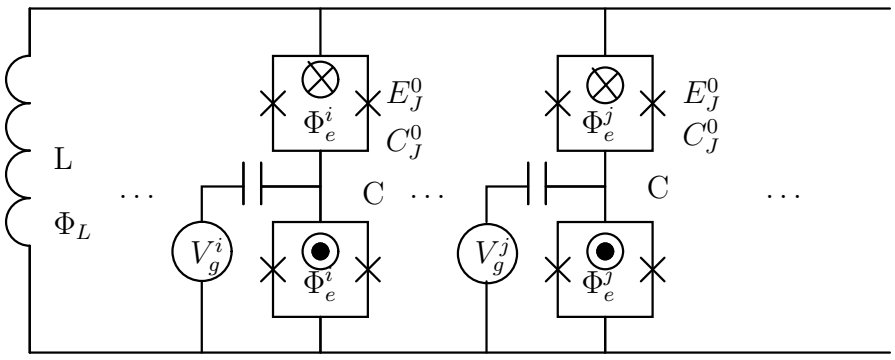


Fig. 3.14. An improved design to couple charge qubits with inductance. The top and bottom magnetic fluxes piercing through each of the two SQUID are designed to have the same amplitude but different directions. Similar to the design in Fig.3.13, the JJ term is tunable through the magnetic fluxes, and the interaction term now has the form of $\sigma_x^i\sigma_x^j$, which is more preferable.

You, et al. [103] improved this design further and obtained a simpler pulse sequence for two qubit operations. In fact, the conditional phase gate can be achieved

with just one 2-qubit pulse combined with several 1-qubit operations, leading to a much more efficient scheme. The improved design has two dc-SQUID instead of one, see Fig. 3.14. Similar to the previous design, the JJ term is tunable through the magnetic field:

$$H_J^i = -E_J^i(\Phi_e^i)(\cos(\phi_A^i) + \cos(\phi_B^i)),$$

where ϕ_A^i and ϕ_B^i are the effective phase drops of the top and bottom SQUID, respectively, in Fig. 3.14. The new effective junction energy is given by

$$E_J^i(\Phi_e^i) = 2E_J^0 \cos(\pi\Phi_e^i/\Phi_0)$$

as previously. The inductance couples all qubits and the whole system Hamiltonian of n qubits now is

$$H = \sum_{k=1}^{n} H_k + \frac{1}{2}LI^2,$$

where $H_k = E_C^k(\hat{n}_k - n_{gk})^2 - E_J^k(\Phi_e^k)(\cos(\phi_A^i) + \cos(\phi_B^i))$, E_C^k is the Coulomb energy of qubit k, and I is the persistent current through the superconducting inductance. Written in Pauli matrices form, the new overall Hamiltonian is

$$H = \sum_{k=i,j} [\epsilon_k(V_g^k)\sigma_z^k - \bar{E}_J^k(\Phi_e^k, \Phi_L, L)\sigma_x^k] + \Pi_{ij}\sigma_x^i\sigma_x^j. \tag{3.79}$$

The $\sigma_x^i\sigma_x^j$ forms the interaction term which brings the advantage that it commutes with the Josephson term, and we will show later that it make the 2-qubit gate design much more straightforward and simple. Also, note that the effective junction energy $\bar{E}_J^k$ in (3.79) is not the same as the E_J in (3.78) and also depends on the inductance L and its magnetic flux Φ_L, although it is still tunable through Φ_e^k. Similarly, the interaction coefficients Π_{ij} are also functions of Φ_L, Φ_e^i and Φ_e^j. Thus all terms are *switchable*.

By setting $\Phi_e^k = \frac{1}{2}\Phi_0$ and $n_g^k = 0.5$ for all qubits, we can let all terms vanish and

obtain $H = 0$. The system state will not change. If we need to perform an operation on qubit i, we change the corresponding Φ_e^i from $\frac{1}{2}\Phi_0$ and n_g^i from 0.5, and then the Hamiltonian becomes

$$H = \epsilon_i(V_g^i)\sigma_z^i - \bar{E}_J^i(\Phi_e^i, \Phi_L, L)\sigma_x^i.$$

Because both n_g^i and Φ_e^i can be tuned separately, the 1-qubit operators $e^{i\alpha\sigma_z^i}$ and $e^{i\beta\sigma_x^i}$ can be obtained easily by choosing $\bar{E}_J^i = 0$ or $\epsilon_i(V_g^i) = 0$, with an appropriate time duration. Any other 1-qubit operations can be constructed by combining these two operators.

Two-qubit operations can now be performed by tuning Φ_e^i and Φ_e^j away from $\Phi_0/2$. Then the Hamiltonian becomes

$$H = -\bar{E}_J^i\sigma_x^i - \bar{E}_J^j\sigma_x^j + \Pi_{ij}\sigma_x^i\sigma_x^j. \tag{3.80}$$

Theorem E.5. *For the Hamiltonian (3.80) with tunable coefficients $\bar{E}_J^i$, $\bar{E}_J^j$ and Π_{ij}, we can construct the 2-bit quantum phase gate Q_π and the CNOT gate U_{CNOT} in conjunction with 1-bit Rabi gate $U_{\theta,\phi}$ (cf. (3.58), as warranted by Corollary E.4), where*

$$Q_\pi = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}, \quad U_{CNOT} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}. \tag{3.81}$$

Proof. We choose the control parameters such that $\bar{E}_J^i = \bar{E}_J^j = \Pi_{ij} = \delta$. Then the evolution matrix for the Hamiltonian (3.80) becomes

$$U = e^{-iH\tau/\hbar} = e^{-(i\delta\tau/\hbar)(-\sigma_x^i - \sigma_x^j + \sigma_x^i\sigma_x^j)}. \tag{3.82}$$

It is easy to check that the eigenvalue equations for H now are:

$$\begin{aligned} &H|++\rangle = -\delta|++\rangle, \quad H|+-\rangle = -\delta|+-\rangle, \\ &H|-+\rangle = -\delta|-+\rangle, \quad H|--\rangle = 3\delta|--\rangle, \\ &(|\pm\rangle = \frac{1}{\sqrt{2}}(|0\rangle \pm |1\rangle)) \end{aligned} \tag{3.83}$$

By choosing $\delta\tau/\hbar = \pi/4$ in (3.82), we see that (3.82) gives the evolution matrix

$$\tilde{U} = e^{i\pi/4} \begin{bmatrix} 1 & & & \\ & 1 & & \\ & & 1 & \\ & & & -1 \end{bmatrix} \tag{3.84}$$

with respect to the ordered basis $\{|++\rangle, |+-\rangle, |-+\rangle, |--\rangle\}$. We can convert the matrix representation (3.84) to a representation with respect to the standard ordered basis $\{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}$ by

$$Q_\pi = H_i^\dagger H_j^\dagger \tilde{U} H_i H_j,$$

where H_i and H_j are, respectively, the Walsh-Hadamard gate for the i-th and j-th qubit. Since the Walsh-Hadamard gate satisfies

$$H_i = H_j = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} = e^{-i\pi/2} R_{y,\pi/2} R_{z,\pi}, \tag{3.85}$$

we have obtained Q_π as promised.

From Q_π, we have

$$U_{CNOT} = U^2_{\pi/4,\pi/2} \tilde{U} U^2_{\pi/4,-\pi/2}, \tag{3.86}$$

we also have the CNOT-gate. □

Corollary 6.5 Superconducting 1-bit gates $U_{\theta,\phi}$ obtained in Corollary E.4 together

with 2-bit gates Q_π or U_{CNOT} obtained in Theorem E.5 are universal.

Proof. This is a consequence of a result of J. Brylinski and R. Brylinski [104]. □

F. Measurement of charge qubit

The energy level of the first excited state $|1\rangle$ changes with the offset charge; when it is higher than the superconducting gap, the Cooper pair is broken apart into two quasi-particles. In Fig. 3.6, a read pulse applied on the probe gate will break the pair and let them tunnel through the junction. Repeating the experiment and measurement at frequency ν and assuming that the probability of observing the qubit at state $|1\rangle$ is P_1, we can obtain a classical current through the probe gate which is proportional to P_1:

$$I = 2eP_1\nu.$$

This measurement is *destructive*. Although state $|0\rangle$ is kept unchanged, state $|1\rangle$ is destroyed after measurement. Nakamura has used this method to observe the coherence in a SCB and quantum oscillation in two coupled charge qubits [100, 105, 106].

The above method is easy to apply, but it requires many repeated experiments and measurements. A single shot measurement requires only one measurement and would save much time. One example is realized by a group in Japan [107] using single electron transistor (SET), a sensitive electrometer, and similar setups is also investigated by other groups. See Fig. 3.15. When an appropriate pulse V_p is applied to the probe gate, such as mentioned in the preceding paragraph, the extra Cooper pair in the box is broken into two quasi-particles and tunnels into the trap. If the box is originally in state $|0\rangle$, no electron will tunnel through the junction. Then the extra charge in the trap may be detected by the SET. This completes the measurement.

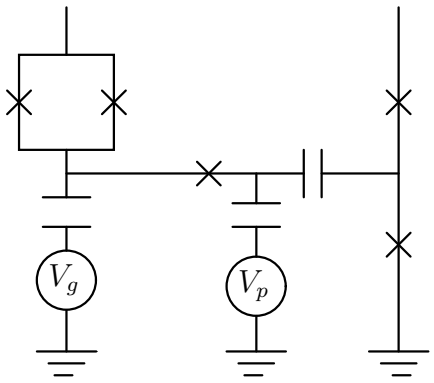


Fig. 3.15. Schematic of a circuit for measuring a charge qubit using low frequency SET. The charge qubit is coupled capacitively to an SET through a charge trap which is connected to the Cooper pair box with a tunnel junction. To reduce dissipation, the junction has high resistance. The SET is in Coulomb blockade state and there is no current through the junctions when there is no charge in the trap. When a read pulse moves extra charges from the charge qubit to the trap, the SET is biased and a current is observed through the SET.

During normal operations, the trap junction is kept unbiased and the charge qubit is isolated from the trap and SET.

The qubit may be coupled to the SET directly through a capacitor without the trap and junction, but this may induce more decoherence to the qubit. The above low frequency SET can be replaced by an rf-SET [108, 109], a more sensitive and fast electrometer, see Fig. 3.16. Different from the low frequency SET where it is the current from the source to the drain to be measured, the rf-SET measures the conductance.

There is a worrisome aspect of measurement due to the effects of noise in hight T_c superconductors as Kish and Svedlindh [110] and others have reported excessively strong magnetic and conductance noise on such superconductors.

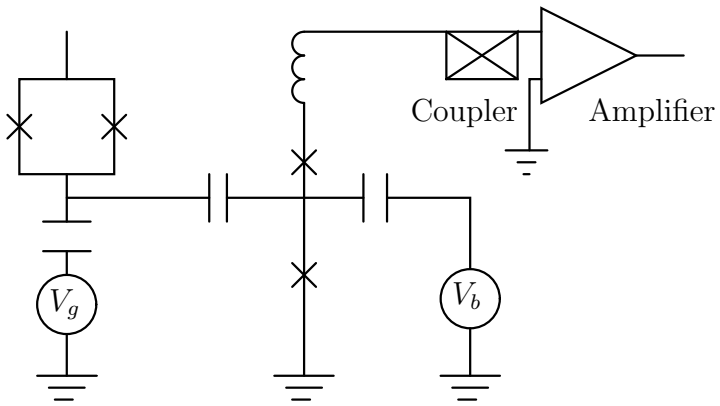


Fig. 3.16. Schematic of circuit for measuring a charge qubit using rf-SET. Different from the low frequency SET where it is the current from the source to the drain to be measured, the rf-SET measures the conductance and this makes it faster and more sensitive. A radio frequency (rf) signal resonant to the SET, referred to as "carrier" but not shown in this figure , is launched toward the SET though the coupler. Then a conductance change of the SET due to the extra charge in the charge qubit results in the change of the damping of the SET circuit, and it is reflected in the output of the amplifier.

CHAPTER IV

FEEDBACK IN COHERENT MAGNETOMETRY

A. Kalman filter and LQG controller

Before we introduce the problem formulation and analyze feedback control in magnetometry, it is proper to first give a short introduction about the LQG controller and the Kalman filter.

A *linear deterministic continuous system* is described by a differential equation:

$$\frac{d}{dt}\boldsymbol{x} = A\boldsymbol{x} + B\boldsymbol{u}, \tag{4.1}$$

where $\boldsymbol{x}$ is an n-dimensional vector which uniquely defines the system state, $\boldsymbol{u}$ is called the control vector (with dimension m), and A and B are two matrices with appropriate dimensions. When A and B do not depend on time, it is called a *linear time-invariant system*. A *controller* defines the control vector $\boldsymbol{u}$ according to the current state $\boldsymbol{x}$ or the history of $\boldsymbol{x}$ to satisfy a certain goal or to optimize a functional of $\boldsymbol{x}$ and $\boldsymbol{u}$. If $\boldsymbol{u}(t)$ only depends on $\boldsymbol{x}(t)$ at time t, it can be expressed as

$$\boldsymbol{u} = \mathcal{U}(\boldsymbol{x}(t), t), \tag{4.2}$$

which is general and covers the time variant case. A *constant linear feedback* controller has the simplest form:

$$\boldsymbol{u}(t) = K\boldsymbol{x}(t), \tag{4.3}$$

where $K \in \mathbf{R}^{m\times n}$. Thus the design of a controller falls into the search of K. An elegant example is the *linear quadratic Gaussian* (LQG) controller, which finds a linear controller in the forms of:

$$\boldsymbol{u}(t) = -K(t)\boldsymbol{x}(t),$$

and minimizes a quadratic cost function given by [111]

$$\mathcal{J} = \int_0^{t_f} (\boldsymbol{x}^T(s)Q\boldsymbol{x}(s) + \boldsymbol{u}^T(s)R\boldsymbol{u}(s))dt + \boldsymbol{x}(t_f)^T P_f \boldsymbol{x}(t_f),$$
$$Q \geq 0, R > 0, P_f \geq 0.$$

By $R > 0$ and $Q \geq 0$, we mean that R is positive definite and Q is positive semidefinite. The optimal feedback gain $K(t)$ is given by

$$K(t) = R^{-1}B^T P(t), \tag{4.4}$$

where $P(t) \geq 0$ and satisfies a matrix Riccati equation:

$$-\frac{d}{dt}P(t) = Q + A^T P(t) + P(t)A - P(t)BR^{-1}B^T P(t), \quad P(t_f) = P_f. \tag{4.5}$$

When $t_f \to \infty$, $P(t) \to P_\infty$ and $K(t)$ becomes constant in our case.

Different combinations of $\{A, B, C\}$ give different systems, and lead to different designs. Controllability and observability are among the important characteristics which must be considered carefully.

We will not bring up more details and we will just list the main results below for future reference. Readers are referred to standard control text books.

In (4.3), the whole state $\boldsymbol{x}(t)$ is needed to define $\boldsymbol{u}(t)$, which implies that the controller knows $\boldsymbol{x}(t)$ completely. In practice, not every state can be obtained directly and instead must be retrieved from the measurement history $\boldsymbol{y}$, which is connected to $\boldsymbol{x}$ and called *output*. For a linear time-invariant system, $\boldsymbol{y}$ is just a linear combination of $\boldsymbol{x}$:

$$\boldsymbol{y}(t) = C\boldsymbol{x}(t), \tag{4.6}$$

where $\boldsymbol{y} \in \mathbf{R}^l$ and $C \in \mathbf{R}^{l \times n}$. The system needs to be observable so that an estimator can be constructed to obtain the system state [112].

The combination of (4.1) and (4.6) forms a deterministic system. The measure-

ment is accurate and $\boldsymbol{x}(t)$ is driven by $\boldsymbol{u}(t)$ exclusively, which is also applied exactly. To incorporate uncertainties into the system, we use a *stochastic differential equation* (SDE) to replace (4.1):

$$d\boldsymbol{x}(t) = A\boldsymbol{x}(t)dt + B\boldsymbol{u}(t)dt + \sigma d\boldsymbol{W}, \tag{4.7}$$

where $\boldsymbol{W}$ is an r-dimensional Brownian motion and $\sigma \in \mathbf{R}^{n\times r}$. Itô form is implied for all the SDEs in this dissertation. The above equation corresponds to (4.1) by adding a "noise" term to the right side and multiplying both sides with dt. Similarly, the Itô process of the measurement with noise corresponding to (4.6) can be given as

$$\boldsymbol{y}(t)dt = C\boldsymbol{x}(t)dt + \sigma_y d\boldsymbol{W}_y, \tag{4.8}$$

where $\boldsymbol{W}_y$ is another s-dimensional Brownian motion independent of $\boldsymbol{W}$ and $\sigma_y \in \mathbf{R}^{l\times s}$. The estimation of $\boldsymbol{x}$ through the observation, or the filtering problem, is to find another n-dimensional stochastic process $\hat{\boldsymbol{x}}(t)$, which is "close" to $\boldsymbol{x}$, based on the history of $\boldsymbol{y}$. The best estimation $\hat{\boldsymbol{x}}(t)$ is defined by [113]

$$\int_\Omega |\boldsymbol{x}(t) - \hat{\boldsymbol{x}}(t)|^2 dP = E[|\boldsymbol{x}(t) - \hat{\boldsymbol{x}}(t)|^2] = \inf\{E[|\boldsymbol{x}(t) - \hat{\boldsymbol{x}}(t)|^2] : Y \in \mathcal{K}\}, \tag{4.9}$$

where $(\Omega, \mathcal{F}, P)$ is the probability space corresponding to the $(r + s)$ dimensional Brownian motion $(\boldsymbol{W}, \boldsymbol{W}_y)$ starting at 0, and $\mathcal{K}$ is the set of all stochastic variables which is measurable to the σ-algebra $\mathcal{G}_t$ and has limited second norm, i.e.,

$$\mathcal{K} = \{Y : \Omega \to \mathbf{R}^n; Y \in L^2(P) \text{ and } Y \text{ is } \mathcal{G}_t \text{ measurable}\}, \tag{4.10}$$

while $\mathcal{G}_t$ is the σ-algebra generated by $\{\boldsymbol{y}_s, s \le t\}$.

The first elegant result about the estimation is that the best estimation coincides with the conditional expectation of $\boldsymbol{x}(t)$ with respect to the measurement history.

Theorem A.1.

$$\hat{\boldsymbol{x}}(t) = E[\boldsymbol{x}(t)|\mathcal{G}_t]. \tag{4.11}$$

The result is not very useful for us in most cases because the conditional expectation is not easy to compute either. In the following, we will list the result of a special case, where we do have an explicit solution in terms of a SDE for $\hat{\boldsymbol{x}}(t)$ [113].

Theorem A.2. *The solution $\hat{\boldsymbol{x}}(t) = E[\boldsymbol{x}(t)|\mathcal{G}_t]$ of a linear system described by (4.7) and (4.8) satisfies the stochastic differential equation*

$$d\hat{\boldsymbol{x}}(t) = (A - \Sigma C^T(\sigma_y\sigma_y^T)^{-1}C)\hat{\boldsymbol{x}}(t)dt + \Sigma C^T(\sigma_y\sigma_y^T)^{-1}y(t)dt, \quad \hat{\boldsymbol{x}}(0) = E[x(0)], \tag{4.12}$$

where $\Sigma(t) = E[(\boldsymbol{x}(t) - \hat{\boldsymbol{x}}(t))(\boldsymbol{x}(t) - \hat{\boldsymbol{x}}(t))^T] \in \mathbf{R}^{n\times n}$ satisfies the matrix Ricatti equation

$$\frac{d}{dt}\Sigma(t) = A\Sigma + \Sigma A^T - \Sigma C^T(\sigma_y\sigma_y^T)C\Sigma + \sigma\sigma^T \tag{4.13}$$

with initial condition:

$$\Sigma(0) = E[(\boldsymbol{x}(0) - \hat{\boldsymbol{x}}(0))(\boldsymbol{x}(0) - \hat{\boldsymbol{x}}(0))^T].$$

The condition on σ_y is that $\sigma_y\sigma_y^T$ is invertible.

The SDE (4.12) can be reorganized as

$$d\hat{\boldsymbol{x}}(t) = A\hat{\boldsymbol{x}}(t)dt + \Sigma C^T(\sigma_y\sigma_y^T)^{-1}(y(t)dt - C\hat{\boldsymbol{x}}(t)dt). \tag{4.14}$$

The stochastic process N_t defined by $dN_t = y(t)dt - C\hat{\boldsymbol{x}}dt$ is called the *innovation process*, and $K_o(t) = \Sigma C^T(\sigma_y\sigma_y^T)^{-1}$ is called the *observer gain vector*. The stochastic process $\dfrac{N_t}{\sigma_y}$ is also a Brownian motion. When $t \to \infty$, both S and K_o converge to their steady states satisfying

$$A\Sigma + \Sigma A^T - \Sigma C^T(\sigma_y\sigma_y^T)C\Sigma + \sigma\sigma^T = 0, \tag{4.15}$$

and

$$K_o = \Sigma C^T (\sigma_y \sigma_y^T)^{-1}. \tag{4.16}$$

B. System setup and its model

In this section, we will set up a magnetometor using a cloud of spins and construct its model. A stochastic master equation (SME) model is first presented using quantum mechanical terminologies. The model is accurate but not suitable for the design of estimator and controller. Then a reduced linear stochastic differential equation (SDE) model is derived which is equivalent to the original model under proper assumptions for the purpose of the magnetic measurement. Wiseman has given a more rigorous derivation connecting SME and SDE [14]. But it requires a lot of knowledge of stochastic quantum mechanics and quantum optics, which is beyond the scope of this dissertation and not necessary for our purpose. The derivation here follows the work of the Caltech group [11, 16].

The system is set up as in Fig. 4.1. To measure an unknown and possibly fluctuating magnetic field oriented along the y-axis, a linearly polarized off-resonant light is introduced to travel through a cloud of nuclei with non-zero spin initialized along the x-axis. Because the light experiences a Faraday rotation proportional to the magnetic field along its propagation direction, the polarization difference measures the z-component of the collective spin. Without another magnetic field to compensate the external field, the spin will rotate within the $x - z$ plane and the z-component will oscillate with the spin's Larmor frequency.

A natural way to describe a quantum system interacting with its environment (bath) is the SME. For the spin system (the nuclear cloud) monitored continuously by an off-resonant light beam, an accurate model of the system has been derived using

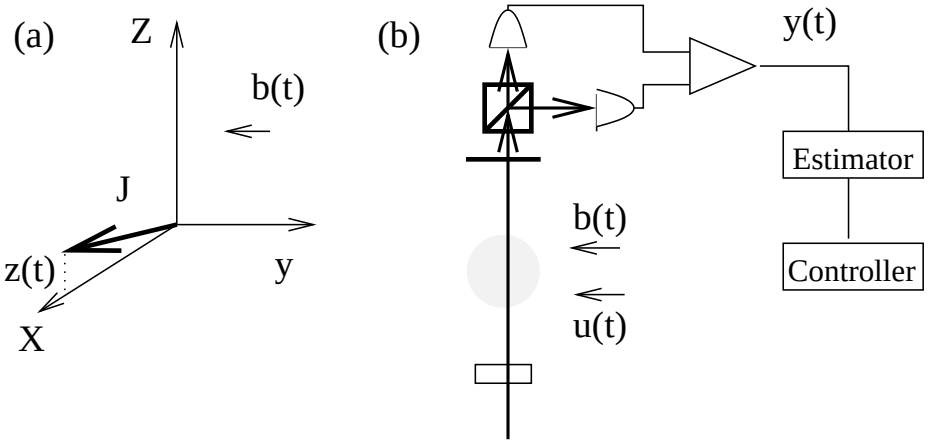


Fig. 4.1. The set up of a magnetometor by the Caltech group. The external magnetic field $b(t)$ is directed along the y-axis which is unknown and possibly fluctuating. To measure it, an ensemble of nuclei with non-zero spin are put in the field and initially polarized along the x-axis. A linearly polarized light beam travels through the nuclei cloud along the z-axis. The signal of the polarization difference is picked up at the end by a pair of photodetectors. A spin ensemble is shown in (a) with spins polarized among a small angle of x-axis. The Kalman filter (estimator) and controller are shown in (b). The picture is excerpted from [16].

a density matrix [114, 13, 14] in the form of SME:

$$d\rho(t) = -idt[H(t), \rho(t)] + \mathcal{D}[\sqrt{M}J_z]\rho(t)dt + \sqrt{\eta}\mathcal{H}[\sqrt{M}J_z]\rho(t)d\bar{W}(t), \tag{4.17}$$

$$y(t)dt = \langle J_z \rangle(t)dt + \sqrt{\sigma_m}d\bar{W}(t), \tag{4.18}$$

where $H = \gamma h J_z$, γ is the gyromagnetic ratio, $h = b(t) + u(t)$ is the total magnetic field, J_z is the z-component of the spin, ρ is the density matrix conditioned on the history of $\boldsymbol{y}$, M is the measurement rate and η is the quantum efficiency of the

detection. Operators $\mathcal{H}$ and $\mathcal{D}$ are two super operators defined by

$$\begin{aligned} \mathcal{D}(c)\rho &= c\rho c^\dagger - (c^\dagger c\rho + \rho c^\dagger c)/2, \\ \mathcal{H}(c)\rho &= c\rho + \rho c^\dagger - Tr[(c+c^\dagger)\rho]\rho, \end{aligned} \tag{4.19}$$

while $d\bar{W}(t)$ is a stochastic quantity representing the shot noise in the photodetection process.

The model of (4.17) and (4.18) is accurate but not very useful, especially when the number of spins J is huge. An observer can not see the inside of the system. He can only obtain information about the system through some measurements, or expectations of operators, and the Heisenburg uncertainty law forbids him to measure exactly two conjugate variables at the same time. In fact, a model connecting the magnetic field and the photo current through the photodectors will satisfy all our needs and only the expectation of the spin angular momentum J_z is necessary. Using

$$< J_z >= Tr(J_z\rho), \tag{4.20}$$

we have the SDE of $< J_z >$ and its second moment $< \Delta J_z^2 >$ [16]:

$$\begin{aligned} d\langle J_z\rangle(t) &= \gamma\langle J_x\rangle(t)h(t)dt + \frac{\langle \Delta J_z^2\rangle(t)}{\sqrt{\sigma_m}}d\bar{W}(t), \\ d\langle \Delta J_z^2\rangle(t) &= -\frac{\langle \Delta J_z^2\rangle^2}{\sigma_m}dt - i\gamma\langle[\Delta J_z^2, J_y]\rangle(t)h(t)dt + \frac{\langle \Delta J_z^3\rangle(t)}{\sqrt{\sigma_m}}d\bar{W}(t). \end{aligned} \tag{4.21}$$

The scalar $\langle J_z\rangle$ is a stochastic process driven by $d\bar{W}$, and according to (4.20), it is actually the conditional expectation of J_z based on the measurement history, and thus the best estimation. More than that, we can find a one-dimensional linear stochastic system whose optimal estimator is equivalent to (4.21) under some assumption.

Theorem B.1. *Assume that $\langle J_z\rangle$ is Gaussian. A linear stochastic process*

$$\begin{cases} dz(t) &= \gamma J h(t) dt \\ y(t)dt &= z(t)dt + \sqrt{\sigma_m} dW(t) \end{cases} \tag{4.22}$$

where $W(t)$ is a Brownian motion, is equivalent to the system described by (4.17) and (4.18) under the small angle assumption. The equivalence is in the meaning that its optimal estimator (Kalman filter) is same to (4.21) by replacing $\langle J_z\rangle$ by the optimal estimation of $z(t)$ and $d\bar{W}$ by a multiple of the innovation process. Specifically, the best estimation of $z(t)$ and $\langle J_z\rangle$ satisfy the same SDE when $d\bar{W}$ is replaced by $\frac{N_t}{\sqrt{\sigma_m}}$, where N_t is the innovation process of the Kalman filter.

Proof. Assume that the system is Gaussian. Then both $\langle \Delta J_z^3\rangle$ and $\langle [\Delta J_z^2, J_y]\rangle$ vanish. If the spin angle $\langle J_z\rangle/\langle J_x\rangle$ is kept small, $\langle J_x\rangle(t) \cong Je^{-Mt/2} \cong J$ when $t \ll 1/M$. The differential equation (4.21) can be simplified to

$$\begin{aligned} d\langle J_z\rangle(t) &= \gamma J h(t)dt + \frac{\langle \Delta J_z^2\rangle(t)}{\sqrt{\sigma_m}} d\bar{W}(t) \\ d\langle \Delta J_z^2\rangle(t) &= -\frac{\langle \Delta J_z^2\rangle^2(t)}{\sigma_m} dt. \end{aligned} \tag{4.23}$$

The last differential equation can be solved analytically:

$$\langle \Delta J_z^2\rangle(t) = \frac{\Delta\langle J_z^2\rangle^2(0)\sigma_m}{\sigma_m + \langle \Delta J_z^2\rangle(0)t},$$

where $\langle \Delta J_z^2\rangle(0) = J/2$ for an initially coherent spin state.

Now, if we think $\langle J_z\rangle$ is the optimal estimation of J_z and $\bar{W}(t)$ is a multiple of the innovation process, (4.23) has the structure of an optimal estimator. More than that, we can find a linear system corresponding to it.

Using (4.13) and (4.14), we can obtain the SDE of the optimal estimation $\hat{\boldsymbol{z}}(t)$

of the linear system described by (4.22):

$$\begin{aligned} d\hat{\boldsymbol{z}}(t) &= \gamma J h(t)dt + \frac{\Sigma}{\sigma_m}(y(t)dt - \hat{\boldsymbol{z}}(t)dt) \\ \frac{d}{dt}\Sigma(t) &= -\frac{\Sigma^2}{\sigma_m}. \end{aligned} \tag{4.24}$$

Comparing (4.24) and (4.23), we obtain what we want. □

- Since J is large the Gaussian assumption is practical. Term $\langle[\Delta J_z^2, J_z]\rangle h(t)$ can be ignored also because that $h(t)$ is close to zero when feedback control is applied.

- The equivalence of these two models is based on the fact that we are only interested in $\langle J_z\rangle$, which connects $h(t)$ and $y(t)$.

A direct result of the above theorem is that we can work on a simpler linear SDE model instead of the terrible SME in (4.17) and (4.18).

The magnetic field is classical and we assume that $b(t)$ is fluctuating and can be described by a stochastic process:

$$db(t) = -r_b b(t)dt + \sqrt{\sigma_{bf}}dw_1, \tag{4.25}$$

where dw_1 is another Wiener process and r_b defines the bandwidth of $b(t)$. Both $b(t)$ and $u(t)$ are classical. If $u(t)$ has infinite accuracy, we obtain a two-dimensional process by combining (4.22) and (4.25) and treating $b(t)$ as an internal state:

$$d\boldsymbol{x}(t) = A\boldsymbol{x}(t)dt + Bu(t)dt + \begin{bmatrix} 0 \\ \sqrt{\sigma_{bf}} \end{bmatrix} dw_1, \tag{4.26}$$

$$y(t)dt = C\boldsymbol{x}(t)dt + \sqrt{\sigma_m}dw_2, \tag{4.27}$$

where we use $\boldsymbol{x}$ to denote the state:

$$\boldsymbol{x} = \begin{bmatrix} z(t) \\ b(t) \end{bmatrix},$$

while $z(t)$ is the small z component, $b(t)$ is the magnetic field along the y-axis, and y is the output or measurement.

The matrices can be given as

$$A = \begin{bmatrix} 0 & \gamma J \\ 0 & -r_b \end{bmatrix}, B = \begin{bmatrix} \gamma J \\ 0 \end{bmatrix}, \tag{4.28}$$

$$C = [1, 0]. \tag{4.29}$$

The initial variances of the state is

$$\Sigma_0 = \begin{bmatrix} \sigma_{z0} & 0 \\ 0 & \sigma_{b0} \end{bmatrix}. \tag{4.30}$$

The initial field variant is σ_{bf}, from which we define

$$\begin{aligned} \Sigma_1 &= \begin{bmatrix} 0 & 0 \\ 0 & \sigma_{bf} \end{bmatrix}, \\ \Sigma_2 &= \sigma_m = \frac{1}{4} M \eta. \end{aligned} \tag{4.31}$$

C. Results by the Caltech group

When the control loop is open, the spin will rotate around the y-axis. If the magnetic field is strong compared with the damping speed, an oscillating signal can be obtained before the spin damps significantly. Since the oscillation frequency (Larmor frequency) is proportional to the magnetic field, the amplitude of the field can be found by computing the Fourier transform of the signal. But if the field is weak, only

a noisy sloped line of a small angle rotation can be recorded. By fitting the curve with a line, we can find an estimation of its slope, denoted as s. If we know J, the field can be retrieved through identity $\tilde{b} = s/\gamma J$.

The first problem that the above slope method meets is an unknown J. Suppose we use J' instead of J. Since $\tilde{b} = s/\gamma J' = (J/J')b$, the variance of the estimation error depends on J',

$$E[(b - \tilde{b})^2] = (1 - \frac{1}{f})E[b^2], \tag{4.32}$$

and the expectation depends on J' too,

$$E[b - \tilde{b}] = (1 - \frac{J}{J'})E[b] = (1 - \frac{1}{f})E[b], \tag{4.33}$$

where $f = J'/J$. If the magnetic field is constant, the error is systematic and can be calibrated away. If both b and J are random, the error is no more systematic and the slope does not make any sense.

The linear model in the previous section builds a suitable platform to study the estimation of $b(t)$ under the small angle assumption and with noise-free input. That is what the Caltech group used. Here we first list their results for future comparison. Define $\Theta(t)$ as the total state variance:

$$\Theta(t) = E[\theta(t)\theta(t)^T] = \begin{bmatrix} \sigma_{zz} & \sigma_{zb} & \sigma_{z\tilde{z}} & \sigma_{z\tilde{b}} \\ \sigma_{bz} & \sigma_{bb} & \sigma_{b\tilde{z}} & \sigma_{b\tilde{b}} \\ \sigma_{\tilde{z}z} & \sigma_{\tilde{z}b} & \sigma_{\tilde{z}\tilde{z}} & \sigma_{\tilde{z}\tilde{b}} \\ \sigma_{\tilde{b}z} & \sigma_{\tilde{b}b} & \sigma_{\tilde{b}\tilde{z}} & \sigma_{\tilde{b}\tilde{b}} \end{bmatrix}. \tag{4.34}$$

We can derive its dynamic equation as

$$\frac{d\Theta(t)}{dt} = \alpha(t)\Theta(t) + \Theta(t)\alpha(t)^T + \beta(t)\beta(t)^T, \tag{4.35}$$

with initial condition

$$\Theta_0 = \begin{bmatrix} \sigma_{z0} & 0 & 0 & 0 \\ 0 & \sigma_{b0} & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}. \tag{4.36}$$

The matrices α and β are defined as bellow:

$$\alpha = \begin{bmatrix} A & -BK_c \\ K_oC & A' - BK_c - K_oC \end{bmatrix}, \quad \beta = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & \sqrt{\sigma_b f} & 0 & 0 \\ 0 & 0 & \sqrt{\sigma_m} K_z & 0 \\ 0 & 0 & \sqrt{\sigma_m} K_b & 0 \end{bmatrix}, \tag{4.37}$$

where K_c and K_o are the controller gain and observer gain vector, K_z and K_b are two entries of K_o. The matrix A' is used in the computation of K_c and K_o, while A is the "real" matrix.

The matrix Θ has all the information about the estimation error variance of $b(t)$ and $z(t)$:

$$\begin{aligned} \sigma_{be} &= E[(b(t) - \tilde{b}(t))^2] = \sigma_{bb} + \sigma_{\tilde{b}\tilde{b}} - 2\sigma_{b\tilde{b}}, \\ \sigma_{ze} &= E[(z(t) - z(t))^2] = \sigma_{zz} + \sigma_{\tilde{z}\tilde{z}} - 2\sigma_{z\tilde{z}}. \end{aligned} \tag{4.38}$$

If every parameter is known exactly, it does not matter whether the system is closed or open. The only necessity of the closed control is to keep the small angle assumption. The steady state performance can be approximated by

$$\begin{aligned} \sigma_{zs} &= \sqrt{2\gamma J}\sigma_m^{3/4}\sigma_{bf}^{1/4}, \\ \sigma_{bs} &= \sqrt{\frac{2}{\gamma J}}\sigma_m^{1/4}\sigma_{bf}^{3/4}. \end{aligned} \tag{4.39}$$

It is interesting to note that this is a squeezed state with Heisenberg limit $\sigma_{zs}\sigma_{bs} \geq 2\sigma_{bf}\sigma_m$.

The complicated case happens when J is unknown. If the control loop is closed,

$$\sigma_{bs}(J, J') = \left(\frac{1+f}{2}\right)\sigma_{bs}(J'), \tag{4.40}$$

where $\sigma_{bs}(J')$ is the variance in (4.39) when $J = J'$. The *large* λ *assumption* ($\lambda^2 \gg \sqrt{\sqrt{\sigma_{bf}/\sigma_m}/(2J)}$) and *large* J *assumption* ($r_b^2 \ll J\sqrt{\sigma_{bf}/\sigma_m}$) hold in all the computations.

Using (4.40) and noting that $\frac{1+f}{2}$ is a increasing function when $f > 0$, the Caltech group chose a robust choice of J' as the minimum of J. Then we will do no worse than $\sigma_{bs}(J')$ ($f < 1$), and no better than $1/2\sigma_{bs}(J')$. In the following section, we will only work on the steady solution and replace σ_{bs} by σ_b and σ_{zs} by σ_z without confusion.

D. Steady performance with noised feedback

The setup of the system by the Caltech group is inspirational, but the robust design is too simple and not persuasive. A simple explanation of the robustness is that the precise control maintains a magnetic field $u(t)$ that compensates the external field. The smaller J' we choose, the stronger control we obtain. The result is heavily rested on the assumption that the control is noise free. This is not practical. In our work, we try to answer the following concerns:

- There is always noise accompanying the control signal. When the noise is sufficient small, it will not affect the result. We want to find out how small it has to be before the closed loop control is not worth the robustness.

- The robust choice of J' by the Caltech group is still too ad hoc. A more accurate discussion will be given.

In this section, we will first extend the previous system with an input noise and compute the static Kalman filter gain K_o and controller gain K_c. Then the steady performance of the magnetometor, i.e., the covariance of the estimation errors will be derived. To do that, the system dimension is reduced to three and a six-variable linear equation set is solved. Simplifications are used in the computation based on the large λ and large J assumption.

We introduce a input noise $\sigma_{uf} dw_3$ in the process, and this changes the original system described by (4.26) and (4.25) to

$$\begin{aligned} d\boldsymbol{x}(t) &= A\boldsymbol{x}(t)dt + Bu(t)dt + \begin{bmatrix} \sqrt{\sigma_{uf}} & 0 \\ 0 & \sqrt{\sigma_{bf}} \end{bmatrix} \begin{bmatrix} dw_3 \\ dw_1 \end{bmatrix}, \\ y(t)dt &= C\boldsymbol{x}(t)dt + \sqrt{\sigma_m} dw_2(t), \end{aligned} \tag{4.41}$$

where dw_3 is a Wiener process. A constant feed back is chosen in our model since we are only interested in the steady performance:

$$u(t)dt = -K_c \tilde{x}(t)dt. \tag{4.42}$$

The matrices are listed as bellow:

$$A = \begin{bmatrix} 0 & J \\ 0 & -rb \end{bmatrix}, B = \begin{bmatrix} J \\ 0 \end{bmatrix}, \tag{4.43}$$

$$C = [1, 0]. \tag{4.44}$$

For simplicity, we replace γJ by J and $\gamma J'$ by J' in this section, because only the products, γJ and $\gamma J'$, appear in the computation.

Lemma D.1. *The static solution of the Kalman filter gain vector K_o of the linear*

system defined by (4.41) is

$$\begin{aligned} K_z = &\quad -r_b + \sqrt{r_b^2 + \sigma_{uf}\sigma_m^{-1} + 2\sqrt{J'^2\sigma_{bf}\sigma_m^{-1} + r_b^2\sigma_{uf}\sigma_m^{-1}}}, \\ K_b = &\quad \frac{1}{2J'}(-\sigma_{uf}\sigma_m^{-1} + K_z^2), \end{aligned} \tag{4.45}$$

when J' instead of J is used in the design of the Kalman filter. Under the large J and λ assumption, the LQG controller gain vector K_c can be approximated by:

$$K_c \approx [\lambda, 1], \tag{4.46}$$

where $\lambda = r/q$ for cost function:

$$\mathcal{J} = \int_0^\infty (qz(t)^2 + ru(t)^2)dt.$$

Proof. We first compute K_c. According to (4.4),

$$K_c \equiv R^{-1}B^T V,$$

where V is the solution of

$$P + A^T V + VA - VBR^{-1}B^T V = 0,$$

leading to

$$K_c = [\lambda, \frac{J'\lambda}{r_b + J'\lambda}]. \tag{4.47}$$

Since r_b is far less than $J'\lambda$, the above equation can be simplified to

$$K_c \approx [\lambda, 1], \tag{4.48}$$

as we wanted.

Using the formulae of the Kalman filter, we obtain

$$K_o = \Sigma(t)C^T\Sigma_2^{-1},$$

and

$$\frac{d}{dt}\Sigma(t) = \Sigma_1 + A'\Sigma(t) + \Sigma(t)A'^T - \Sigma(t)C^T\Sigma_2^{-1}C\Sigma(t),$$

where A' is the same as A except that J is replaced by J', Σ_2 is defined in (4.31), and Σ_1 is given by

$$\Sigma_1 = \begin{bmatrix} \sigma_{uf} & 0 \\ 0 & \sigma_{bf} \end{bmatrix}.$$

Other matrices are defined as before. Let $\Sigma(t)$ be constant and denoted as

$$\Sigma(t) = \begin{bmatrix} \sigma_{zs} & \sigma_{cs} \\ \sigma_{cs} & \sigma_{bs} \end{bmatrix},$$

and put in all the numbers, we find a set of equations:

$$\begin{cases} \sigma_{uf} + 2J'\sigma_{cs} - \sigma_{zs}^2\sigma_m^{-1} = 0 \\ J'\sigma_{bs} - r_b\sigma_{cs} - \sigma_{cs}\sigma_{zs}\sigma_m^{-1} = 0 \\ \sigma_{bf} - 2r_b\sigma_{bs} - \sigma_{cs}^2\sigma_m^{-1} = 0 \end{cases} \tag{4.49}$$

Among the three unknown entries, we only need σ_{zs} and σ_{cs},

$$\begin{aligned} \sigma_{zs} = &\quad -r_b\sigma_m + \sigma_m\sqrt{r_b^2 + \sigma_{uf}\sigma_m^{-1} + 2\sqrt{J'^2\sigma_{bf}\sigma_m^{-1} + r_b^2\sigma_{uf}\sigma_m^{-1}}}, \\ \sigma_{cs} = &\quad \frac{1}{2J'}(-\sigma_{uf} + \sigma_{zs}^2\sigma_m^{-1}), \end{aligned} \tag{4.50}$$

and the entries of K_o are obtained from σ_{zs} and σ_{cs}:

$$\begin{aligned} K_z = \frac{\sigma_{zs}}{\sigma_m} = &\quad -r_b + \sqrt{r_b^2 + \sigma_{uf}\sigma_m^{-1} + 2\sqrt{J'^2\sigma_{bf}\sigma_m^{-1} + r_b^2\sigma_{uf}\sigma_m^{-1}}}, \\ K_b = \frac{\sigma_{cs}}{\sigma_m} = &\quad \frac{1}{2J'}(-\sigma_{uf}\sigma_m^{-1} + K_z^2). \end{aligned} \tag{4.51}$$

□

There are several useful identities about K_z and K_b, and we list them in the following lemmas.

Lemma D.2. *1.*

$$\sigma_{uf} = K_z^2\sigma_m - 2J'K_b\sigma_m, \tag{4.52}$$

2.

$$J'\sigma_{bf} = (J'K_b^2 + 2r_b^2K_b + 2r_bK_bK_z)\sigma_m, \tag{4.53}$$

3.

$$K_zr_b + K_bJ' = \sqrt{J'^2\sigma_{bf}\sigma_m^{-1} + r_b^2\sigma_{uf}\sigma_m^{-1}}. \tag{4.54}$$

Proof. The first one is just the first equation in (4.49) in terms of K_z and K_b. We only need to prove the last two. Using (4.52) and (4.51), we have

$$\begin{aligned} K_zr_b + K_bJ' &= \frac{1}{2}(2K_zr_b + K_z^2 - \sigma_{uf}\sigma_m^{-1}) \\ &= \frac{1}{2}((K_z + r_b)^2 - r_b^2 - \sigma_{uf}\sigma_m^{-1}) \\ &= \sqrt{J'^2\sigma_{bf}\sigma_m^{-1} + r_b^2\sigma_{uf}\sigma_m^{-1}}, \end{aligned} \tag{4.55}$$

and this proves the last one. For the second, look into identity:

$$\begin{aligned} \sigma_{bf} &= \sigma_{cs}^2\sigma_m^{-1} + 2r_b\sigma_{bs} \\ &= \sigma_{cs}^2\sigma_m^{-1} + 2r_b\frac{1}{J'}(r_b\sigma_{cs} + \sigma_{cs}\sigma_{zs}\sigma_m^{-1}) \end{aligned} \tag{4.56}$$

which is equal to

$$J'\sigma_{bf} = J'K_b^2\sigma_m + 2r_b^2K_b\sigma_m + 2r_bK_bK_z\sigma_m.$$

□

Lemma D.3.

$$\frac{(K_b^2\sigma_m + \sigma_{bf})(r_bK_z + J'K_b) + K_b^2\sigma_{uf} + K_z^2\sigma_{bf}}{2(K_zr_b + K_bJ')(K_z + r_b)} = \frac{K_b(K_z + r_b)\sigma_m}{J'}. \tag{4.57}$$

Proof. Using lemma D.2, we have

$$\begin{aligned} & 2(K_z r_b + K_b J')(K_z + r_b)^2 K_b \sigma_m - J'(\sigma_m K_b^2 + \sigma_{bf})(r_b K_z + J' K_b) \\ = \; & (K_z r_b + J' K_b)(2K_b(K_z + r_b)^2 \sigma_m - J'(K_b^2 \sigma_m + \sigma_{bf})) \\ = \; & (K_z r_b + J' K_b)(2K_b(K_z + r_b)^2 - 2(J' K_b^2 + r_b^2 K_b + r_b K_b K_z))\sigma_m \\ = \; & 2(K_z r_b + J' K_b)(K_b(K_z^2 + K_z r_b - K_b J'))\sigma_m \\ = \; & 2K_b(r_b K_z^3 + r_b^2 K_z^2 + J' K_b K_z^2 - J'^2 K_b^2)\sigma_m, \end{aligned} \tag{4.58}$$

and

$$\begin{aligned} & K_b^2 J' \sigma_{uf} + K_z^2 J' \sigma_{bf} \\ = \; & \sigma_m(K_b^2 J'(K_z^2 - 2J' K_b) + K_z^2(K_b^2 J' + 2r_b^2 K_b + 2r_b K_b K_z)) \\ = \; & \sigma_m(J' K_z^2 K_b^2 - 2J'^2 K_b^3 + J' K_z^2 K_b^2 + 2r_b^2 K_b K_z^2 + 2r_b K_b K_z^3) \\ = \; & 2K_b(r_b K_z^3 + r_b^2 K_z^2 + J' K_b K_z^2 - J'^2 K_b^2)\sigma_m. \end{aligned} \tag{4.59}$$

Thus

$$\begin{aligned} & 2(K_z r_b + K_b J')(K_z + r_b)^2 K_b \sigma_m - J'(\sigma_m K_b^2 + \sigma_{bf})(r_b K_z + J' K_b) \\ & = K_b^2 J' \sigma_{uf} + K_z^2 J' \sigma_{bf}, \end{aligned} \tag{4.60}$$

and that is equal to (4.57). □

Lemma D.4. *If $J \gg r_b^2 \sqrt{\frac{\sigma_m}{\sigma_{bf}}}$ (large J assumption), $K_z \gg r_b$.*

Proof. If $\beta \gg \alpha > 0$, then

$$\sqrt{\alpha + \beta} - \sqrt{\alpha} = \sqrt{\alpha}(\sqrt{1 + \beta/\alpha} - 1) \approx \sqrt{\beta} \gg \sqrt{\alpha}.$$

Let $\alpha = r_b^2$ and $\beta = 2\sqrt{J'^2 \sigma_{bf}/\sigma_m + r_b^2 \sigma_{uf}/\sigma_m} + \sigma_{uf}/\sigma_m$, and we obtain what we want. □

The large J assumption implies that $r_b^2 \ll \sqrt{J'\sigma_{bf}\sigma_m^{-1} + r_b^2\sigma_{uf}\sigma_m^{-1}}$, leading to an approximation of K_z,

$$K_z \approx \sqrt{\sigma_{uf}\sigma_m + 2\sqrt{J'^2\sigma_{bf}\sigma_m^3 + r_b^2\sigma_{uf}\sigma_m^3}}\frac{1}{\sigma_m}, \tag{4.61}$$

and an approximation of K_b,

$$K_b \approx \sqrt{\sigma_{bf}\sigma_m + r_b^2 J'^{-2}\sigma_{uf}\sigma_m}\frac{1}{\sigma_m}. \tag{4.62}$$

Now we can write the whole system model with the estimator and controller:

$$\begin{aligned} d\begin{bmatrix} z(t) \\ b(t) \\ \tilde{z}(t) \\ \tilde{b}(t) \end{bmatrix} &= \begin{bmatrix} 0 & J & -J\lambda & -J \\ 0 & -r_b & 0 & 0 \\ K_z & 0 & -J'\lambda - K_z & 0 \\ K_b & 0 & -K_b & r_b \end{bmatrix}\begin{bmatrix} z(t) \\ b(t) \\ \tilde{z}(t) \\ \tilde{b}(t) \end{bmatrix} dt \\ &+ \begin{bmatrix} \sqrt{\sigma_{uf}} & & & \\ & \sqrt{\sigma_{bf}} & & \\ & & K_z\sqrt{\sigma_m} & \\ & & K_b\sqrt{\sigma_m} & \end{bmatrix}\begin{bmatrix} dw_3 \\ dw_2 \\ dw_1 \end{bmatrix}, \end{aligned} \tag{4.63}$$

where K_z and K_b are the two entries of the estimator gain vector. This four dimensional system in (4.63) is not necessary for our purpose. To reduce its size, we define $\Delta z = z - \tilde{z}$ and $\Delta b = b - \tilde{b}$, and a three dimensional system is obtained as

$$\begin{aligned} d\begin{bmatrix} \Delta z \\ \Delta b \\ \tilde{z} \end{bmatrix} &= \begin{bmatrix} -K_z & J & \lambda(J'-J) \\ -K_b & -r_b & 0 \\ K_z & 0 & -\lambda J' \end{bmatrix}\begin{bmatrix} \Delta z dt \\ \Delta b dt \\ \tilde{z} dt \end{bmatrix} \\ &+ \begin{bmatrix} \sqrt{\sigma_{uf}} & 0 & -K_z\sqrt{\sigma_m} \\ 0 & \sqrt{\sigma_{bf}} & -K_b\sqrt{\sigma_m} \\ 0 & 0 & K_z\sqrt{\sigma_m} \end{bmatrix}\begin{bmatrix} dw_3 \\ dw_1 \\ dw_2 \end{bmatrix}. \end{aligned} \tag{4.64}$$

We denote the vector and matrices as bellow:

$$\theta = \begin{bmatrix} \Delta z \\ \Delta b \\ \tilde{z} \end{bmatrix}, \quad \alpha = \begin{bmatrix} -K_z & J & \lambda(J'-J) \\ -K_b & -r_b & 0 \\ K_z & 0 & -\lambda J' \end{bmatrix}, \quad \beta = \begin{bmatrix} \sqrt{\sigma_{uf}} & 0 & -K_z\sqrt{\sigma_m} \\ 0 & \sqrt{\sigma_{bf}} & -K_b\sqrt{\sigma_m} \\ 0 & 0 & K_z\sqrt{\sigma_m} \end{bmatrix}. \tag{4.65}$$

The dynamic equation of the whole system state covariance can be derived as we did before:

$$\frac{d\Theta}{dt} = \alpha\Theta + \Theta\alpha^T + \beta\beta^T, \tag{4.66}$$

where $\Theta = E[\theta\theta^T]$. The steady solution can be found by solving the algebraic equation

$$0 = \alpha\Theta + \Theta\alpha^T + \beta\beta^T. \tag{4.67}$$

Because Θ is symmetric, we can assume

$$\Theta = \begin{bmatrix} \theta_{11} & \theta_{12} & \theta_{13} \\ \theta_{12} & \theta_{22} & \theta_{23} \\ \theta_{13} & \theta_{23} & \theta_{33} \end{bmatrix},$$

and (4.67) can be transformed to six linear equations:

$$\begin{cases} -K_z\theta_{11} + J\theta_{12} + \lambda(J'-J)\theta_{13} = -(\sigma_{uf} + K_z^2\sigma_m)/2 \\ -K_z\theta_{12} + J\theta_{22} + \lambda(J'-J)\theta_{23} - K_b\theta_{11} - r_b\theta_{12} = -K_zK_b\sigma_m \\ -K_z\theta_{13} + J\theta_{23} + \lambda(J'-J)\theta_{33} + K_z\theta_{11} - \lambda J'\theta_{13} = K_z^2\sigma_m \\ K_b\theta_{12} + r_b\theta_{22} = (\sigma_{bf} + K_b^2\sigma_m)/2 \\ -K_b\theta_{13} - r_b\theta_{23} + K_z\theta_{12} - \lambda J'\theta_{23} = K_zK_b\sigma_m \\ K_z\theta_{13} - \lambda J'\theta_{33} = -K_z^2\sigma_m/2 \end{cases} \tag{4.68}$$

Note that we only need $\sigma_b = E[(\Delta b)^2] = \theta_{22}$ and $\sigma_z = E[(\Delta z)^2] = \theta_{11}$. The solution of θ_{22} can be found as

$$\theta_{22} = \frac{Poly_N}{Poly_D}, \tag{4.69}$$

where

$$\begin{aligned} Poly_D = \; & 2J(r_bK_z + J'K_b)((J'K_b + (K_z + \lambda J')(r_b + \lambda J'))(K_z + r_b) \\ & +(J - J')(\lambda K_z(\lambda J' + K_z) + K_b(r_b + K_z))), \end{aligned} \tag{4.70}$$

and

$$\begin{aligned} Poly_N = \; & J((\sigma_m K_b^2 + \sigma_{bf})(r_bK_z + J'K_b) + K_b^2\sigma_{uf} + K_z^2\sigma_{bf}) \\ & (J'K_b + (K_z + \lambda J')(r_b + \lambda J')) \\ & +(J - J')(\lambda J\sigma_{bf}K_z^3 + \lambda J\sigma_m r_b K_z^2 K_b^2 + \lambda^2 JJ'\sigma_{bf}K_z^2 \\ & +r_bJ\sigma_m K_z K_b^3 + \lambda JJ'\sigma_m K_z K_b^3 - \lambda J'\sigma_{uf}K_z K_b^2 \\ & +r_bJ\sigma_{bf}K_z K_b - \lambda JJ'\sigma_{bf}K_z K_b + JJ'\sigma_m K_b^4 + JJ'\sigma_{bf}K_b^2 \\ & -\lambda J'\sigma_{uf}r_b K_b^2 - \lambda^2 J'^2\sigma_{uf}K_b^2) \\ = \; & J(\cdots) \\ & +(J - J')(J(K_b^2\sigma_m + \sigma_{bf})K_b(r_bK_z + J'K_b) \\ & -2\lambda\frac{J}{J'}K_z K_b r_b(K_z + r_b)(r_bK_z + J'K_b)\sigma_m \\ & +\lambda(K_z + r_b + \lambda J')(JK_z^2\sigma_{bf} - J'K_b^2\sigma_{uf})). \end{aligned} \tag{4.71}$$

Observing that

$$\frac{\alpha\mathcal{A} + \beta\mathcal{B}}{\alpha + \beta\mathcal{C}} = \mathcal{A} + \frac{\beta(\mathcal{B} - \mathcal{A}\mathcal{C})}{\alpha + \beta\mathcal{C}},$$

comparing it with (4.69) through (4.71) and using lemma D.3, we find

$$\theta_{22} = \frac{K_b(K_z + r_b)}{J'}\sigma_m + \frac{(J - J')Poly_N^2}{Poly_D}, \tag{4.72}$$

where

$$\begin{aligned} Poly_N^2 = \; & J(K_b^2\sigma_m + \sigma_{bf})K_b(r_bK_z + J'K_b) \\ & -2\lambda\frac{J}{J'}K_zK_br_b(K_z + r_b)(r_bK_z + K_bJ')\sigma_m \\ & +\lambda(K_z + r_b + \lambda J')(JK_z^2\sigma_{bf} - J'K_b^2\sigma_{uf}) \\ & -\frac{K_b(K_z + r_b)}{J'}\sigma_m(K_br_b + K_bK_z + \lambda^2J'K_z + \lambda K_z^2)2J(K_zr_b + J'K_b). \end{aligned} \tag{4.73}$$

Since

$$\begin{aligned} J(K_b^2\sigma_m + \sigma_{bf}) &= \frac{J}{J'}(K_b^2J'\sigma_m + J'\sigma_{bf}) \\ &= \frac{2J}{J'}(K_b^2J' + r_b^2K_b + r_bK_bK_z)\sigma_m, \end{aligned} \tag{4.74}$$

the numerator $Poly_N^2$ can be rewritten as

$$\begin{aligned} Poly_N^2 = \; & 2\frac{J}{J'}\sigma_mK_b(r_bK_z + J'K_b)(\lambda(K_z + r_b)K_z(-r_b - K_z - \lambda J') \\ & +K_b^2J' - K_bK_z^2 - K_bK_zr_b) \\ & +\frac{J}{J'}(K_z + r_b + \lambda J')\lambda(J'K_z^2\sigma_{bf} - \frac{J'^2}{J}K_b^2\sigma_{uf}). \end{aligned} \tag{4.75}$$

By using lemma D.2, we can further write $Poly_N^2$ as a multiple of σ_m,

$$\begin{aligned} \frac{Poly_N^2}{\sigma_m} = \; & \lambda K_b^2(K_z + r_b + \lambda J')(2J'^2K_b - J'K_z^2 - 2r_bJK_z - JK_z^2) \\ & +2\frac{J}{J'}K_b(K_bJ' + r_bK_z)(K_b^2J' - K_bK_z^2 - r_bK_zK_b), \end{aligned} \tag{4.76}$$

and we write (4.72) in a multiple of σ_m too:

$$\frac{\theta_{22}}{\sigma_m} = \frac{K_b(K_z + r_b)}{J'} + \frac{(J - J')Poly_N^2}{Poly_D\sigma_m}. \tag{4.77}$$

Before we use the large λ and large J assumption to simplify (4.77), we first look into two identities:

$$\begin{aligned} & 2J'^2K_b - J'K_z^2 - JK_z^2 - 2r_bJK_z \\ &= J'(2J'K_b - K_z^2) - JK_z(K_z + 2r_b) \\ &= -J'\sigma_{uf}\sigma_m^{-1} - JK_z(K_z + 2r_b) \\ &= -(J + J')\sigma_{uf}\sigma_M^{-1} - 2J\sqrt{T}, \end{aligned} \tag{4.78}$$

and

$$\begin{aligned} & K_b^2J' - K_bK_z^2 - r_bK_zK_b \\ &= K_b(J'K_b - K_z^2 - r_bK_z) \\ &= K_b(\frac{1}{2}(K_z^2 - \sigma_{uf}\sigma_m^{-1}) - K_z^2 - r_bK_z) \\ &= K_b(-\frac{1}{2}(K_z^2 + 2r_bK_z + r_b^2) + \frac{r_b^2}{2} - \frac{\sigma_{uf}}{2\sigma_m}) \\ &= \frac{K_b}{2}(-2\frac{\sigma_{uf}}{\sigma_m} - 2\sqrt{T}), \end{aligned} \tag{4.79}$$

where $T = J'^2\sigma_{bf}\sigma_m^{-1} + r_b^2\sigma_{uf}\sigma_m^{-1}$. Now (4.76) can be rewritten as:

$$\begin{aligned} \frac{Poly_N^2}{\sigma_m} = & -(\lambda K_b^2J((1 + \frac{J'}{J})\frac{\sigma_{uf}}{\sigma_m} + 2\sqrt{T})(K_z + r_b + \lambda J') \\ & + \frac{J}{J'}K_b^2(r_bK_z + K_bJ')(2\frac{\sigma_{uf}}{\sigma_m} + 2\sqrt{T})) \\ = & -2K_b^2(\frac{\sigma_{uf}}{\sigma_m} + \sqrt{T})(\frac{1}{f}\sqrt{T} + \lambda J(K_z + r_b + \lambda J')) \\ & + (1 - f)\lambda K_b^2J(K_z + r_b + \lambda J')\frac{\sigma_{uf}}{\sigma_m}, \end{aligned} \tag{4.80}$$

where $f = J'/J$.

An approximation of $Poly_D$ can be obtained by reorganizing its terms:

$$\begin{aligned}
Poly_D &= 2J(K_z r_b + J' K_b)((K_z + r_b)(K_z r_b + J' K_b) \\
&+\lambda J'(K_z + r_b + \lambda J')(K_z + r_b) \\
&+(J - J')(\lambda K_z(\lambda J' + K_z) + K_b(r_b + K_z))) \\
&= 2J\sqrt{T}((K_z + r_b)\sqrt{T} + \lambda J'(K_z + r_b)^2 + \lambda^2 J'^2(K_z + r_b) \\
&+(J - J')(\lambda K_z(\lambda J' + K_z) + K_b(r_b + K_z))) \\
&= 2J\sqrt{T}((K_z + r_b)(\sqrt{T} + (J - J')K_b) \\
&+\lambda(JK_z^2 + \lambda J'(JK_z + J' r_b) + J' r_b^2 + 2J' r_b K_z)) \\
&\approx 2J\sqrt{T}(K_z(\sqrt{T} + (\frac{1}{f} - 1)\sqrt{T}) + \lambda J K_z(K_z + \lambda J')) \\
&\approx 2J\sqrt{T}K_z(\frac{J}{J'}\sqrt{T} + \lambda J(K_z + \lambda J')).
\end{aligned} \tag{4.81}$$

We have used the large λ and large J assumption in above approximation.

Combining (4.80) and (4.81), we obtain:

$$\frac{Poly_N^2}{Poly_D} = -\frac{K_b^2(\sigma_{uf} + \sigma_m\sqrt{T})}{J\sqrt{T}K_z} + (f - 1)\frac{\lambda K_b^2(K_z + r_b + \lambda J')\sigma_{uf}}{2J\sqrt{T}K_z(\frac{\sqrt{T}}{J'} + \lambda(K_z + \lambda J'))}. \tag{4.82}$$

The second term in the above equation can be simpler if $\lambda(K_z + \lambda J') \gg \frac{\sqrt{T}}{J'}$:

$$\frac{\lambda K_b^2(K_z + r_b + \lambda J')\sigma_{uf}}{2J\sqrt{T}K_z(\frac{\sqrt{T}}{J'} + \lambda(K_z + \lambda J'))} \approx \frac{K_b^2\sigma_{uf}}{2J\sqrt{T}K_z}, \tag{4.83}$$

and σ_b can be computed by substituting (4.82) and (4.83) back to (4.73). We state this in our first theorem.

Theorem D.5. *When J' is used in the design of the estimator (Kalman filter) and the controller, the steady estimation error variance σ_b obtained for the linear system*

(4.63) is given by

$$\sigma_b/\sigma_m = \frac{(K_z + r_b)K_b}{J'} - (1-f)\frac{K_b^2}{K_z} + \frac{K_b^2\sigma_{uf}\sigma_m^{-1}}{2\sqrt{T}K_z}(f^2-1). \tag{4.84}$$

The large λ and large J assumption are supposed to hold as in the Theorem E.2 and $\lambda(K_z + \lambda J') \gg \frac{\sqrt{T}}{J'}$. Entries K_z and K_b are defined as before.

- The result is based on the large λ and large J assumption.
- The inequality $\lambda(K_z + \lambda J') \gg \frac{\sqrt{T}}{J'}$ is reasonable, since λ^2 is far larger than J^{-1} and $\sqrt{T}$ is in the order of J.

E. Steady performance with open control loop

In this section, we will find the steady performance of the Kalman filter when $J \neq J'$ and the control loop is open. The result will be used for comparison in the future. The dynamic function of the system with an open control loop can be given as

$$\begin{cases} dz(t) = Jb(t)dt \\ db(t) = -r_b b(t)dt + \sqrt{\sigma_{bf}}dw_1 \\ d\tilde{z}(t) = J'\tilde{b}(t)dt + K_z(y(t) - z(t))dt \\ d\tilde{b}(t) = -r_b\tilde{b}(t)dt + K_b(y(t) - z(t))dt \end{cases} \tag{4.85}$$

and

$$y(t)dt = z(t)dt + \sqrt{\sigma_m}dw_2, \tag{4.86}$$

where all the variables are defined as before. As in the last section, we still replace γJ and $\gamma J'$ by J and J', respectively. Similarly, a reduced three-dimensional model can be obtained as:

$$d\theta = G\theta dt + \beta \begin{bmatrix} dw_1 \\ dw_2 \end{bmatrix}. \tag{4.87}$$

The matrices are defined as:

$$\theta = \begin{bmatrix} \Delta z \\ \Delta b \\ b \end{bmatrix}, \quad G = \begin{bmatrix} -K_z & J' & J - J' \\ -K_b & -r_b & 0 \\ 0 & 0 & -r_b \end{bmatrix}, \\ \beta = \begin{bmatrix} 0 & -K_z\sqrt{\sigma_m} \\ \sqrt{\sigma_{bf}} & -K - b\sqrt{\sigma_m} \\ \sqrt{\sigma_{bf}} & 0 \end{bmatrix}. \tag{4.88}$$

We also define

$$\Theta = E[\theta\theta^T] = \begin{bmatrix} \theta_{11} & \theta_{12} & \theta_{13} \\ \theta_{12} & \theta_{22} & \theta_{23} \\ \theta_{13} & \theta_{23} & \theta_{33} \end{bmatrix},$$

and this leads to the algebraic Riccati equation:

$$G\Theta + \Theta G^T + \beta\beta^T = 0. \tag{4.89}$$

It is linear and can be solved. Let us ignore the tedious derivation and just list the results. First note that K_z and K_b are different from the closed loop case:

$$K_z \approx \sqrt{2J'}\sigma_{bf}^{1/4}\sigma_m^{-1/4}, \\ K_b \approx \frac{K_z^2}{2J'} = \sigma_{bf}^{1/2}\sigma_m^{-1/2}. \tag{4.90}$$

The solution of (4.89) gives us the steady performance of the estimator,

$$\begin{aligned} \sigma_z &= E[(z - \tilde{z})^2] = \theta_{11} \\ &= \frac{1}{4}\frac{4\sigma_{bf}J^2 + K_z^2(3K_z^2 + 8K_z + 4r_b^2)\sigma_m}{K_z(K_z + r_b)(K_z + 2r_b)}, \end{aligned} \tag{4.91}$$

and

$$\begin{aligned} \sigma_b &= E[(b - \tilde{b})^2] = \theta_{22} \\ &= \frac{1}{8}\frac{K_z^4\sigma_m}{(K_z + r_b)J'^2} + \frac{1}{2}\sigma_{bf}\left(\frac{K_z(K_z(J - J') - r_bJ')^2}{(K_z + r_b)(K_z + 2r_b)^2 r_b J'^2} + \frac{3K_z + 4r_b}{(K_z + 2r_b)^2}\right). \end{aligned} \tag{4.92}$$

We have used the fact that $K_b = \frac{1}{2J'}K_z^2$.

The large J assumption still implies that $K_z \gg r_b$, and this leads to an approximation of σ_z:

$$\begin{aligned}\sigma_z &= \frac{1}{4}\frac{4\sigma_m K_z^2(K_z+r_b)^2 - \sigma_m K_z^3 + 4\sigma_{bf}^2 J^2}{K_z(K_z+2r_b)(K_z+r_b)} \\ &\approx \frac{3}{4}K_z\sigma_m + \sigma_{bf}\frac{J^2}{K_z^3} \\ &= \sigma_m^{3/4}\sigma_{bf}^{1/4}(\frac{3}{4}\sqrt{2J'} + \frac{J^2}{2J'\sqrt{2J'}}).\end{aligned} \tag{4.93}$$

Theorem E.1. *The variance σ_z in (4.93) reaches its minimum when $J = J'$.*

Proof. Let $v = \sqrt{2J'}$ and define $g(v) = \frac{3}{4}v + \frac{J^2}{v^3}$. The first two derivatives of g can be computed as

$$\begin{aligned}g'(x) &= \frac{3}{4} - \frac{3J^2}{v^4}, \\ g''(x) &= \frac{12J^2}{v^5} > 0.\end{aligned} \tag{4.94}$$

Equation $g'(v) = 0$ has two solutions: $v = \sqrt{2J}$ and $v = -\sqrt{2J}$. The last one can be discarded since $v > 0$ according to it definition. Thus f obtains its minimum when $v = \sqrt{2J}$ since $g''(v)$ is always positive. Identity $v = \sqrt{2J}$ implies $J = J'$ and $\sigma_z = \sqrt{2J}\sigma_m^{3/4}\sigma_{bf}^{1/4}$. □

Approximation of σ_b can also benefit from the large J assumption:

$$\sigma_b \approx \frac{K_z^3}{8J'^2}\sigma_m + \frac{1}{2}\sigma_{bf}(\frac{3}{K_z} + \frac{(K_z(J-J') - r_bJ')^2}{K_z^2J'^2r_b}). \tag{4.95}$$

By substituting K_z in terms of σ_{bf} and σ_m, we can further write σ_b as

$$\sigma_b \approx \sqrt{\frac{2}{J'}}\sigma_{bf}^{3/4}\sigma_m^{1/4} + \frac{\sigma_{bf}}{2r_b}(\frac{1-f}{f})^2 - \frac{\sigma_{bf}^{3/4}\sigma_m^{1/4}}{\sqrt{2J'}}(\frac{1-f}{f}) + \frac{\sigma_{bf}r_b}{2K_z^2}.$$

The last term can be discarded as $K_z \gg r_b$, and the above expression becomes

$$\sigma_b \approx \frac{\sigma_{bf}^{3/4}\sigma_m^{1/4}}{\sqrt{2J}}(\frac{3}{\sqrt{f}} - \frac{1}{f^{3/2}}) + \frac{\sigma_{bf}}{2r_b}(\frac{1-f}{f})^2. \tag{4.96}$$

Theorem E.2. *When J' is used in the design of the estimator (Kalman filter) and the control loop is open, the steady estimation error variance σ_b obtained for the linear system (4.26) is approximated under the large J assumption by*

$$\sigma_b = \frac{1}{\sqrt{2J}}\sigma_{bf}^{3/4}\sigma_m^{1/4}\left(\frac{3}{\sqrt{f}} - \frac{1}{f^{3/2}}\right) + \frac{\sigma_{bf}}{2r_b}\left(\frac{1-f}{f}\right)^2, \tag{4.97}$$

where $f = J'/J$. Specifically, when $J = J'$,

$$\begin{aligned} \sigma_z &= \sqrt{2J}\sigma_{bf}^{1/4}\sigma_m^{3/4}, \\ \sigma_b &= \sqrt{\frac{2}{J}}\sigma_{bf}^{3/4}\sigma_m^{1/4}. \end{aligned} \tag{4.98}$$

Proof. We have already shown most of the theorem. For the specific case of σ_b when $J = J'$, replacing f by 1 in (4.96), we obtain

$$\sigma_b = \sqrt{\frac{2}{J}}\sigma_{bf}^{3/4}\sigma_m^{1/4}.$$

□

We have several remarks on the above theorem:

- Note that we use $f = J'/J$, which is different from what the Caltech group used in their paper.
- The result is an extension of the Caltech group result and more accurate when f is near 1.
- Because of the large J assumption, $r_b^2 \ll J\sqrt{\sigma_{bf}/\sigma_m}$, the second term in (4.97) quickly dominates the first one when f is away from 1, which is the result of the Caltech group. It vanishes only when $f = 1$, so σ_b obtains its minimum near $f = 1$.

F. Effect of the input noise

The Caltech group has shown that feedback increases the robustness of magnetometry when the spin number is unknown. It is true when the input is noise free or the noise is very weak. When the control variable, the compensate field $u(t)$, is not exact, additional uncertainty is brought into the system. It will influence and finally damage the measurement. That leaves us the questions of how large its effect is and how strong a noise the system can tolerate.

An example is useful and can show us some idea. Fig. 4.2 through Fig. 4.5 show σ_b curves with respect to f for different σ_{uf}. Numbers are chosen as below:

$$\begin{cases} \sigma_m = \sigma_{bf} = 10^{-4} \\ J = 10^{10} \\ \lambda = 0.5 \\ r_b = 10^3 \end{cases} \tag{4.99}$$

Fig. 4.6 shows the 3D mesh of σ_b with respect to f and σ_{uf}. For most σ_{uf}, σ_b obtains its minimum near $f = 1$. But when σ_{uf} is too large, for example, $\sigma_{uf} = 10^{10}$, σ_b increases with $f = J'/J$.

A comparison between the open loop and closed loop results is shown in Fig.4.7 through Fig.4.12 by drawing the ratio σ_b/σ_{bo} where σ_{bo} is the estimation variance from Theorem E.2 when no feedback is applied. When σ_{uf} is small, $\sigma_b/\sigma_{bo} \leq 1$ for most J'. When σ_{uf} increases, σ_b/σ_{bo} first becomes larger than 1 near $f = 1$ and then the range of J' where $\sigma_b/\sigma_{bo} > 1$ expends to the whole scope of J'. Note that the figures are not accurate near $f = 1$ because of simplifications in the computation and round-off errors.

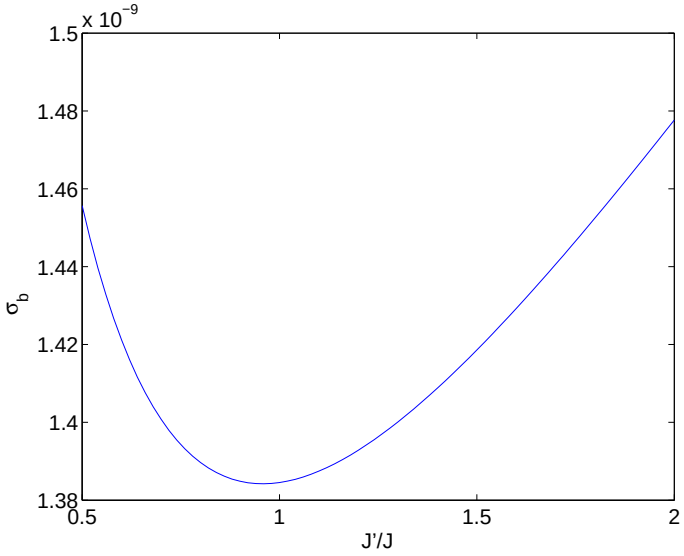


Fig. 4.2. Curve of σ_b with respect to f, when $\sigma_{uf} = 10^{-4}$. Parameters are chosen as (4.99).

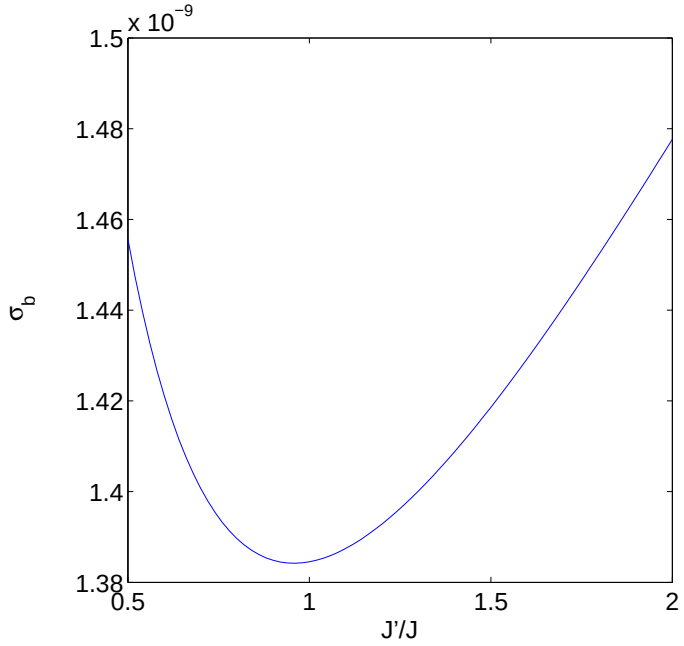


Fig. 4.3. Curve of σ_b with respect to f, when $\sigma_{uf} = 1.0$. Parameters are chosen as (4.99).

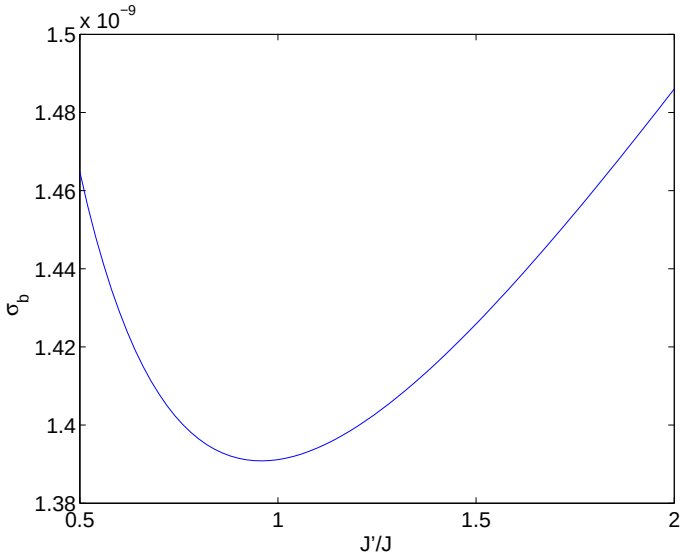


Fig. 4.4. Curve of σ_b with respect to f, when $\sigma_{uf} = 2.0 \times 10^4$. Parameters are chosen as (4.99).

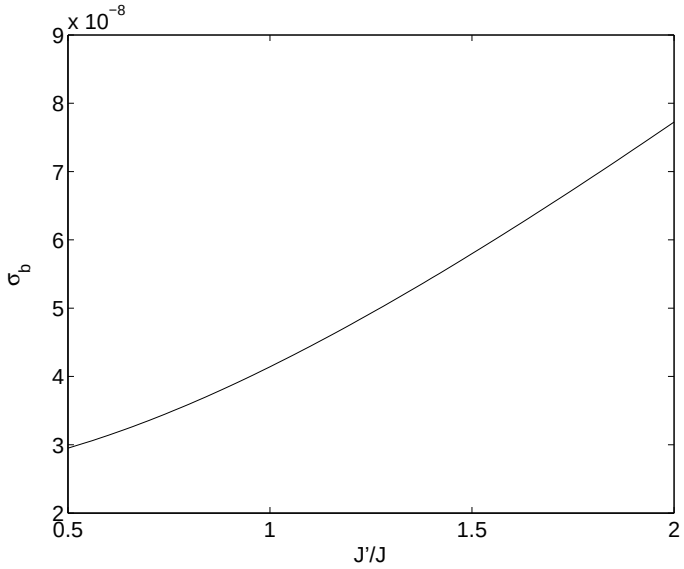


Fig. 4.5. Curve of σ_b with respect to f, when $\sigma_{uf} = 10^{10}$. Parameters are chosen as (4.99).

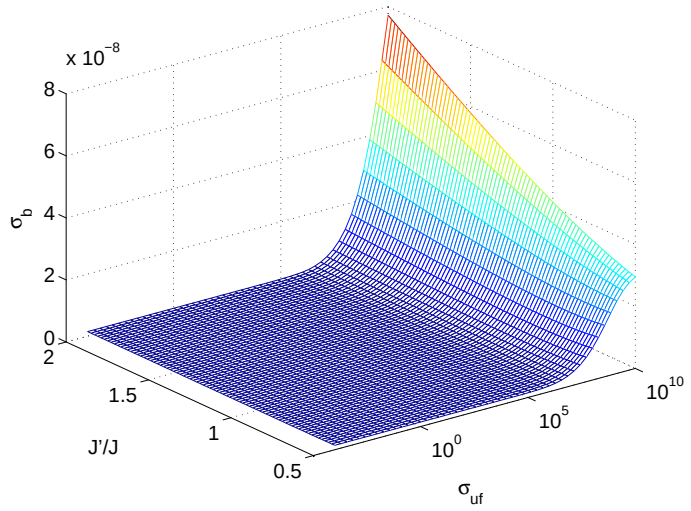


Fig. 4.6. 3D mesh of σ_b with respect to f and σ_{uf}. Parameters are chosen as (4.99).

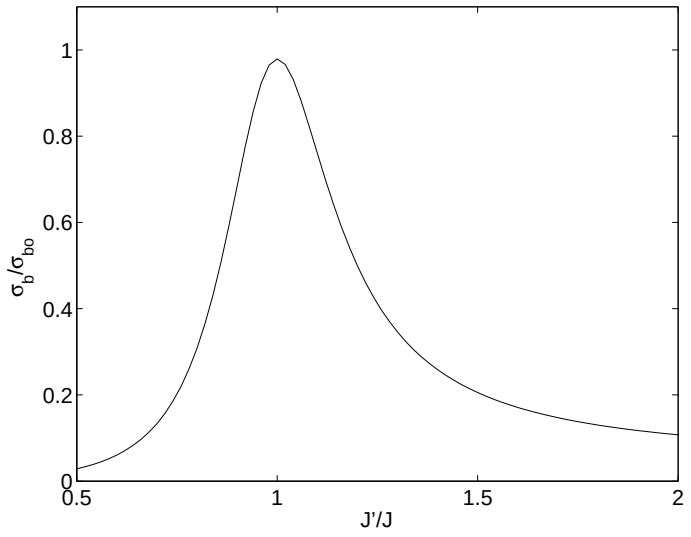


Fig. 4.7. Curve of σ_b/σ_{bo} with respect to f, when $\sigma_{uf} = 10^{-4}$. Parameters are chosen as (4.99).

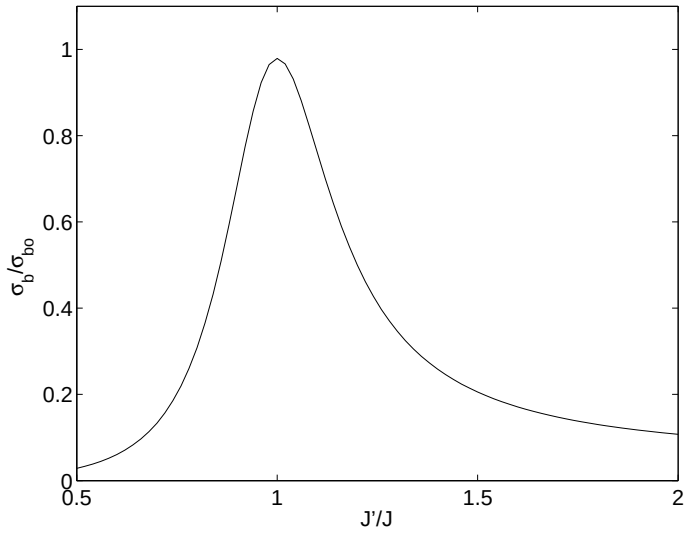


Fig. 4.8. Curve of σ_b/σ_{bo} with respect to f, when $\sigma_{uf} = 1.0$. Parameters are chosen as (4.99).

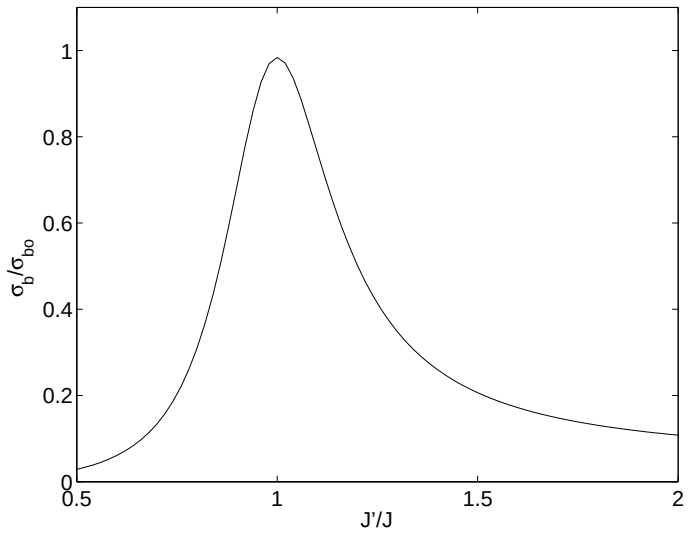


Fig. 4.9. Curve of σ_b/σ_{bo} with respect to f, when $\sigma_{uf} = 2.0 \times 10^4$. Parameters are chosen as (4.99).

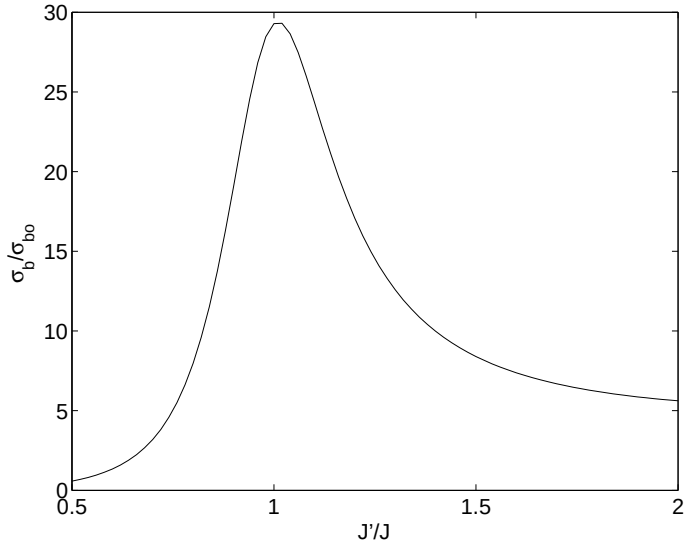


Fig. 4.10. Curve of σ_b/σ_{bo} with respect to f, when $\sigma_{uf} = 10.0^{10}$. Parameters are chosen as (4.99).

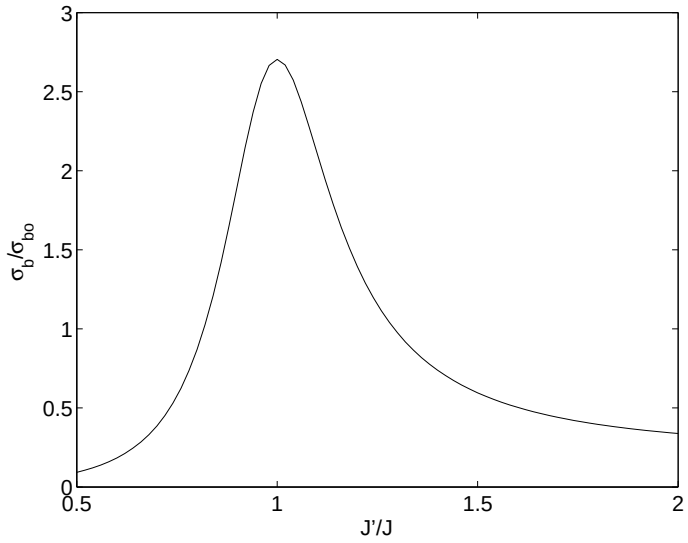


Fig. 4.11. Curve of σ_b/σ_{bo} with respect to f, when $\sigma_{uf} = 1.4 \times 10^7$. Parameters are chosen as (4.99).

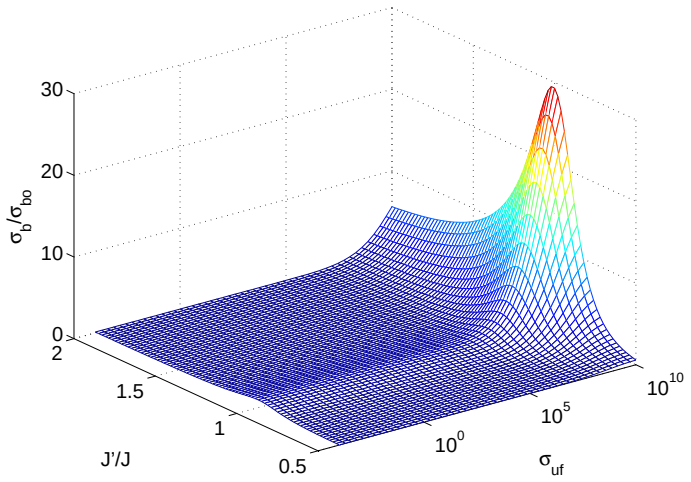


Fig. 4.12. 3D mesh of σ_b/σ_{bo} with respect to f and σ_{uf}. Parameters are chosen as (4.99).

Now let us come back to the formulae in Theorem D.5. If we can assume that

$$\frac{\sigma_{uf}}{\sigma_m} \ll \sqrt{T}, \tag{4.100}$$

K_z and K_b have a simpler form:

$$\begin{aligned} K_z &\approx \sqrt{2\sqrt{T}}, \\ K_b &= \frac{1}{2J'}(K_z^2 - \frac{\sigma_{uf}}{\sigma_m}) \\ &\approx \frac{1}{J'}\sqrt{T}. \end{aligned} \tag{4.101}$$

Substitute (4.101) back to (4.84), and we find an approximation of σ_b:

$$\begin{aligned} \frac{\sigma_b}{\sigma_m} &\approx \frac{K_z K_b}{J'} + \frac{K_b^2}{K_z}(f-1) \\ &\approx \frac{\sqrt{2}T^{3/4}}{J'^2} + \frac{T^{3/4}}{J'^2\sqrt{2}}(f-1) \\ &= (\frac{\sigma_{bf}}{\sigma_m} + \frac{r_b^2\sigma_{uf}}{J'^2\sigma_m})^{3/4}\frac{\sqrt{2}}{2}(\frac{1}{\sqrt{J'}} + \frac{\sqrt{J'}}{J}). \end{aligned} \tag{4.102}$$

The last term in (4.84) is discarded because $\frac{\sigma_{uf}}{\sigma_m} \ll \sqrt{T}$.

The above approximation explains the change of σ_b curves with part of the σ_{uf}.

When $\frac{r_b^2\sigma_{uf}}{J'^2\sigma_m} \ll \frac{\sigma_{bf}}{\sigma_m}$, or $\sigma_{uf} \ll \frac{J'^2\sigma_{bf}}{r_b^2}$, the σ_{uf} term can be further discarded and

$$\frac{\sigma_b}{\sigma_m} \approx (\frac{\sigma_{bf}}{\sigma_m})^{3/4}\frac{\sqrt{2}}{2}(\frac{1}{\sqrt{J'}} + \frac{\sqrt{J'}}{J}), \tag{4.103}$$

which is the same as the noise-free case. Thus σ_b obtains its minimum near $J = J'$. In our example, $\frac{J'^2\sigma_{bf}}{r_b^2} = 10^{10}$ and $\sigma_m\sqrt{J'^2\sigma_{bf}\sigma_m^{-1}} = 10^6$, which is a good approximation of $\sigma_m\sqrt{T}$. Thus when $\sigma_{uf} \ll 10^6$, the input will not have much effect on the measurement. This can be seen from in Fig. 4.6 and 4.12, where 10^5 is a point below which σ_{uf} have no large effect.

Theorem F.1. *Under the assumption of theorem D.5, and if σ_{uf} is not too large such that (4.102) holds, σ_b is larger than σ_{bo} in a range enclosing point $f = 1$, and bounded by two zero points of the function*

$$g(f) = \alpha(f^2 + \beta)^{3/4}\frac{1+f}{(1-f)^2} - 1, \tag{4.104}$$

where

$$\begin{aligned} \alpha &= \frac{2r_b}{\sqrt{2J}}(\frac{\sigma_m}{\sigma_{bf}})^{1/4}, \\ \beta &= \frac{r_b^2\sigma_{uf}}{J^2\sigma_{bf}}. \end{aligned} \tag{4.105}$$

If the above equation has no zero point, $\sigma_b > \sigma_{bo}$ everywhere. If it has only one zero point, $\sigma_b > \sigma_{bo}$ on the same side of the zero point as the point 1. In the above statements, we constrain f in the interval $(0, 6]$.

Proof. We only need to show that $\sigma_b \geq \sigma_{bo}$ at $f = 1$, and σ_b/σ_{bo} increases on the left side of 1 and decreases on the right side of 1.

When $J = J'$, from Theorem E.2 and D.5, we find

$$\sigma_b = \sigma_m^{1/4}(\sigma_{bf} + \frac{r_b^2}{J'^2}\sigma_{uf})^{3/4}\sqrt{\frac{2}{J}}, \tag{4.106}$$

and

$$\sigma_{bo} = \sigma_m^{1/4}\sigma_{bf}^{3/4}\sqrt{\frac{2}{J}}. \tag{4.107}$$

Clearly, $\sigma_b \geq \sigma_{bo}$ when $J = J'$, and the closed loop control is necessary only because it keeps the small angle assumption. Its behavior is not better than the that of the open loop when we know J exactly. It is also interesting to notice that the same result is obtained when σ_{bf} is amplified to $\sigma_{bf} + \frac{r_b^2}{J'^2}\sigma_{uf}$.

When J' is away from J, according to our previous discussion, the second term in (4.97) dominates quickly and the ratio σ_b/σ_{bo} can be well approximated by

$$\begin{aligned}
\frac{\sigma_b}{\sigma_{bo}} &= \frac{(1+\frac{r_b^2\sigma_{uf}}{J'^2\sigma_{bf}})^{3/4}\frac{1}{\sqrt{2J}}(\frac{\sigma_m}{\sigma_{bf}})^{1/4}(\frac{1}{\sqrt{f}}+\sqrt{f})}{\frac{1}{2r_b}(\frac{1-f}{f})^2} \\
&= \frac{2r_b}{\sqrt{2J}}(\frac{\sigma_m}{\sigma_{bf}})^{1/4}(f^2+\frac{r_b^2\sigma_{uf}}{J^2\sigma_{bf}})^{3/4}\frac{1+f}{(1-f)^2} \\
&= \alpha(f^2+\beta)^{3/4}\frac{1+f}{(1-f)^2},
\end{aligned} \tag{4.108}$$

where

$$\begin{aligned}
\alpha &= \frac{2r_b}{\sqrt{2J}}(\frac{\sigma_m}{\sigma_{bf}})^{1/4}, \\
\beta &= \frac{r_b^2\sigma_{uf}}{J^2\sigma_{bf}}.
\end{aligned}$$

Define $h(f) = \alpha(f^2+\beta)^{3/4}\frac{1+f}{(1-f)^2}$, and we compute its first order derivative:

$$\frac{d}{df}h = \alpha(f^2+\beta)^{-1/4}\frac{1}{(1-f)^3}(3\beta + f(\frac{3}{2}+\beta) + f^2(3-\frac{f}{2})).$$

When $f \leq 6$, the derivative is positive on the left of 1 and negative on the right of 1 and that is what we want. □

We conclude from the above theorem:

- When $J' = J$, the closed loop behaves not better than the open loop. But if σ_{uf} is small, this make little difference.

- The range of f where the closed loop behaves better depends on σ_{uf}, and that is defined by the two zero points of the function $g(f)$.

- We have found two criteria to check if σ_{uf} can be ignored: $\sigma_{uf} \ll \sigma_m\sqrt{T}$ and $\sigma_{uf} \ll \dfrac{J'^2\sigma_b}{r_b^2}$.

G. Robust and optimal choice of J'

In this section, besides giving an explanation of the robust choice of J' by the Caltech group, we will further dicuss its optimal choice when J is treated as a random number. We assume that σ_{uf} is not too large so that we can use (4.102).

When σ_{uf} is small enough so that it can be neglected, let $J = aJ_0$ and $J' = bJ_0$, $0.5 \le a \le 2.0$ where J_0 is some constant, we can rewrite (4.46) as:

$$\sigma_b\sigma_m^{-1} = c(\frac{1}{\sqrt{b}} + \frac{\sqrt{b}}{a}), \tag{4.109}$$

where $c = (\dfrac{\sigma_{bf}}{\sigma_m})^{3/4} \cdot \dfrac{\sqrt{2}}{2} \cdot \dfrac{1}{\sqrt{J_0}}$. Denote the minimum of a as $\min(a)$, we can find the worst result for a fixed b:

$$\sigma_b\sigma_m^{-1} \le c(\frac{1}{\sqrt{b}} + \frac{\sqrt{b}}{\min(a)}). \tag{4.110}$$

The minimum of the right side is obtained when $b = \min(a)$ or $J' = \min(a) \cdot J_0$, corresponding to the result by the Caltech group. Thus the Caltech design is robust in the maximum-minimum meaning.

An optimal design can be found by choosing the optimal objective function as

$$\mathcal{J} = E[\sigma_b],$$

where the expectation is taken with respect to the random variable J. From (4.109),

we have

$$E[\sigma_b] = \sigma_m c(\frac{1}{\sqrt{b}} + \sqrt{b}E[\frac{1}{a}]) \geq 2c\sigma_m\sqrt{E[\frac{1}{a}]}, \tag{4.111}$$

and the identity is only obtained when $J' = \dfrac{1}{E[1/J]}$.

When σ_{uf} can not be ignored, $c = (\frac{\sigma_{bf}}{\sigma_m} + \frac{r_b^2\sigma_{uf}}{\sigma_m J'^2})^{3/4} \cdot \dfrac{\sqrt{2}}{2} \cdot \dfrac{1}{\sqrt{J_0}}$ and the optimal choice drifts away from what we obtained above.

CHAPTER V

CONCLUSIONS

At present, several types of elementary quantum computing devices have been developed, based on AMO (atomic, molecular and optical) or semiconductor physics and technologies. We may roughly classify them into the following:

atomic — ion and atom traps, cavity QED;

molecular — NMR;

semiconductor — coupled quantum dots, silicon (Kane);

crystal structure — nitrogen-vacancy (NV) diamond;

superconductivity — SQUID.

The above classification is not totally rigorous as new types of devices, such as quantum dots, or ion traps imbedded in cavity-QED, have emerged which are of the *hybrid* nature. Also, laser pulse control, which is of the optical nature, seems to be omnipresent.

We still do not know which proposal will win the race to build a reliable and useful quantum computer. Liquid NMR is so far the most successful and the only technology to realize a seven qubit algorithm in laboratory. It forms a test bed for quantum algorithms and the techniques developed with it can be easily used in the control of other quantum devices. The severe obstacle to its future applicability is its lack of scalability. Solid state NMR may solve the problem, but more improvements are still needed. Superconducting quantum computer utilizes the quantum effects in superconducting. The devices are composed of Josephson junction and CPB (Cooper Pair Box). Coherence oscillations have been observed and two qubit CNOT gate has

been realized. Technologies used for the superconducting quantum devices are those well developed in semiconductor industry, but to manufacture a superconducting quantum computer with several qubits still has a long way to go.

In Chapter IV, we move to an example of "quantum metrology". We have investigated an extended model of a magnetometer to evaluate the effect of input noise on the measurement, which is originally ignored by earlier researches. Our computation shows the benchmark under which the input noise has a negligible effect. Besides keeping the small angle assumption, the closed-loop feedback makes the measurement robust to an unknown parameter. If we only look into the steady performances of the closed-loop and open-loop cases, the input noise decides when the closed-loop is better than the open-loop. When the input noise is small, the closed-loop always behaves better. But when the input noise is large, there is a range including $f = 1$ where the closed-loop behaves worse. The open-loop estimation is preferred when we have a good estimation of the parameter J and the spins do not rotate too much. A simulation is given to support our results. The input noise also influences the robust and optimal choices of J'. The optimal choice drifts away when the noise increases.

REFERENCES

[1] D. Deutsch and R. Jozsa, Proc. R. Soc. London. A **439**, 553 (1992).

[2] A. Ekert and R. Jozsa, Rev. Mod. Phys. **68**, 733 (1996).

[3] L.G. Grover, Phys. Rev. Lett. **79**, 325 (1997).

[4] P.W. Shor, SIAM J. Comput. **26**, 1484 (1997).

[5] L.M.K. Vandersypen, M. Steffen, G. Breyta, C. Yannon, M. Sherwood, et al., Nature **414**, 883 (2001).

[6] R.L. Rivest, A. Shamir, and L. Adleman, Commun. of the ACM **21**, 120 (1978).

[7] S.J. Lomonaco, e-print quant-ph/0010034.

[8] C. Ahn, A.C. Doherty, and A.J. Landahl, Phys. Rev. A **65**, 042301 (2002).

[9] A.J. Berglund and H. Mabuchi, Appl. Phys. B **78**, 653 (2004).

[10] A.C. Doherty, S. Habib, K. Jacobs, H. Mabuchi and S. Tan, Phys. Rev. A **62**, 012105 (2000).

[11] J.M. Geremia, J.K. Stockton, and H. Mabuchi, Science **304**, 270 (2004).

[12] H. Mabuchi and N. Khaneja, Int. J. Robust Nonlinear Control **15**, 647 (2005).

[13] L.K. Thomsen, S. Mancini, and H.W. Wiseman, Phys. Rev. A **65**, 061801 (2002).

[14] H.W. Wiseman, Phys. Rev. A **49**, 2133 (1994).

[15] A.C. Doherty and K. Jacobs, Phys. Rev. A **60**, 2700 (1999).

[16] J.K. Stockton, J.M. Geremia, A.C. Doherty, and H. Mabuchi, Phys. Rev. A **69**, 032109 (2004).

[17] F. Verstraete, A.C. Doherty, and H. Mabuchi, Phys. Rev. A **64**, 032111 (2001).

[18] Z. Zhang, G. Chen, Z. Diao, and P. Hemmer, in *Advances in Mechanics and Mathematics (AMMA)* edited by Gao and R.W. Ogden (Springer-Verlag, New York, 2006), Vol. 3, p. 135.

[19] P. Benioff, J. Stat. Phys. **22**, 563 (1980).

[20] R.P. Feynman, Int. J. Theor. Phys. **21**, 467 (1982).

[21] E. Bernstein and U. Vazirani, SIAM J. Comput. **26**, 1411 (1997).

[22] D. Simon, SIAM J. Comput., **26**, 1474 (1997).

[23] D. Deutsch, Proc. R. Soc. London A **425**, 73 (1989).

[24] C.C. Yao, in *Proceedings of the 34th Annual IEEE Symposium on Foundations of Computer Science*, Los Alamitos, CA, (1993) (IEEE Press, 1993), p. 352.

[25] V. Vedral, A. Barenco, and A. Ekert, Phys. Rev. A **54**, 147 (1996).

[26] A. Barenco, C.H. Bennett, R. Cleve, D.P. DiVincenzo, N. Margolus, et al., e-print quant-ph/9503016.

[27] D.P. DiVincenzo, Phys. Rev. A **51**, 1015 (1995).

[28] G.D. Mateescu and A. Valeriu, *2D NMR : Density Matrix and Product Operator Treatment* (PTR Prentice-Hall, Englewood Cliffs, NJ, 1993).

[29] R. Laflamme, E. Knill, C. Negrevergne, R. Martinez, S. Sinha, et al, in *Experimental Quantum Computation and Information, Proceeding of the International*

School of Physics "Enrico Fermi", edited by F. De Martini and C. Monroe (IOS Press, Amsterdam, Netherlands, 2002), p. 411.

[30] J. Normand, *A Lie Group: Rotations in Quantum Mechanics* (North-Holland, New York, 1980).

[31] J.W. Emsley, J. Feeney and L.H. Sutcliffe, *High Resolution Nuclear Magnetic Resonance Spectroscopy* (Pergamon Press, Oxford, 1965).

[32] C.P. Poole, and H.A. Farach, *Theory of Magnetic Resonance* (Wiley, New York, 1987).

[33] M.L. Martin, and G.J. Martin, *Practical NMR Spectroscopy* (Heyden, London, U.K., 1980).

[34] J.C. Edwards, Principles of NMR, *http://www.process-nmr.com/nmr.htm*, accessed June 23, 2006.

[35] J.P. Hornak, The basics of NMR, *http://www.cis.rit.edu/htbooks/nmr/inside.htm*, accessed June 23, 2006.

[36] Wikipedia, Nuclear magnetic resonance, *http://en.wikipedia.org/wiki/Nuclear_magnetic_resonance*, accessed June 23, 2006.

[37] F. Bloch, Phys. Rev. **70**, 460 (1946).

[38] M.A. Nielsen and I.L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, U.K., 2000).

[39] I.L. Chuang, in *Introduction to Quantum Computation and Information*, edited by H. Lo, S. Popescu, and T. Spiller (World Scientific, Singapore, 1998), p. 311.

[40] N. Linden, B. Herve, R.J. Carbajo and R. Freeman, Chem. Phys. Lett. **305**, 28 (1999).

[41] J.A. Jones, and E. Knill, J. of Magnetic Resonance **141**, 322 (1999).

[42] D.W. Leung, I.L. Chuang, F. Yamaguchi, and Y. Yamamoto, Phys. Rev. A **61**, 042310 (2000).

[43] S.J. Glaser, T. Schulte-Herbruggen, M. Sieveking, O. Schedletzky, N.C. Nielsen, et al., Science **280**, 421 (1998).

[44] N. Khaneja, R. Brockett, and S.J. Glaser, Phys. Rev. A **63**, 032308 (2001).

[45] N. Khaneja, S.J. Glaser, and R. Brockett, Phys. Rev. A **65**, 032301 (2002).

[46] O.W. Sorenson, Progress of Nuclear Magnetic Resonance Spectroscopy **21**, 503 (1989).

[47] H.K. Cummings, and J.A. Jones, New Journal of Physics **2**, 6.1 (2000).

[48] E.M. Fortunato, M.A. Pravia, N. Boulant, G. Taklemariam, T. Havel, et al., J. of Chem. Phys. **116**, 7599 (2002).

[49] M.H. Levitt, Progress in Nuclear Magnetic Resonance Spectroscopy **18**, 61 (1986).

[50] J.A. Jones, Phil. Trans. Roy. Soc. Lond. A **361**, 1429 (2003).

[51] J.A. Jones, Phys. Rev. A **67**, 012317 (2003).

[52] T. Yamamoto, Y.A. Pashkin, O. Astaflev, Y. Nakamura, and J.S. Tsai, Nature **425**, 941 (2003).

[53] L.M.K. Vandersypen, M. Steffen, M.H. Sherwood, C. Yannoni, G. Breyta, and I.L. Chuang, Appl. Phys. Lett. **76**, 646 (2000).

[54] S.L. Patt, J. of Magnetic Resonance **96**, 94 (1992).

[55] R. Freeman, J. of Progress in Nuclear Magnetic Resonance Spectroscopy **32**, 59 (1998).

[56] E. Kupce, and R. Freeman, J. of Magnetic Resonance Series A **112**, 261 (1995).

[57] E. Knill, I. Chuang, and R. Laflamme, Phys. Rev. A **57**, 3348 (1998).

[58] D.G. Cory, A.F. Fahmy, and T.F. Havel, Proc. Natl. Acad. Sci. USA **94**, 1634 (1997).

[59] B. Fung, and V.L. Ermakov, J. of Chem. Phys. **121**, 8410 (2004).

[60] N. Gershenfeld, and I.L. Chuang, Science **275**, 350 (1997).

[61] L.M.K. Vandersypen, C.S. Yannoni, M.H. Sherwood, and I.L. Chuang, Phys. Rev. Lett. **83**, 3085 (1999).

[62] A.K. Khitrin, H. Sun, and B.M. Fung, Phys. Rev. A **63**, 020301 (2001).

[63] A.M. Childs, I.L. Chuang, and D.W. Leung, Phys. Rev. A **64**, 012314 (2001).

[64] I.L. Chuang, L.M.K. Vandersypen, X. Zhou, D.W. Leung, and S. Lloyd, Nature **393**, 143 (1998).

[65] A.O. Pittenger, *An Introduction to Quantum Computing Algorithms* (Birkhauser, Boston, 2000).

[66] J. Yepez, International Journal of Modern Physics C **12**, 1273 (2001).

[67] G.P. Berman, A.A. Ezhov, D.I. Kamenev, and J. Yepez, Phys. Rev. A **66**, 012310 (2002).

[68] M.A. Pravia, Z. Chen, J. Yepez, and D.G. Cory, Comput. Phys. Commun. **146**, 339 (2002).

[69] J. Yepez, International Journal of Modern Physics C **12**, 1285 (2001).

[70] J. Yepez, and B. Boghosian, Comput. Phys. Commun. **146**, 280 (2002).

[71] L. Vahala, G. Vahala, and J. Yepez, Phys. Lett. A **306**, 227 (2003).

[72] G. Vahala, J. Yepez, and L. Vahala, Phys. Lett. A **310**, 187 (2003).

[73] C. Cosmelli, P. Carelli, M.G. Castellano, F. Chiarello, Palazzi G. Diambrini, et al., Phys. Rev. Lett. **82**, 5357 (1999).

[74] J.R. Friedman, V. Patel, W. Chen, S.K. Tolpygo, and J.E. Lukens, Nature **406**, 43 (2000).

[75] J.M. Martinis, M.H. Devoret, and J. Clark, Phys. Rev. B **35** 4682 (1987).

[76] R. Rouse, S. Han, and J.S. Lukens, Phys. Rev. Lett. **75** 514 (1995).

[77] P. Silvestrini, V.G. Palmieri, B. Ruggiero, and M. Russo, Phys. Rev. Lett. **79**, 3046 (1997).

[78] R.F. Voss and R.A. Webb, Phys. Rev. Lett. **47**, 265 (1981).

[79] Y. Nakamura, Y.A. Pashkin, and J.S. Tsai, Nature **398**, 786 (1999).

[80] M.H. Devoret and J.M. Martinis, e-print arXiv: cond-mat/041174.

[81] M.H. Devoret, A. Wallraff, and J.M. Martinis, e-print arXiv: cond-mat/0411174.

[82] D. Esteve and D. Vion, in *Les Houches Summer School-Session LXXXI on Nanoscopic Quantum Physics, 2004* (Elsevier Science, Amsterdam, 2005).

[83] G. Wendin, Phil. Trans. R. Soc. Lond. A **361**, 1323 (2003).

[84] J. Bardeen, L.N. Cooper and J.R. Schrieffer, Phys. Rev. **108**, 1175 (1957).

[85] J.B. Ketterson and S.N. Song, *Superconductivity* (Cambridge University Press, Cambridge, U.K., 1999).

[86] C.P. Poole, H.A. Farach, and R.J. Creswick, *Superconductivity* (Academic Press, New York, 1996).

[87] A.C. Rose-Innes and E.H. Rhoderick, *Introduction to Superconductivity*, Second Edition (Pergamon Press, Oxford, 1978).

[88] M. Tinkham, *Introduction to Superconductivity* (McGraw-Hill, New York, 1996).

[89] Oak Ridge National Laboratory, Historical background, *http://www.ornl.gov/info/reports/m/ornlm3063r1/pt2.html*, accessed June 23, 2006.

[90] G. Wendin and V.S. Shmeiko, e-print arXiv:cond-mat/0508729v1.

[91] J.Q. You and F. Nori, Phys. Today **Nov. 2005**, p. 42.

[92] S. Saito, M. Thorwart, H. Tanaka, M. Ueda, H. Nakano, et al., Phys. Rev. Lett. **93**, 037001 (2004).

[93] Caspar H. van der Wal, A.C.J. ter Haar, F.K. Wilhelm, R.N. Schouten, C.J.P.M. Harmans, et al., Science **290**, 773 (2000).

[94] J.Q. You and F. Nori, Phys. Rev. B **68**, 064509 (2003).

[95] E. Collin, G. Ithier, A. Aassime, P. Joyez, D. Vion, and D. Esteve, Phys. Rev. Lett. **93**, 157005 (2004).

[96] D. Vion, A. Aassime, A. Cottet, P. Joyez, H. Pothier, et al., Science **296**, 886 (2002).

[97] J.M. Martinis, S. Nam, and J. Aumentado, Phys. Rev. Lett. **89**, 117901 (2002).

[98] R.W. Simmonds, K.M. Lang, D.A. Hite, S. Nam, D.P. Pappas, and J.M. Martinis, Phys. Rev. Lett. **93**, 077003 (2004).

[99] Y. Yu, S. Han, X. Chu, S. Chu, and Z. Wang, Science **296**, 889 (2002).

[100] Y. Pashkin, T. Yamamoto, O. Astaflev, Y. Nakamura, D.V. Averin, and J.S. Tsai, Nature **421**, 823 (2003).

[101] A. Shnirman, G. Schön, and Z. Hermon, Phys. Rev. Lett. **79**, 2317 (1997).

[102] Y. Makhin, G. Schön, and A. Shnirman, Nature **398**, 305 (1999).

[103] J.Q. You, J.S. Tsai, and F. Nori, Phys. Rev. Lett. **89**, 197902 (2002).

[104] J.L. Brylinski and R. Brylinski, in *Mathematics of Quantum Computation*, edited by R. Brylinski and G. Chen (Chapman & Hall/CRC, Boca Raton, Florida, 2002), p. 101.

[105] Y. Nakamura, Y.A. Pashkin, and J.S. Tsai, Physica B **280**, 405 (2000).

[106] Y. Nakamura and J.S. Tsai, J. of Low Temperature Physics **118**, 765 (2000).

[107] O. Astafiev, Y.A. Pashkin, T. Yamamoto, Y. Nakamura, and J.S. Tsai, Phys. Rev. B **69**, 180507 (2004).

[108] A. Aassime, G. Johansson, G. Wendin, R.J. Schoelkopf, and P. Delsing, Phys. Rev. Lett. **86**, 3376 (2001).

[109] R.J. Schoelkof, P. Wahlgren, A.A. Kozhevnikov, P. Delsing, and D.E. Prober, Science **280**, 1238 (1998).

[110] L.B. Kish and P. Svedlindh, IEEE Trans. Electron Devices **41**, 2112 (1994).

[111] M. Grimble and M. Johnson, *Optimal Control and Stochastic Estimation: Theory and Application* (Jonh Wiley & Sons, New York, 1988), Vol. 1.

[112] G.C. Goodwin and K.S. Sin, *Adaptive Filtering, Prediction and Control* (Prentice-Hall, Inc., Englewood, NJ, 1984).

[113] B. Øksendal, *Stochastic Differential Equation: an Introduction with Applications* (Springer, Berlin, Germany, 2003).

[114] M. Scully and M.S. Zubairy, *Quantum Optics* (Cambridge University Press, Cambridge, U.K., 1997).

VITA

Name:	Zhigang Zhang
Address:	Department of Mathematics, TAMU,
	College Station, TX77843-3348
Email Address:	zgzhang@math.tamu.edu
Education:	B.S., Automation,
	University of Science and Technology of China, 1992
	M.S., Theory and Applications of Automatic Control
	University of Science and Technology of China, 1995
	Ph.D., Mathematics,
	Texas A&M University, August 2006