Quantum Computation and Quantum Information

Yazhen Wang

Abstract. Quantum computation and quantum information are of great current interest in computer science, mathematics, physical sciences and engineering. They will likely lead to a new wave of technological innovations in communication, computation and cryptography. As the theory of quantum physics is fundamentally stochastic, randomness and uncertainty are deeply rooted in quantum computation, quantum simulation and quantum information. Consequently quantum algorithms are random in nature, and quantum simulation utilizes Monte Carlo techniques extensively. Thus statistics can play an important role in quantum computation and quantum simulation, which in turn offer great potential to revolutionize computational statistics. While only pseudo-random numbers can be generated by classical computers, quantum computers are able to produce genuine random numbers; quantum computers can exponentially or quadratically speed up median evaluation, Monte Carlo integration and Markov chain simulation. This paper gives a brief review on quantum computation, quantum simulation and quantum information. We introduce the basic concepts of quantum computation and quantum simulation and present quantum algorithms that are known to be much faster than the available classic algorithms. We provide a statistical framework for the analysis of quantum algorithms and quantum simulation.

Key words and phrases: Quantum algorithm, quantum bit (qubit), quantum Fourier transform, quantum information, quantum mechanics, quantum Monte Carlo, quantum probability, quantum simulation, quantum statistics.

1. INTRODUCTION

For decades computer hardware has grown in power approximately according to Moore's law, which states that the computer power doubles for constant cost roughly once every two years. However, because of the fundamental difficulties of size in conventional computer technology, this dream run is ending. The conventional approaches to the fabrication of computer technology are to make electronic devices smaller and smaller in order to increase the computer power. As the sizes of the electronic devices get close to the atomic scale, quantum effects are starting to interfere in their functioning, and thus the conventional

approaches run up against the size limit. One possible way to get around the difficulties is to move to a new computing paradigm provided by quantum information science. Quantum information science is based on the idea of using quantum devices to perform computation and manipulate and transmit information, instead of electronic devices following the laws of classical physics, see Deutsch (1985), DiVincenzo (1995), Feynman (1981/82). Quantum mechanics and information theory are two of the great scientific developments and technological revolutions in the 20th century, and quantum information science is to marry the two previously disparate fields and form a single unifying viewpoint. Quantum information science studies the preparation and control of the quantum states of physical systems for the purposes of information transmission and manipulation. It includes quantum computation, quantum communication and quantum cryptography. This revolutionary field will enable a range of

Yazhen Wang is Professor, Department of Statistics,
University of Wisconsin–Madison, Madison, Wisconsin
53706, USA (e-mail: yzwang@stat.wisc.edu).

exotic new devices to be possible. There is now a general agreement that quantum information science will likely lead to the creation of a quantum computer to solve problems that could not be efficiently solved on a classical computer. Already scientists have built rudimentary quantum computers in the research laboratory to run quantum algorithms and perform certain calculations. Intensive research efforts are under way around the world to investigate a number of technologies that could lead to more powerful and more prevalent quantum computers in the near future. It is believed that quantum information and quantum bits are to lead to a 21st century technological revolution much as classic information and classic bits did to the 20th century. Since the theory of quantum mechanics is fundamentally stochastic, randomness and uncertainty are deeply rooted in quantum computation and quantum information. As a result, quantum algorithms are of random nature in the sense that they yield correct solutions only with some probabilities, and Monte Carlo methods are widely employed in quantum simulation. Thus statistics has an important role to play in quantum computation, quantum simulation and quantum information. On the other hand, quantum computation and quantum simulation have tremendous potential to revolutionize computational statistics.

A quantum system is generally described by its state, and the state is mathematically defined to be a unit vector in some complex Hilbert space. The number of complex numbers required to characterize the quantum state usually grows exponentially with the size of the system, rather than linearly, as occurs in classical systems. As a consequence, it takes an exponential number of bits of memory on a classical computer to store the quantum state, which puts classical computers in a difficult position to simulate a quantum system. On the other hand, nature quantum systems are able to store and keep track of an exponential number of complex numbers and perform data manipulations and calculations as the systems evolve. Quantum information science is to grapple with understanding how to take advantage of the enormous information hidden in the quantum systems and to harness the immense potential computational power of atoms and molecules for the purpose of performing computation and processing information. Already it has been shown that quantum algorithms like Grover's search algorithm and Shor's factoring algorithm provide great advantage over known classical algorithms.

Contemporary scientific studies often rely on understanding complex quantum systems, such as those

in biochemistry and nanotechnology for the design of biomolecules and nano-materials. Quantum simulation is to use computers to simulate a quantum system and its time evolution. Classical computers are being used for quantum simulation in designing novel molecules and creating innovative nano-products. Quantum computers built upon quantum systems may excel in simulating naturally occurring quantum systems, while large quantum systems may be impossible to simulate in an efficient manner by classical computers. A quantum system with b distinct components may be described with b quantum bits in a quantum computer, while a classical computer requires 2^b bits of memory to store its quantum state. This advantage allows quantum computers to efficiently simulate general quantum systems that are not efficiently simulatable on classical computers.

In this article we review the concepts of quantum computation and introduce quantum algorithms and quantum simulation. The quantum algorithms are known to be much faster than the available classical algorithms. Statistical analyses of quantum algorithms and quantum simulation are provided. We give a brief description on quantum information. The article sections start with presentations in broad brushstrokes, followed by specific discussions along with some mathematical derivations if necessary. The intention is to give each topic first an overview and then a general description and a precise characterization. It is recommended to focus on the qualitative discussions but skip the derivations for the readers who would like to get a quick picture of quantum computation and quantum simulation.

The rest of the paper proceeds as follows. Section 2 briefly introduces quantum mechanics, quantum probability and quantum statistics. Section 3 reviews basic concepts of quantum computation and entanglement. Section 4 illustrates some widely known quantum algorithms and provides a statistical framework for the study of quantum algorithms. Section 5 presents quantum simulation and discusses its statistical analysis. Section 6 gives a short description on quantum information theory. Section 7 features concluding remarks and lists some open research problems.

2. BRIEF BACKGROUND REVIEW ON QUANTUM THEORY

Quantum mechanics has been applied to everything under and inside the Sun, from chemical reaction and superconductor to the structure of DNA and nuclear fusion in stars. Although the significant difference be-

tween classical physics and quantum physics lies in the quantum prediction of physical entity when the scale of observations becomes comparable to the atomic or sub-atomic scale, many macroscopic properties of systems can only be fully explained and understood by quantum physics. The quantum world is extremely strange, and quantum theory is completely counterintuitive. Light waves behave like particles and particles behave like waves (wave particle duality); matter can go from one spot to another without moving through the intermediate space (quantum tunneling); information can be moved across a vast distance without transmitting it through the intervening space (quantum teleportation). Quantum theory provides a mathematical description of wave particle duality and interaction of matter and energy. It describes the time evolutions of physical systems via wave functions. The wave functions encapsulate the probabilities that particles are to be found in a given state at a given time. For example, the probability of finding a photon in some region is the square of the modulus of a wave function, and, since at some point the sum of two wave functions can be zero but neither wave function is zero, probabilities appear to cancel out each other in a way totally unexpected from classical probability. The intrinsic stochastic nature of quantum theory indicates a deep connection between quantum mechanics and probability. Since the main focus of this paper is on quantum computation and quantum information, we give a brief description of quantum theory in this section to provide some quantum background for the purpose of reviewing quantum computation and quantum simulation in subsequent sections. For further reading on the subjects we recommend textbooks by Sakurai and Napolitano (2010) at the graduate level and Griffiths (2004) at the undergraduate level for quantum mechanics, Holevo (1982), Parthasarathy (1992) and Wang (1994) for quantum probability and quantum stochastic processes, and Artiles, Gill and Guţă (2005) and Barndorff-Nielsen, Gill and Jupp (2003) for quantum statistics.

2.1 Hilbert Space and Operator

For the sake of simplicity we choose to work with comparatively easy finite-dimensional situations. Denote by $\mathbb{C}$ the set of all complex numbers. We start with vector space in linear algebra. A simple example of vector space is $\mathbb{C}^k$ consisting of all k-tuples of complex numbers $(z_1, \ldots, z_k)$. The elements of a vector space are called vectors. As in quantum mechanics and quantum computation, we use Dirac notations $|\cdot\rangle$ (which is called ket) and $\langle\cdot|$ (which is called bra) to in-

dicate that the objects are column vectors or row vectors in the vector space, respectively. Denote by superscripts $*$, $\prime$ and $\dagger$ the conjugate of a complex number, the transpose of a vector or matrix, and conjugate transpose operation, respectively. We define an inner product on the vector space to be a function that takes as input two vectors from the vector space and produces a complex number as output. For $|u\rangle$ and $|v\rangle$ in the vector space, we denote their inner product by $\langle u|v\rangle$. The inner product must satisfy (i) conjugate symmetry, $\langle u|v\rangle = (\langle v|u\rangle)^*$; (ii) linearity in the second argument, $\langle u|v + w\rangle = \langle u|v\rangle + \langle u|w\rangle$; (iii) positive-definiteness, $\langle u|u\rangle \geq 0$ with equality only for $u = 0$. For example, $\mathbb{C}^k$ has a natural inner product

$$\langle u|v\rangle = \sum_{j=1}^k u_j^* v_j = (u_1^*, \ldots, u_k^*)(v_1, \ldots, v_k)',$$

where $\langle u| = (u_1, \ldots, u_k)$ and $|v\rangle = (v_1, \ldots, v_k)'$. An inner product induces a norm $\|u\| = \sqrt{\langle u|u\rangle}$, and a distance $\|u - v\|$ between $|u\rangle$ and $|v\rangle$. For the finite-dimensional case, a Hilbert space $\mathcal{H}$ is simply a vector space with an inner product.

An operator $\mathbf{A}$ on $\mathcal{H}$, denoted by $\mathbf{A}(|u\rangle)$ for $|u\rangle \in \mathcal{H}$, is a function mapping from $\mathcal{H}$ to $\mathcal{H}$ that satisfies $\mathbf{A}(a|u\rangle + b|v\rangle) = a\mathbf{A}(|u\rangle) + b\mathbf{A}(|v\rangle)$ for any $|u\rangle, |v\rangle \in \mathcal{H}$ and $a, b \in \mathbb{C}$. We can represent an operator through a matrix. Suppose that $\mathbf{A}$ is an operator on $\mathcal{H}$ and $\mathbf{e}_1, \ldots, \mathbf{e}_k$ form an orthonormal basis in $\mathcal{H}$. Then each $\mathbf{A}(|\mathbf{e}_j\rangle) \in \mathcal{H}$ and there exists a unique $k \times k$ matrix $(a_{j\ell})$ such that $\mathbf{A}(|\mathbf{e}_j\rangle) = \sum_{\ell=1}^k |\mathbf{e}_\ell\rangle a_{\ell j}$, $j = 1, \ldots, k$. We will identify operator $\mathbf{A}$ with matrix $(a_{j\ell})$ and use $\mathbf{A}$ for both operator and matrix $(a_{j\ell})$. An operator $\mathbf{A}$ on $\mathcal{H}$ is said to be self-adjoint if its corresponding matrix $\mathbf{A}$ is Hermitian, that is, $\mathbf{A} = \mathbf{A}^\dagger$. We also refer to self-adjoint operators as Hermitian operators. An operator $\mathbf{U}$ is said to be unitary if its corresponding matrix $\mathbf{U}$ is unitary, that is, $\mathbf{U}\mathbf{U}^\dagger = \mathbf{U}^\dagger\mathbf{U} = \mathbf{I}$. We say an operator $\mathbf{A}$ is semi-positive (or positive) definite if its corresponding matrix $\mathbf{A}$ is semi-positive (or positive) definite, that is, $\langle u|\boldsymbol{\rho}|u\rangle \geq 0$ for $|u\rangle \in \mathcal{H}$ (or $\langle u|\boldsymbol{\rho}|u\rangle \geq 0$ for $|u\rangle \in \mathcal{H}$ with equality only for $|u\rangle = 0$). The trace of an operator $\mathbf{A}$, denoted by $\mathrm{Tr}(\mathbf{A})$, is defined to be the trace of its corresponding matrix $\mathbf{A} = (a_{j\ell})$, that is, $\mathrm{Tr}(\mathbf{A}) = \sum_{j=1}^k a_{jj}$.

2.2 Quantum System

Quantum mechanics depicts phenomena at microscopic level such as position and momentum of an individual particle like an atom or electron, spin of an electron, detection of light photons, and the emission and absorption of light by atoms. Unlike classical mechanics where physical entities like position and momentum

can be measured precisely, the theory of quantum mechanics is intrinsically stochastic in a sense that we can only make probabilistic prediction about the results of the measurements performed.

Quantum mechanics is mathematically described by a Hilbert space $\mathcal{H}$ and self-adjoint operators on $\mathcal{H}$. A quantum system is completely characterized by its state and the time evolution of the state. A state is defined to be a unit vector in $\mathcal{H}$. Let $|\psi(t)\rangle$ be the state of the quantum system at time t, which is also referred to as a wave function. The states $|\psi(t_1)\rangle$ and $|\psi(t_2)\rangle$ at t_1 and t_2 are connected through $|\psi(t_2)\rangle = \mathbf{U}(t_1, t_2)|\psi(t_1)\rangle$, where $\mathbf{U}(t_1, t_2)$ is a unitary operator depending only on time t_1 and t_2. In fact, there exists a self-adjoint operator $\mathbf{H}$, which is known as the Hamiltonian of the quantum system, such that $\mathbf{U}(t_1, t_2) = \exp[-i\mathbf{H}(t_2 - t_1)]$. With Hamiltonian $\mathbf{H}$, we may describe the continuous time evolution of $|\psi(t)\rangle$ by the Schrödinger equation

$$(1) \qquad i\frac{\partial|\psi(t)\rangle}{\partial t} = \mathbf{H}|\psi(t)\rangle, \quad i = \sqrt{-1}.$$

Alternatively a quantum system can be described by a density operator (or density matrix). A density operator $\boldsymbol{\rho}$ is an operator on $\mathcal{H}$ which (1) is self-adjoint; (2) is semi-positive definite; (3) has unit trace [i.e., $\text{Tr}(\boldsymbol{\rho}) = 1$]. Following the convention in quantum information science, we reserve notation $\boldsymbol{\rho}$ for state, density operator or density matrix. A state is often classified as a pure state or an ensemble of pure states. A pure state is a unit vector $|\psi\rangle$ in $\mathcal{H}$, which corresponds to a density operator $\boldsymbol{\rho} = |\psi\rangle\langle\psi|$, and an ensemble of pure states corresponds to the case that the quantum system is in one of states $|\psi_j\rangle$, $j = 1, \ldots, J$, with probability p_j being in state $|\psi_j\rangle$, and the corresponding density operator

$$(2) \qquad \boldsymbol{\rho} = \sum_{j=1}^{J} p_j |\psi_j\rangle\langle\psi_j|.$$

See Griffiths (2004), Sakurai and Napolitano (2010) and Shankar (1994).

2.3 Quantum Probability

We can test the theory of quantum mechanics by checking its predictions with experiments of performing measurements on quantum systems in the laboratory. The usual quantum measurements are on observables such as position, momentum, spin, and so on, where an *observable* $\mathbf{X}$ is defined as a self-adjoint operator on Hilbert space $\mathcal{H}$. The *observable* definition

is motivated from the fact that the eigenvalues of self-adjoint operators are real. Assume that an observable $\mathbf{X}$ has a discrete spectrum with the following diagonal form

$$(3) \qquad \mathbf{X} = \sum_{a=1}^{p} x_a \mathbf{Q}_a,$$

where x_a are real eigenvalues of $\mathbf{X}$ and $\mathbf{Q}_a$ are the corresponding one-dimensional projections onto the orthogonal eigenvectors of $\mathbf{X}$. Consider such an observable in the quantum system prepared in state $\boldsymbol{\rho}$. Measure space $(\Omega, \mathcal{F})$ is used to describe possible measurement outcomes of the observable, and the result of the measurement is a random variable on $(\Omega, \mathcal{F})$ with probability distribution $P_{\boldsymbol{\rho}}$. We denote by X the result of the measurement of observable $\mathbf{X}$ given by (3). Then X is a random variable taking values in $\{x_1, x_2, \ldots, \}$, and under pure state $|\psi\rangle$, the probability that measurement outcome x_a occurs is defined to be

$$P(a) = P_{\boldsymbol{\rho}}(X = x_a)$$
$$= \langle\psi|\mathbf{Q}_a|\psi\rangle = \text{Tr}(\mathbf{Q}_a|\psi\rangle\langle\psi|), \quad a = 1, 2, \ldots.$$

With the probability we derive the expectation under pure state $|\psi\rangle$,

$$\mathbb{E}_\psi(\mathbf{X}) = \sum_a x_a P(a) = \sum_a x_a \langle\psi|\mathbf{Q}_a|\psi\rangle$$
$$= \langle\psi|\mathbf{X}|\psi\rangle = \text{Tr}(\mathbf{X}|\psi\rangle\langle\psi|).$$

Note the difference between an observable $\mathbf{X}$ which is a Hermitian matrix and its measurement result X which is a real-valued random variable.

Measuring observable $\mathbf{X}$ will alter the state of the quantum system (Kiefer, 2004; von Neumann, 1955). If the quantum system is prepared with initial state $|\psi\rangle$, the state of the system after the measurement result x_a is defined to be

$$(4) \qquad \frac{\mathbf{Q}_a|\psi\rangle}{\sqrt{P(a)}}.$$

For an ensemble state with density operator $\boldsymbol{\rho}$ given by (2), if the quantum state is $|\psi_j\rangle$, the probability that result x_a occurs is

$$P(a|j) = \langle\psi_j|\mathbf{Q}_a|\psi_j\rangle = \text{Tr}(\mathbf{Q}_a|\psi_j\rangle\langle\psi_j|).$$

Applying the law of total probability, we obtain that under state $\boldsymbol{\rho}$, the probability that x_a occurs is equal to

$$P(a) = P_{\boldsymbol{\rho}}(X = x_a) = \sum_{j=1}^{J} p_j P(a|j)$$

$$= \sum_{j=1}^{J} p_j \text{Tr}(\mathbf{Q}_a|\psi_j\rangle\langle\psi_j|) = \text{Tr}(\mathbf{Q}_a\boldsymbol{\rho}).$$

The expectation of $\mathbf{X}$ under state $\boldsymbol{\rho}$,

$$\mathbb{E}_{\boldsymbol{\rho}}[\mathbf{X}] = \sum_{a=1}^{p} x_a P_{\boldsymbol{\rho}}[X = x_a] = \sum_{a=1}^{p} x_a \operatorname{Tr}(\mathbf{Q}_a \boldsymbol{\rho})$$
$$= \operatorname{tr}(\mathbf{X}\boldsymbol{\rho}),$$

and variance

$$\operatorname{Var}_{\boldsymbol{\rho}}[\mathbf{X}] = \operatorname{tr}[\mathbf{X}^2 \boldsymbol{\rho}] - (\operatorname{tr}[\mathbf{X}\boldsymbol{\rho}])^2.$$

We may derive the density operator of the quantum system after obtaining the measurement result x_a by conditional probability arguments as follows. If the quantum system is in pure state $|\psi_j\rangle$ before the measurement, the quantum state after measurement result x_a has occurred is

$$|\psi_j^a\rangle = \frac{\mathbf{Q}_a |\psi_j\rangle}{\sqrt{P(a|j)}}.$$

If the quantum state is $\boldsymbol{\rho}$ before the measurement, after observing measurement outcome x_a we have the following ensemble of states: the quantum system is in pure state $|\psi_j^a\rangle$ with probability $P(j|a)$, where Bayes's theorem shows

$$P(j|a) = p_j P(a|j) / P(a).$$

Thus after measurement x_a the density operator for the ensemble state is given by

$$\boldsymbol{\rho}_a = \sum_{j=1}^{J} P(j|a) |\psi_j^a\rangle\langle\psi_j^a|$$
$$= \sum_{j=1}^{J} P(j|a) \frac{\mathbf{Q}_a |\psi_j\rangle\langle\psi_j| \mathbf{Q}_a}{P(a|j)}$$
$$= \sum_{j=1}^{J} p_j \frac{\mathbf{Q}_a |\psi_j\rangle\langle\psi_j| \mathbf{Q}_a}{P(a)} = \frac{\mathbf{Q}_a \boldsymbol{\rho} \mathbf{Q}_a}{\operatorname{Tr}(\mathbf{Q}_a \boldsymbol{\rho})}.$$

See Holevo (1982), Parthasarathy (1992) and Sakurai and Napolitano (2010).

2.4 Quantum Statistics

For a given quantum system, it is very important but difficult to know its state. If we do not know in advance the state of the quantum system, we may infer the quantum state by the measurement results of some observables obtained from the quantum system and show that a certain state has been created. In statistical terminology, we want to estimate density matrix $\boldsymbol{\rho}$ based on measurements on an often large number of systems which are identically prepared in the state $\boldsymbol{\rho}$. That is, after measuring observables on some identical quantum systems, we can make statistical inference

about probability distribution $P_{\boldsymbol{\rho}}$ of the measurements and thus indirectly about density matrix $\boldsymbol{\rho}$. In the literature of quantum physics, quantum tomography is referred to as the reconstruction of the underlying density matrix $\boldsymbol{\rho}$ by probing identically prepared quantum systems from some different angles. Specifically, suppose that we perform measurements of observables on identically prepared quantum systems in an unknown state $\boldsymbol{\rho}$ and obtain measurement results $X_1, \ldots, X_n$. Assume that $\boldsymbol{\rho}$ is known up to some unknown parameter θ; then $X_1, \ldots, X_n$ are i.i.d. observations with distributions $P_{\boldsymbol{\rho}}$ which depend on θ. This gives a quantum parametric statistical model. We may then define quantum likelihood and Fisher quantum information and establish quantum point estimation and quantum hypothesis testing theory. Alternatively we may model $\boldsymbol{\rho}$ nonparametrically by assuming that $\boldsymbol{\rho}$ is an infinite matrix and then use nonparametric methods to estimate the density matrix. For details see Artiles, Gill and Guţă (2005), Barndorff-Nielsen, Gill and Jupp (2003), Butucea, Guţă and Artiles (2007) and Nussbaum and Szkoła (2009).

3. QUANTUM COMPUTING CONCEPTS

Unlike classical computers using transistors to crunch the ones and zeroes individually, quantum computers can handle both one and zero simultaneously via what are known as superposition quantum states. A superposition state is a state of matter which we may think of as both one and zero at the same time. Quantum computers use the strange superposition states and quantum entanglements to do the trick of performing simultaneous calculations and extracting the calculated results. The spooky phenomena of quantum entanglement and superposition are the key that enables quantum computers to be superfast and vastly outperform classical computers.

3.1 Quantum Bit

Analogous to the fundamental concept of bit in classical computation and classical information, we have its counterpart, quantum bit, in quantum computation and quantum information. Quantum bit is called qubit for short. Just like a classical bit with state either 0 or 1, a qubit has states $|0\rangle$ and $|1\rangle$. However, the real difference between a bit and a qubit is that besides states $|0\rangle$ and $|1\rangle$, a qubit may take the superposition states,

$$|\psi\rangle = \alpha_0 |0\rangle + \alpha_1 |1\rangle,$$

where α_0 and α_1 are complex numbers and called amplitudes satisfying $|\alpha_0|^2 + |\alpha_1|^2 = 1$. That is, the states of a qubit are unit vectors in a two-dimensional complex vector space, and states $|0\rangle$ and $|1\rangle$ consist of an

orthonormal basis for the space and are often referred to as computational basis states. For a classical bit we can examine it to determine whether it is in the state 0 or 1. However, for a qubit we cannot determine its state and find the values of α_0 and α_1 by examining it. The stochastic nature of quantum theory shows that we can measure a qubit and obtain either the result 0, with probability $|\alpha_0|^2$, or the result 1, with probability $|\alpha_1|^2$. Physical experiments have realized qubits as physical objects in different physical systems, such as the two states of an electron orbiting a single atom, the two different polarizations of a photon, or the alignment of a nuclear spin in a uniform magnetic field. Consider the case of atom model by corresponding $|0\rangle$ and $|1\rangle$ with the so-called "ground" and "excited" states of the electron, respectively. As the atom is shined by light with suitable energy and for a proper amount of time, the electron can be moved from the $|0\rangle$ state to the $|1\rangle$ state and vice versa. Furthermore, by shortening the length of time shining the light on the atom, we may move the electron initially in the state $|0\rangle$ to "halfway" between $|0\rangle$ and $|1\rangle$, say, into a state $(|0\rangle + |1\rangle)/\sqrt{2}$.

Note that for qubit state $|\psi\rangle$ the only measurable quantities are the probabilities $|\alpha_0|^2$ and $|\alpha_1|^2$; since $|e^{i\theta}\alpha_x|^2 = |\alpha_x|^2$, where $x = 0, 1$, $i = \sqrt{-1}$, and θ is a real number, from the viewpoint of the qubit measurements, states $e^{i\theta}|\psi\rangle$ and $|\psi\rangle$ are identical. That is, multiplying a qubit state by a global phase factor $e^{i\theta}$ bears no observational consequence.

Note the distinction between superposition states and probability mixtures (or ensemble of pure states defined in Section 2.1). Consider superposition $(|0\rangle + |1\rangle)/\sqrt{2}$ as a pure state. Its density matrix is given by

$$\frac{1}{2}(|0\rangle + |1\rangle)(\langle 0| + \langle 1|)$$

$$= \frac{1}{2}(|0\rangle\langle 0| + |1\rangle\langle 1|) + \frac{1}{2}(|0\rangle\langle 1| + |1\rangle\langle 0|),$$

while the first term on the right-hand side of the above equation corresponds to the ensemble of pure states $|0\rangle$ and $|1\rangle$, that is, a probabilistic mixture of states $|0\rangle$ and $|1\rangle$ with equal probability.

Similar to classic bits, we can define multiple qubits. The states of b qubits are unit vectors in a 2^b-dimensional complex vector space with 2^b computational basis states of the form $|x_1 x_2 \cdots x_b\rangle$, $x_j = 0$ or 1, $j = 1, \ldots, b$. For example, the states of two qubits are unit vectors in a four-dimensional complex vector space, with four computational basis states labeled by $|00\rangle$, $|01\rangle$, $|10\rangle$ and $|11\rangle$. The computational basis states $|00\rangle$, $|01\rangle$, $|10\rangle$ and $|11\rangle$ generate the four-dimensional complex vector space, and the superposition states are all unit vector in the space with the forms

$$|\psi\rangle = \alpha_{00}|00\rangle + \alpha_{01}|01\rangle + \alpha_{10}|10\rangle + \alpha_{11}|11\rangle,$$

where amplitudes α_x are complex numbers satisfying $|\alpha_{00}|^2 + |\alpha_{01}|^2 + |\alpha_{10}|^2 + |\alpha_{11}|^2 = 1$. As in the single qubit case, when two qubits are measured we get result x being one of 00, 01, 10, 11, with probability $|\alpha_x|^2$. Moreover, we may measure just the first qubit of the two-qubit system and obtain either the result 0, with probability $|\alpha_{00}|^2 + |\alpha_{01}|^2$, or the result 1, with probability $|\alpha_{10}|^2 + |\alpha_{11}|^2$. As quantum measuring changes the quantum state, if the measurement result on the first qubit is 0, after the measurement the qubits are in the state

$$(5) \qquad \frac{\alpha_{00}|00\rangle + \alpha_{01}|01\rangle}{\sqrt{|\alpha_{00}|^2 + |\alpha_{01}|^2}}.$$

A qubit is the simplest quantum system. The quantum system of b qubits is described by a 2^b-dimensional complex vector space with each superposition state specified by 2^b amplitudes. As 2^b increases exponentially in b, it is very easy for such a system to have an enormously big vector space. A quantum system with even a few dozens of "qubits" will strain the resources of even the largest supercomputers. Consider a quantum system of 50 qubits. $2^{50} \approx 10^{15}$ complex amplitudes are needed to specify its quantum states. With 128 bits of precision, it requires approximately 32 thousand terabytes of information to store all 10^{15} complex amplitudes. Such storage capacity may be available in future supercomputers. For a quantum system with $b = 500$ qubits we need to specify 2^{500} complex amplitudes for its states. It is unimaginable to store all 2^{500} complex numbers in any classical computers. In principle, a quantum system with only a few hundred atoms can manage such an enormous amount of data and execute calculations as the system evolves. Quantum computation and quantum information are to find ways to utilize the immense potential computational power in quantum systems.

3.2 Quantum Circuit Model

As a classical computer is built from an electrical circuit consisting of wires for carrying information around the circuit and logic gates for performing simple computational tasks, a quantum computer can be created from a quantum circuit with quantum gates to perform quantum computation and manipulate quantum information. A number of physical systems are being investigated for building quantum computers. These include optical photon, optical cavity quantum electrodynamics, ion traps, nuclear magnetic resonance with molecules, quantum dots, and superconductors (Nielsen and Chuang, 2000). In fact, prim-

itive solid-state quantum processors have been created in research laboratories to run quantum algorithms (DiCarlo et al., 2009; Johnson et al., 2011; Mariantoni et al., 2011; Sayrin et al., 2011). The circuit model is particularly important in quantum computation and quantum information, and a quantum computer is often synonymous with the quantum circuit model. A quantum circuit operates on b qubits for some integer b. The state takes a form of $|x_1 \cdots x_b\rangle$, with state space being a 2^b-dimensional complex Hilbert space. When $x_i = 0$ or 1, states $|x_1 \cdots x_b\rangle$ are the computational basis states of the quantum computer and often written as $|x\rangle$, where x is the integer with binary representation $x_1 \cdots x_b$.

As a classical logic gate converts classical bits from one form to another such as $0 \to 1$ and $1 \to 0$, a quantum gate operates on qubits. Quantum mechanics dictates that quantum gates operating on b qubits are 2^b by 2^b unitary matrices on the 2^b-dimensional Hilbert space. For example, a Hadamard gate on one qubit is the 2×2 unitary matrix that realizes the following transformation:

$$|0\rangle \to \frac{|0\rangle + |1\rangle}{\sqrt{2}}, \quad |1\rangle \to \frac{|0\rangle - |1\rangle}{\sqrt{2}}.$$

Consider another important gate on two qubits which is called control-NOT gate. It takes the two input qubits as control qubit and target qubit, respectively, and the output target qubit of the gate retains the input target qubit if the control qubit is $|0\rangle$ and is flipped if the control qubit is $|1\rangle$, that is,

$$|00\rangle \to |00\rangle, \quad |01\rangle \to |01\rangle,$$
$$|10\rangle \to |11\rangle, \quad |11\rangle \to |10\rangle.$$

Generally for any single qubit unitary operation $\mathbf{U}$, a control-$\mathbf{U}$ gate is a two-qubit gate, with one control qubit and one target qubit. If the control qubit is $|1\rangle$, $\mathbf{U}$ is applied to the target qubit; if the control qubit is $|0\rangle$, the target qubit is left alone, that is,

$$|0\rangle|0\rangle \to |0\rangle|0\rangle, \quad |0\rangle|1\rangle \to |0\rangle|1\rangle,$$
$$|1\rangle|0\rangle \to |1\rangle\mathbf{U}|0\rangle, \quad |1\rangle|1\rangle \to |1\rangle\mathbf{U}|1\rangle.$$

If $f(x)$ maps $\{0, 1\}^b$ onto $\{0, 1\}$, we define a unitary transformation $\mathbf{U}_f$ that operates on $b + 1$ qubit state

$$(6) \qquad |x, y\rangle \to |x, y \oplus f(x)\rangle,$$

where $x = x_1 \cdots x_b$ with $x_j = 0$ or 1 is the data register, $y = 0$ or 1 is the target register, $\oplus$ denotes additional modulo 2. If $y = 0$, after the transformation $\mathbf{U}_f$, the state of the last qubit is the value of $f(x)$.

3.3 Entanglement

Quantum entanglement is one of the most mind-bending creatures known to science. It is referred to as the phenomenon that two qubits behave like twins that are connected by an invisible wave to share each other's properties.

3.3.1 *Bell states.* Consider a quantum gate on two-qubit basis states $|00\rangle$, $|01\rangle$, $|10\rangle$ and $|11\rangle$ that is composed of a Hadamard gate on the first qubit and then is followed by a control-NOT gate. The output states of the gate are as follows:

$$|00\rangle \to \frac{|00\rangle + |11\rangle}{\sqrt{2}}, \quad |01\rangle \to \frac{|01\rangle + |10\rangle}{\sqrt{2}},$$
$$|10\rangle \to \frac{|00\rangle - |11\rangle}{\sqrt{2}}, \quad |11\rangle \to \frac{|01\rangle - |10\rangle}{\sqrt{2}}.$$

Physicists Bell, Einstein, Podolsky and Rosen discovered the amazing properties of these four states, which are often referred to as the Bell states, EPR states or EPR pairs (Bell, 1964; Einstein, Podolsky and Rosen, 1935). In general states such as these four states that cannot be expressed as products of some single qubits are called entangled states. Entangled states, which are not fully understood in quantum physics, have remarkable properties.

For the two-qubit system consider a Bell state

$$|\psi\rangle = \frac{|01\rangle - |10\rangle}{\sqrt{2}},$$

and an observable

$$\mathbf{M} = a_x \boldsymbol{\sigma}_x + a_y \boldsymbol{\sigma}_y + a_z \boldsymbol{\sigma}_z,$$

where (a_x, a_y, a_z) is a real unit vector (i.e., $a_x^2 + a_y^2 + a_z^2 = 1$), and $\boldsymbol{\sigma}_x$, $\boldsymbol{\sigma}_y$ and $\boldsymbol{\sigma}_z$ are Pauli matrices given by

$$(7) \qquad \boldsymbol{\sigma}_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \boldsymbol{\sigma}_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix},$$
$$\boldsymbol{\sigma}_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad i = \sqrt{-1}.$$

It is easy to show that $\mathbf{M}$ has eigenvalues ± 1 for any real unit vector (a_x, a_y, a_z). If measuring observable $\mathbf{M}$ on each qubit of $|\psi\rangle$, we will obtain a measurement result of $+1$ or -1. Surprisingly, no matter what choice of (a_x, a_y, a_z), the measurement results on the two qubits are always opposite of each other, that is, when the first qubit measurement is -1, then the second qubit measurement will be $+1$, and vice versa.

The two-qubit system can be realized by the spins of two particles, and the measurement of $\mathbf{M}$ is referred to

as a measurement of spin along the (a_x, a_y, a_z) axis. After the two-particle system is prepared in the Bell state $|\psi\rangle$, the two particles drift far apart. Alice measures the spin of the first particle and Bob measures the spin of the second particle. The above opposite measurement phenomenon corresponds to that due to the entangled state $|\psi\rangle$, if Alice gets a result $+1$ from her spin measurement on the first particle, then the state of the system immediately jumps to the untangled state so that the second particle now has definite spin state and Bob's spin measurement on the second particle gives definite result -1. This phenomenon is often referred to as anti-correlation in entanglement experiments (Neumann et al., 2008; Sakurai and Napolitano, 2010).

The mathematical arguments for the anti-correlation phenomenon are as follows. The measurement of $\mathbf{M}$ on the first (or second) qubit of $|\psi\rangle$ corresponds to the spin measurement of Alice's (or Bob's) particle along the (a_x, a_y, a_z) axis in the above two-particle spin model. From Sections 2.3 and 3.1 we have that the two-qubit system is described by the Bell state $|\psi\rangle$ in $\mathbb{C}^4$; measuring $\mathbf{M}$ on the first qubit of $|\psi\rangle$ means performing measurement on observable $\mathbf{M} \otimes \mathbf{I}$ in the Bell state $|\psi\rangle$, which alters the quantum state of the two-qubit system; measuring $\mathbf{M}$ on the second qubit corresponds to measuring observable $\mathbf{I} \otimes \mathbf{M}$ in the altered quantum state, where $\mathbf{I}$ is the 2 by 2 identity matrix, and $\mathbf{M} \otimes \mathbf{I}$ and $\mathbf{I} \otimes \mathbf{M}$ are matrix tensor products.

Denote by $|\varphi_\pm\rangle$ the two orthonormal eigenvectors of $\mathbf{M}$ corresponding to eigenvalues ± 1, respectively, and let $\mathbf{Q}_\pm$ be the respective projections onto the eigenvectors $|\varphi_\pm\rangle$. Following (3)–(5) we have a diagonal representation $\mathbf{M} = \mathbf{Q}_+ - \mathbf{Q}_-$; when we measure observable $\mathbf{M}$ on each qubit, the possible measurement results are ± 1; measuring $\mathbf{M}$ on the first qubit changes the state of the two-qubit system, and after the measurement result ± 1 on the first qubit, the post-measurement state of the two-qubit system is $\mathbf{Q}_\pm \otimes \mathbf{I}|\psi\rangle / \|\mathbf{Q}_\pm \otimes \mathbf{I}|\psi\rangle\|$. Below we will evaluate the post-measurement state and show that measuring $\mathbf{I} \otimes \mathbf{M}$ in the post-measurement state always yields measurement results opposite to the measurement results on the first qubit.

Since $(|0\rangle, |1\rangle)$ and $(|\varphi_+\rangle, |\varphi_-\rangle)$ are two bases for the one-qubit system in $\mathbb{C}^2$, then

$$\begin{pmatrix} |0\rangle \\ |1\rangle \end{pmatrix} = \begin{bmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{bmatrix} \begin{pmatrix} |\varphi_+\rangle \\ |\varphi_-\rangle \end{pmatrix},$$

where $(\alpha_{j\ell})$ forms a 2×2 unitary matrix with determinant equal to a phase factor $e^{i\theta}$ ($i = \sqrt{-1}$) for some

real θ. Substituting the above expressions into the entangled state and ignoring a global phase factor $e^{i\theta}$ (which has no effects on measurement results; see Section 3.1), we obtain

$$\begin{aligned} |\psi\rangle &= \frac{|01\rangle - |10\rangle}{\sqrt{2}} = e^{i\theta} \frac{|\varphi_+\varphi_-\rangle - |\varphi_-\varphi_+\rangle}{\sqrt{2}} \\ &\sim \frac{|\varphi_+\varphi_-\rangle - |\varphi_-\varphi_+\rangle}{\sqrt{2}}. \end{aligned}$$
(8)

From the definitions of $|\varphi_\pm\rangle$ and $\mathbf{Q}_\pm$, $\mathbf{Q}_+ \otimes \mathbf{I}|\varphi_+\varphi_-\rangle = |\varphi_+\varphi_-\rangle$, $\mathbf{Q}_- \otimes \mathbf{I}|\varphi_-\varphi_+\rangle = -|\varphi_-\varphi_+\rangle$, $\mathbf{Q}_+ \otimes \mathbf{I}|\varphi_-\varphi_+\rangle = 0$, and $\mathbf{Q}_- \otimes \mathbf{I}|\varphi_+\varphi_-\rangle = 0$. If the measurement result of $\mathbf{M}$ on the first qubit is $+1$ (or -1), from (8) we obtain the post-measurement state of the two-qubit system as follows:

$$\begin{aligned} &\frac{\mathbf{Q}_+ \otimes \mathbf{I}|\psi\rangle}{\|\mathbf{Q}_+ \otimes \mathbf{I}|\psi\rangle\|} \\ &= e^{i\theta} \frac{\mathbf{Q}_+ \otimes \mathbf{I}|\varphi_+\varphi_-\rangle - \mathbf{Q}_+ \otimes \mathbf{I}|\varphi_-\varphi_+\rangle}{\sqrt{2}\|\mathbf{Q}_+ \otimes \mathbf{I}|\psi\rangle\|} \\ &= e^{i\theta} \frac{|\varphi_+\varphi_-\rangle}{\||\varphi_+\varphi_-\rangle\|} = e^{i\theta}|\varphi_+\varphi_-\rangle \sim |\varphi_+\varphi_-\rangle \end{aligned}$$

(or $\mathbf{Q}_- \otimes \mathbf{I}|\psi\rangle / \|\mathbf{Q}_- \otimes \mathbf{I}|\psi\rangle\| = e^{i\theta}|\varphi_-\varphi_+\rangle \sim |\varphi_-\varphi_+\rangle$). Since

$$\begin{aligned} \mathbf{I} \otimes \mathbf{M}|\varphi_+\varphi_-\rangle &= \mathbf{I} \otimes \mathbf{Q}_+|\varphi_+\varphi_-\rangle - \mathbf{I} \otimes \mathbf{Q}_-|\varphi_+\varphi_-\rangle \\ &= -|\varphi_+\varphi_-\rangle, \end{aligned}$$

$$\mathbf{I} \otimes \mathbf{M}|\varphi_-\varphi_+\rangle = |\varphi_-\varphi_+\rangle,$$

that is, the post-measurement state $|\varphi_+\varphi_-\rangle$ (or $|\varphi_-\varphi_+\rangle$) is the eigenvector of $\mathbf{I} \otimes \mathbf{M}$ corresponding to eigenvalue -1 (or $+1$), performing measurement on $\mathbf{I} \otimes \mathbf{M}$ in the post-measurement state must always yield measurement result -1 (or $+1$). Thus, the measurement results of $\mathbf{M}$ on the two qubits of $|\psi\rangle$ are always opposite to each other.

3.3.2 *Quantum teleportation.* Quantum teleportation is a process by which we can transfer the state of a qubit from one location to another, without transmitting it through the intervening space. We illustrate the phenomenon as follows. Alice and Bob together generated a Bell state long ago. Each took one qubit of the Bell state when they split. Now they are far away from each other. The mission for Alice is to deliver a qubit $|\psi\rangle$ to Bob, while he is hiding, and she can only send classical information to Bob but does not know the state of the qubit $|\psi\rangle$. Quantum teleportation is a way that Alice utilizes the entangled Bell state to send a qubit of unknown state to Bob, with only a

small overhead of classical communication. Recently a breakthrough in quantum teleportation has been made by successfully transferring complex quantum data instantaneously from one place to another, paving the way for real-world applications of quantum communications (Lee et al., 2011).

Here is how it works. Alice interacts the qubit $|\psi\rangle$ to be teleported with her half of the Bell state, and then performs a measurement on the two interacted qubits to obtain one of four possible two-classical-bit results: $00, 01, 10$ and 11. She sends the two-bit information via classical communication to Bob. Depending on Alice's classical message, Bob performs one of four operations on his half of the Bell state. Surprisingly, the described procedure allows Bob to recover the original state $|\psi\rangle$.

Specifically assume that the state to be teleported is $|\psi\rangle = \alpha_0|0\rangle + \alpha_1|1\rangle$, where α_0 and α_1 are unknown amplitudes. First, consider a three-qubit state

$$|\varphi_0\rangle = |\psi\rangle \frac{|00\rangle + |11\rangle}{\sqrt{2}}$$

$$= \frac{1}{\sqrt{2}}[\alpha_0|0\rangle(|00\rangle + |11\rangle) + \alpha_1|1\rangle(|00\rangle + |11\rangle)],$$

where the first two qubits (on the left) belong to Alice, and the third qubit to Bob. Note that Alice's second qubit and Bob's third qubit are from the entangled Bell state. Second, Alice applies a control-NOT gate to her qubits in $|\varphi_0\rangle$ and obtains

$$|\varphi_1\rangle = \tfrac{1}{\sqrt{2}}[\alpha_0|0\rangle(|00\rangle + |11\rangle) + \alpha_1|1\rangle(|10\rangle + |01\rangle)].$$

Third, she applies a Hadamard gate to the first qubit in $|\varphi_1\rangle$ and gets

$$|\varphi_2\rangle = \tfrac{1}{2}[\alpha_0(|0\rangle + |1\rangle)(|00\rangle + |11\rangle)$$
$$+ \alpha_1(|0\rangle - |1\rangle)(|10\rangle + |01\rangle)].$$

We regroup the terms of $|\varphi_2\rangle$ and rewrite it as follows:

$$|\varphi_2\rangle = \tfrac{1}{2}[|00\rangle(\alpha_0|0\rangle + \alpha_1|1\rangle) + |01\rangle(\alpha_0|1\rangle + \alpha_1|0\rangle)$$
$$+ |10\rangle(\alpha_0|0\rangle - \alpha_1|1\rangle)$$
$$+ |11\rangle(\alpha_0|1\rangle - \alpha_1|0\rangle)].$$

The new expression has four terms, and each term has Alice's qubits in one of four possible states $|00\rangle$, $|01\rangle$, $|10\rangle$ and $|11\rangle$, and Bob's qubit is in the state related to the original state $|\psi\rangle$. If Alice performs a measurement on her qubits and informs Bob of the measurement result, then his post-measurement state is completely determined. For example, the first term has Alice's qubits

in the state $|00\rangle$ and Bob's qubit in state $|\psi\rangle$. Therefore, if Alice's measurement result on her qubits is 00, then Bob's qubit will be in state $|\psi\rangle$. Below is a list of Bob's four post-measurement states corresponding to the results of Alice's measurements:

$$00 \to \alpha_0|0\rangle + \alpha_1|1\rangle, \quad 01 \to \alpha_0|1\rangle + \alpha_1|0\rangle,$$
$$10 \to \alpha_0|0\rangle - \alpha_1|1\rangle, \quad 11 \to \alpha_0|1\rangle - \alpha_1|0\rangle.$$

As Alice's measurement outcome on her qubits is one of $00, 01, 10$ and 11, depending on her measurement outcome Bob's qubit will be one of the above four possible states. Once Alice sends to Bob her two-classical-bit measurement outcome through a classical channel, he applies appropriate quantum gates to his state and recovers $|\psi\rangle$. For example, if her measurement is 00, Bob's state is $|\psi\rangle$, and he does not need to apply any quantum gate. If her measurement is 01, then Bob needs to apply a σ_x gate to his state $\alpha_0|1\rangle + \alpha_1|0\rangle$ and yields $|\psi\rangle$. If her measurement is 10, then applying a σ_z gate to his state $\alpha_0|0\rangle - \alpha_1|1\rangle$ Bob recovers $|\psi\rangle$. If her measurement is 11, then Bob can fix up his state $\alpha_0|1\rangle - \alpha_1|0\rangle$ to recover $|\psi\rangle$ by applying first a σ_x gate and then a σ_z gate. Here the σ_x and σ_z gates are defined by Pauli matrices σ_x and σ_z given by (7). In summary, according to Alice's measurement outcome, applying some appropriate quantum gates to his qubit Bob will recover the state $|\psi\rangle$.

A few important remarks about quantum teleportation are in the line. First, quantum teleportation does not involve any transfer of matter or energy. Alice's particle has not been physically moved to Bob; only its state has been transferred. Second, after the teleportation Bob's qubit will be on the teleported state, while Alice's qubit will become some undefined part of an entangled state. In other words, what the teleportation does is that a qubit was destroyed in one place but instantaneously resurrected in another. Teleportation does not copy any qubits, and hence is consistent with the no-cloning theorem (which forbids the creation of identical copies of an arbitrary unknown quantum state; see Wootters and Zurek, 1982). Third, in order to teleportate a qubit, Alice has to inform Bob of her measurement by sending him two classical bits of information. These two classical bits do not carry complete information about the qubit being teleported. If the two bits are intercepted by an eavesdropper, he or she may know exactly what Bob needs to do in order to recover the desired state. However, this information is useless if the eavesdropper cannot interact with the entangled particle in Bob's possession. Also the requirement of sending two bits of information via classical

channel prevents quantum teleportation from transmitting information faster than the speed of light.

3.3.3 *Bell's inequality.* The Bell test experiments are designed to investigate the validity of the entanglement effect in quantum mechanics through Bell's inequality. Over the past four decades many physical experiments on quantum systems were conducted to check the validity of Bell's inequality and resulted in some violation of the inequality. For example, Aspect, Grangier and Roger (1981, 1982a, 1982b) provided overwhelming support to the violation of Bell's inequality. The experimental results are often invoked as the proof of quantum non-locality and lack of realism that no particle has definite form until it is measured and measuring a quantum entity can instantaneously influence another far away. See Aspect, Grangier and Roger (1981, 1982a, 1982b), Bohm (1951), Bell (1964), Clauser et al. (1969) and Einstein, Podolsky and Rosen (1935). Below we describe the CHSH version of the Bell's inequality (Clauser et al., 1969).

Suppose X_i, $i = 1, 2, 3, 4$, are four random variables taking values ± 1. Consider an ordinary experiment with two people, Alice and Bob. In the experiment Alice observes X_1 or X_2 while Bob measures X_3 or X_4. Consider the quantity $X_1 X_3 + X_2 X_3 + X_2 X_4 - X_1 X_4$. It is equal to

$$(X_1 + X_2)X_3 + (X_2 - X_1)X_4 = \pm 2 \le 2.$$

Regardless of the distributions of X_i, taking expectation on both sides of the above inequality we arrive at the famous Bell inequality,

$$\begin{aligned} & E(X_1 X_3) + E(X_2 X_3) + E(X_2 X_4) \\ (9) \quad & \qquad - E(X_1 X_4) \le 2. \end{aligned}$$

The violation of Bell's inequality demonstrates entanglement effect in quantum mechanics. In fact, quantum experiments yield a quantum version of the inequality. Consider that a quantum system of two qubits is prepared in a Bell state

$$|\psi\rangle = \frac{|01\rangle - |10\rangle}{\sqrt{2}}.$$

Alice takes the first qubit of $|\psi\rangle$ while Bob gets its second qubit. Define four observables with eigenvalues ± 1,

$$\mathbf{X}_1 = \boldsymbol{\sigma}_z, \quad \mathbf{X}_2 = \boldsymbol{\sigma}_x,$$

on the first qubit and

$$\mathbf{X}_3 = -\frac{\boldsymbol{\sigma}_z + \boldsymbol{\sigma}_x}{\sqrt{2}}, \quad \mathbf{X}_4 = \frac{\boldsymbol{\sigma}_z - \boldsymbol{\sigma}_x}{\sqrt{2}},$$

on the second qubit, where $\boldsymbol{\sigma}_x$ and $\boldsymbol{\sigma}_z$ are Pauli matrices given by (7). Again Alice performs measurements on $\mathbf{X}_1$ or $\mathbf{X}_2$ while Bob measures $\mathbf{X}_3$ or $\mathbf{X}_4$. The quantum expectations of $\mathbf{X}_1\mathbf{X}_3$, $\mathbf{X}_2\mathbf{X}_3$, $\mathbf{X}_2\mathbf{X}_4$, $\mathbf{X}_1\mathbf{X}_4$ in the state $|\psi\rangle$ are calculated below:

$$\mathbb{E}_\psi(\mathbf{X}_1\mathbf{X}_3) = \frac{1}{\sqrt{2}}, \quad \mathbb{E}_\psi(\mathbf{X}_2\mathbf{X}_3) = \frac{1}{\sqrt{2}},$$

$$\mathbb{E}_\psi(\mathbf{X}_2\mathbf{X}_4) = \frac{1}{\sqrt{2}}, \quad \mathbb{E}_\psi(\mathbf{X}_1\mathbf{X}_4) = -\frac{1}{\sqrt{2}}.$$

Here the observable product is in the sense of tensor product. Thus we obtain a value in the quantum framework for the analog quantity on the left-hand side of the Bell's inequality (9)

$$\begin{aligned} & \mathbb{E}_\psi(\mathbf{X}_1\mathbf{X}_3) + \mathbb{E}_\psi(\mathbf{X}_2\mathbf{X}_3) + \mathbb{E}_\psi(\mathbf{X}_2\mathbf{X}_4) \\ & \qquad - \mathbb{E}_\psi(\mathbf{X}_1\mathbf{X}_4) = 2\sqrt{2}, \end{aligned}$$

which exceeds 2 and hence violates the Bell's inequality. In fact, the quantum version of the Bell's inequality is the Tsirelson's inequality (Tsirelson, 1980) which shows that in any quantum state $\boldsymbol{\rho}$,

$$\begin{aligned} (10) \quad & \mathbb{E}_{\boldsymbol{\rho}}(\mathbf{X}_1\mathbf{X}_3) + \mathbb{E}_{\boldsymbol{\rho}}(\mathbf{X}_2\mathbf{X}_3) + \mathbb{E}_{\boldsymbol{\rho}}(\mathbf{X}_2\mathbf{X}_4) \\ & \qquad - \mathbb{E}_{\boldsymbol{\rho}}(\mathbf{X}_1\mathbf{X}_4) \le 2\sqrt{2}. \end{aligned}$$

3.4 Quantum Parallelism

Quantum computation has an amazing feature termed as quantum parallelism, which may be heuristically explained by the following oversimplifying description: a quantum computer can simultaneously evaluate the whole range of a function $f(x)$ at many different values of x.

For function $f(x)$ with b bit input $x = x_1 \cdots x_b$ and 1 bit output $f(x)$, we illustrate quantum parallel evaluation of its values at many different x simultaneously as follows. First we apply b Hadamard gates to the first b qubits of $|0 \cdots 0\rangle|0\rangle$ to obtain

$$\frac{|0\rangle + |1\rangle}{\sqrt{2}} \cdots \frac{|0\rangle + |1\rangle}{\sqrt{2}}|0\rangle = \frac{1}{\sqrt{2^b}} \sum_x |x\rangle|0\rangle,$$

$$x = x_1 \cdots x_b, x_j = 0, 1,$$

where the sum is over all possible 2^b values of x. Second, apply quantum circuit $\mathbf{U}_f$ defined in (6) to the obtained $b + 1$ qubit state to yield

$$\frac{1}{\sqrt{2^b}} \sum_x |x\rangle|f(x)\rangle.$$

The quantum circuit with b Hadamard gates is extremely efficient in producing an equal superposition

of all 2^b computational basis states with only b gates; and quantum parallelism enables simultaneous evaluation of the whole range of the function f, although we evidently evaluate f just once with single quantum circuit $\mathbf{U}_f$ applied to the superposition state. To make it more clear we consider the case of $b = 1$. Apply circuit $\mathbf{U}_f$ to a superposition state as follows:

$$\mathbf{U}_f\left(\frac{|0\rangle + |1\rangle}{\sqrt{2}}|0\rangle\right) = \mathbf{U}_f\left(\frac{|00\rangle + |10\rangle}{\sqrt{2}}\right)$$
$$= \frac{|0f(0)\rangle + |1f(1)\rangle}{\sqrt{2}}.$$

One application of a single circuit $\mathbf{U}_f$ results in a superposition state whose two components contain information about both $f(0)$ and $f(1)$, as if we have evaluated $f(x)$ at values 0 and 1 simultaneously. The quantum parallelism is in contrast with classical parallelism, where multiple circuits each built to compute one value of $f(x)$ are executed simultaneously. Quantum parallelism arises from superposition states. A superposition state has many components, each of which may be thought of as a single argument to function $f(x)$. Because of quantum nature, a single circuit $\mathbf{U}_f$ applied once to the superposition state is actually performed on each of the components of the superposition, and the whole range of the values of function $f(x)$ is stored in the resulted outcome superposition state.

The quantum parallelism can be a potentially powerful tool for computational statistics. For example, Bayesian analysis often encounters the problems of evaluating sums over 2^b quantities, with b proportional to sample size or the number of variables. For moderate to large b, the evaluation of such sums is computationally prohibitive by classical computers (Vidakovic, 1999). Because of the quantum parallelism, it is possible for quantum computers to perform such computing tasks.

4. QUANTUM ALGORITHMS

Quantum algorithms are described by quantum circuits that take input qubits and yield output measurements for the solutions of the given problems. As a classical algorithm is a step-by-step problem-solving procedure, with each step performed on a classical computer, a quantum algorithm is a step-by-step procedure to solve a problem, with each step executed by a quantum computer. Although all classical algorithms can also be carried out on a quantum computer, we refer to quantum algorithms as the algorithms that utilize essential quantum features such as quantum superposition and quantum entanglement. While it is true that all problems solvable on a quantum computer are solvable on a classical computer, and problems undecidable by classical computers remain undecidable on quantum computers, what makes quantum algorithms exciting is the faster speed that quantum algorithms might be able to achieve, compared to classical algorithms, for solving some tough problems. The well-known quantum algorithms are Shor's factoring algorithm and Grover's search algorithm. Shor's algorithm and Grover's algorithm run, respectively, exponentially faster and quadratically faster than the best known classical algorithms for the same tasks. Common techniques used in quantum algorithms include quantum Fourier transform, phase estimation and quantum walk.

4.1 Quantum Fourier Transform

The quantum Fourier transform is defined to be a linear transformation on n qubits that maps the computational basis states $|j\rangle$, $j = 0, 1, \ldots, 2^n - 1$, to superposition states as follows:

$$|j\rangle \longrightarrow \frac{1}{\sqrt{2^n}} \sum_{k=0}^{2^n-1} e^{2\pi i jk/2^n} |k\rangle, \quad i = \sqrt{-1}.$$

The inverse of quantum Fourier transform is given by

$$|k\rangle \longrightarrow \frac{1}{\sqrt{2^n}} \sum_{j=0}^{2^n-1} e^{-2\pi i jk/2^n} |j\rangle.$$

We use the binary representation to express the state $j = j_1 2^{n-1} + j_2 2^{n-2} + \cdots + j_n 2^0$ as $j = j_1 j_2 \cdots j_n$ and represent binary fraction $j_\ell/2 + j_{\ell+1}/2^2 + \cdots + j_m/2^{m-\ell+1}$ as $0.j_\ell j_{\ell+1} \cdots j_m$, where $1 \leq \ell \leq m \leq 2^n$. Then the quantum Fourier transform of state $|j\rangle = |j_1 j_2 \cdots j_n\rangle$ has the following useful product representation:

$$|j_1 j_2 \cdots j_n\rangle \rightarrow \frac{1}{2^{n/2}}(|0\rangle + e^{2\pi i 0.j_n}|1\rangle)$$
$$\cdot (|0\rangle + e^{2\pi i 0.j_{n-1}j_n}|1\rangle) \cdots$$
$$\cdot (|0\rangle + e^{2\pi i 0.j_1 j_2 \cdots j_n}|1\rangle).$$

It can be easily checked from the product representation that with quantum parallelism the quantum Fourier transform can be realized as a quantum circuit with only $O(n^2)$ operations, while classically the fast Fourier transform requires $O(n2^n)$ operations for processing 2^n data, which indicates an exponential speedup (Nielsen and Chuang, 2000). Realizing such an exponential saving accommodated by quantum parallelism requires clever measurement schemes. Successful examples include quantum phase estimation and Shor's algorithms for factoring and discrete logarithm.

4.2 Phase Estimation

Quantum algorithms are of random nature in the sense that they are able to produce correct answers only with some probabilities. Consider quantum phase estimation which provides the key to many quantum algorithms. Assume that a unitary operator $\mathbf{U}$ has an eigenvector $|x\rangle$ with eigenvalue $e^{2\pi i\varphi}$. The phase φ of the eigenvalue is unknown and the goal of the phase estimation algorithm is to estimate φ based on the assumption that the state $|x\rangle$ can be prepared and the controlled-$\mathbf{U}^{2^j}$ operations [see Section 3.2 for control gate] can be performed for suitable nonnegative integers j.

The registers are used in phase estimation. The first register consists of b qubits initially in the state $|0\rangle$. The second register starts in the state $|x\rangle$ and involves enough qubits to store $|x\rangle$. The phase estimation procedure is performed in two stages.

First, we apply Hadamard transform to the first register and then controlled-$\mathbf{U}$ operations on the second register, with $\mathbf{U}$ raised to successive powers of 2, to obtain the final state with the second register unchanged and the first register given by

$$\frac{1}{2^{b/2}}(|0\rangle + e^{2\pi i 2^{b-1}\varphi}|1\rangle)$$

$$(11) \qquad \cdot (|0\rangle + e^{2\pi i 2^{b-2}\varphi}|1\rangle)\cdots(|0\rangle + e^{2\pi i 2^0\varphi}|1\rangle)$$

$$= \frac{1}{2^{b/2}}\sum_{k=0}^{2^b-1}e^{2\pi i\varphi k}|k\rangle.$$

If φ is expressed exactly in b bits as $\varphi = 0.\varphi_1\cdots\varphi_b$, (11) becomes

$$\frac{1}{2^{b/2}}(|0\rangle + e^{2\pi i 0.\varphi_b}|1\rangle)$$

$$(12) \qquad \cdot(|0\rangle + e^{2\pi i 0.\varphi_{b-1}\varphi_b}|1\rangle)\cdots$$

$$\cdot(|0\rangle + e^{2\pi i 0.\varphi_1\varphi_1\cdots\varphi_b}|1\rangle),$$

which is the quantum Fourier transform of the product state $|\varphi_1\varphi_2\cdots\varphi_b\rangle$.

The second stage of phase estimation is to take the inverse quantum Fourier transform on the first register. For $\varphi = 0.\varphi_1\cdots\varphi_b$, the output state from the second stage is $|\varphi_1\varphi_2\cdots\varphi_b\rangle$, and a measurement in the computational basis yields $\varphi_1\cdots\varphi_b$ and dividing the measurement by 2^b gives $\varphi_1\cdots\varphi_b/2^b = 0.\varphi_1\cdots\varphi_b = \varphi$. We obtain a perfect estimate of φ.

Now we consider the case that φ cannot be expressed exactly with a b bit binary expansion. Take $0 \le \eta < 2^b$ to be the integer that its binary fraction $\eta/2^b =$

$0.\eta_1\eta_2\cdots\eta_b$ is the first b bit representation in the binary expansion of φ, which satisfies $0 \le \varphi - \eta/2^b \le 2^{-b}$.

Perform the inverse quantum Fourier transform on the first register given by (11), which is obtained in the first stage results, and get

$$\frac{1}{2^b}\sum_{k,\ell=0}^{2^b-1}e^{-2\pi i k\ell/2^b}e^{2\pi i\varphi k}|\ell\rangle = \sum_{\ell=0}^{2^b-1}\beta_\ell|\ell\rangle,$$

where amplitudes of $|(\eta + \ell)(\mathrm{mod}\, 2^b)\rangle$ are

$$\beta_\ell = \frac{1}{2^b}\sum_{k=0}^{2^b-1}\{e^{2\pi i[\varphi-(\eta+\ell)/2^b]}\}^k$$

$$= \frac{1}{2^b}\left(\frac{1 - e^{2\pi i(2^b\varphi-\eta-\ell)}}{1 - e^{2\pi i(\varphi-\eta/2^b-\ell/2^b)}}\right).$$

Assume that the result of the final measurement from phase estimation is $\tilde{\eta}$ and dividing the result by 2^b gives $\tilde{\varphi} = \tilde{\eta}/2^b$. Let ζ be the specified accuracy for the phase estimation procedure. By adding up $|\beta_\ell|^2$ with ℓ being within $\zeta 2^b$, we bound the probability that the obtained $\tilde{\varphi}$ is within ζ from φ:

$$P(|\tilde{\varphi} - \varphi| \le \zeta) \ge P(|\tilde{\eta} - \eta| \le \zeta 2^b - 1)$$

$$\ge 1 - \frac{1}{2(\zeta 2^b - 2)}.$$

For $\epsilon > 0$, set

$$(13) \qquad b = \left\lceil\log_2\left(\frac{1}{\zeta}\right)\right\rceil + \left\lceil\log_2\left(2 + \frac{1}{2\epsilon}\right)\right\rceil.$$

Then $P(|\tilde{\varphi} - \varphi| \le \zeta) \ge 1 - \epsilon$, that is, with probability at least $1 - \epsilon$ the phase estimation procedure can successfully produce $\tilde{\varphi}$ within ζ from the true φ. See Nielsen and Chuang (2000).

4.3 Statistical Analysis

The phase estimation algorithm requires b qubits for the first register to achieve $[-\log_2\zeta]$ bit accuracy and success probability $1 - \epsilon$. With accuracy fixed, to increase the success probability the required qubits

$$b \sim 1 - \log_2\zeta + \frac{1}{4\epsilon\log 2},$$

which grows at a very fast rate. For example, an increase in success probability from 90% to 99% requires eighteen times of qubit increase compared to the change from 80% to 90%.

Quantum algorithms are of random nature in the sense that they often produce correct answers only with

certain probabilities. The success probabilities depend upon the schemes of the algorithms as well as the context of applications. Given a quantum algorithm for solving a problem, a common practice is to repeatedly run the quantum algorithm to achieve high probability of successfully obtaining a correct answer. Consider that the phase estimation procedure is repeatedly run n times to obtain results $\tilde{\varphi}_1, \ldots, \tilde{\varphi}_n$. Then $\tilde{\varphi}_1, \ldots, \tilde{\varphi}_n$ may be treated as i.i.d. random variables with each $\tilde{\varphi}_j$ satisfying

$$P(|\tilde{\varphi}_j - \varphi| > \zeta) \leq \epsilon.$$

We may statistically model $\tilde{\varphi}_j$ by the gross error model (Huber and Ronchetti, 2009) as follows. Assume that $\tilde{\varphi}_j$ are independently and identically generated from $(1 - \epsilon)F(x) + \epsilon H(x)$, where $F(x)$ is the distribution of the correct answers that are within ζ from true φ, and $H(x)$ is the distribution of wrong answers that are at least ζ away from true φ. Then $\tilde{\varphi}_j$ are correct with probability $1 - \epsilon$ and incorrect with probability ϵ.

If the outcome result of the algorithm is verifiable to be a correct answer or not [as in the case of Shor's algorithms for factoring and order-finding in Section 4.4 below], the obtained result from each run is checked to be a correct answer or not. Then the number of times required to run the algorithm for obtaining a correct answer follows a geometric distribution. Thus the probability that we obtain a correct answer in n repetitions is equal to

$$P(\text{obtain a correct answer in } n \text{ trials})$$
$$= 1 - P(\text{no success in the } n \text{ trials}) = 1 - \epsilon^n.$$

Since ϵ^n goes to zero geometrically fast, we may choose a moderate ϵ with fewer qubits to achieve very high probability of successfully obtaining a correct answer by repeatedly running the algorithm enough times.

On the other hand, if the outcome result is not verifiable to be a correct answer or not [as in the case of phase estimation], careful analysis is needed to design ways for obtaining a correct answer with very high probability. As wrong answers are far away from true φ, estimators like sample average of $\tilde{\varphi}_1, \ldots, \tilde{\varphi}_n$ may not estimate φ well. We adopt a robust statistical method to estimate φ by α-trimmed mean $\bar{\varphi}$, which is defined as follows. Ordering $\tilde{\varphi}_1, \ldots, \tilde{\varphi}_n$ and then removing $[n\alpha]$ largest and $[n\alpha]$ smallest ones, we take the average of the remaining $\tilde{\varphi}_j$ as α-trimmed mean, where α is chosen to be greater than $\epsilon/2$. One example is the sample median of $\tilde{\varphi}_1, \ldots, \tilde{\varphi}_n$. The probability

that $\bar{\varphi}$ is within ζ from φ can be calculated from the binomial probability as follows:

$$P(|\bar{\varphi} - \varphi| \leq \zeta) \geq P\big(\text{more than } n(1 - 2\alpha) \text{ number}$$
$$\text{of } \tilde{\varphi}_j \text{ are within } \zeta \text{ from } \varphi\big)$$
$$= \sum_{k=[n(1-2\alpha)]-1}^{n} \binom{n}{k}(1 - \epsilon)^k \epsilon^{n-k}.$$

As $n \to \infty$, $n\epsilon$ approaches to infinity. The binomial probability can be approximated by resorting to a normal approximation, yielding

$$\sum_{k=[n(1-2\alpha)]-1}^{n} \binom{n}{k}(1 - \epsilon)^k \epsilon^{n-k}$$
$$\sim 1 - \Phi\left(\frac{\sqrt{n}(\epsilon - 2\alpha)}{\sqrt{\epsilon(1 - \epsilon)}}\right) = \Phi\left(\frac{\sqrt{n}(2\alpha - \epsilon)}{\sqrt{\epsilon(1 - \epsilon)}}\right),$$

where $\Phi(\cdot)$ is the standard normal distribution function. Since $2\alpha - \epsilon > 0$, as n increases, $P(|\bar{\varphi} - \varphi| \leq \zeta)$ approaches to 1 exponentially fast. Combining the two cases together, we arrive at the following theorem.

THEOREM 1. *Suppose that the outcome $\tilde{\varphi}$ of a quantum algorithm obeys the gross error model that with probability $1 - \epsilon$ it produces a correct answer and probability ϵ it gives a wrong answer. Then by repeatedly running the quantum algorithm we will obtain a correct answer with probability approaching 1 exponentially fast in the number of repetitive runs.*

For a quantum algorithm that produces a correct answer with probability 70% and $\alpha = 0.2$, in order to obtain a correct answer with 0.999 probability we need to run the algorithm five times and 20 times, respectively, for the cases that the outcome results are verifiable and not verifiable.

4.4 Factoring and Order-Finding Algorithms

The factoring problem is to find all prime factors of a given positive composite number such that the product of these prime numbers is equal to the composite number. Factoring is known to be a very hard problem for classical computers. Shor (1994, 1997) developed a quantum algorithm for the factoring problem that is exponentially faster than the most efficient known classical factoring algorithm.

Shor's quantum algorithms work as follows. Mathematically the factoring problem is equivalent to the order-finding problem that for two positive integers x and N, $x < N$, with no common factors, find the smallest integer r such that dividing x^r by N we obtain a

reminder 1 (Shor, 1997; Nielsen and Chuang, 2000). The quantum algorithm for factoring is reduced to a quantum algorithm for order-finding. The quantum algorithm for order-finding is to apply the phase estimation algorithm to the unitary operator

$$\mathbf{U}|y\rangle = |xy(\bmod N)\rangle.$$

The eigenvectors of $\mathbf{U}$ are

$$|u_s\rangle = \frac{1}{\sqrt{r}} \sum_{k=0}^{r-1} \exp\left(\frac{-2\pi i s k}{r}\right)|x^k \bmod N\rangle,$$

$$s = 0, 1, \ldots, r-1, i = \sqrt{-1},$$

with corresponding eigenvalues $\exp(2\pi i s/r)$. Using the phase estimation algorithm we can obtain the eigenvalues $\exp(2\pi i s/r)$ with high accuracy and thus find the order r with certain probability.

While the quantum factoring algorithm can accomplish the task of factoring an n-bit integer with operations of order $n^2 \log n \log \log n$, the current best known classical algorithm requires operations of order $\exp(n^{1/3} \log^{2/3} n)$ to factor an n-bit composite number (Crandall and Pomerance, 2001). Note that the number of operations required in the best classical algorithm grows exponentially in the size of the number being factored. Because of the exponential complexity, the factoring problem is generally regarded as an intractable problem on classical computers.

The factoring problem plays an important role in cryptography. Cryptography is to enable two parties, Alice and Bob, to communicate privately, while it is very difficult for the third parties to "eavesdrop" on the contents of the communications. Examples include ATM cards, computer passwords, internet commences, clandestine meetings and military communications. Two cryptographic protocols used in the communications are private key cryptosystem and public key cryptosystem. A private key cryptosystem requires the two communicating parties to share a private key. Alice uses the key to encrypt the information, sends the encrypted information to Bob who uses the key to decrypt the received information. The severe drawback of the private key cryptosystem is that the parties have to safeguard the key transmission from being eavesdropped. A public key cryptosystem invented in the 1970s requires no sharing secret key in advance. Bob publishes a "public key" available to the general public, and Alice uses the public key to encrypt information and sends the encrypted information to Bob. The encryption transformation is specially created such that with only the public key, it is extremely difficult, though not impossible, to invert the encryption transformation. When publishing the public key Bob keeps a matched secret key for easy inversion of the encryption transformation and decryption of the received information. One of the most widely used public key cryptosystems is the RSA cryptosystem, which is named after its creator Rivest, Shamir and Adleman (Menezes, van Oorschot and Vanstone, 1996; Rivest, Shamir and Adleman, 1978). RSA is built on the mathematical asymmetry of factoring: it is easy to multiply large prime numbers and obtain their product as a composite number but hard to find the prime factors of a given large composite number. RSA encryption keeps the large primes as a secret key and uses their product to make a "public key." Because of its exponential complexity, tremendous efforts tried to break the RSA system so far have resulted in vain, and there is a widespread belief that the RSA system is secure against any classical computer based attacks. As the factoring problem can be efficiently solved by Shor's quantum factoring algorithm, a quantum computer can break the RSA system easily. Fortunately, while quantum mechanics takes away with one hand, it gives back with the other. A quantum procedure known as quantum cryptography or quantum key distribution can do key distribution so that the communication security cannot be compromised. The idea is based on the quantum principle that observing a quantum system will disturb the system being observed. If there is an eavesdropper during the transmission of the quantum key between Alice and Bob, eavesdropping will disturb the quantum communication channel that is used to establish the key, and the disturbance will make eavesdropping visible. Alice and Bob will throw away the compromised key and keep only the secured key for their communication.

4.5 Quantum Search Algorithm

Suppose that you would like to find the name corresponding to a given phone number in a telephone directory; or suppose that there are some locations in a given city you would like to visit and wish to find the shortest route passing through all the locations. If there are N names in the telephone directory or N possible routes to pass through all the locations, search algorithms by classical computers usually require operations of order N. One such simple classical algorithm is to check exhaustively all names to find a name matching with the given phone number or to search all possible routes and then find the shortest route among all routes. However, Grover (1996, 1997) developed a quantum search

algorithm that needs only operations of order $\sqrt{N}$ to find a solution to the search problem.

The quantum search algorithm works as follows. Suppose that the search space has N elements and the search problem has exactly M solutions. Assume $M \leq N/2$. (For the silly case of $M > N/2$, we either search for the solution by doing random selection from the search space or double the number of the elements in the search space by adding N extra non-solution elements to the search space.) The algorithm works by creating superposition state with Hadamard gate,

$$|\psi\rangle = \frac{1}{N^{1/2}} \sum_{x=0}^{N-1} |x\rangle,$$

and then applying a so-called Grover iteration (or operator) repeatedly. Set

$$|\phi\rangle = \frac{1}{\sqrt{N-M}} \sum_{x'} |x'\rangle, \quad |\varphi\rangle = \frac{1}{\sqrt{M}} \sum_{x''} |x''\rangle,$$

where the summations over x' and x'' denote sums over all non-solutions and solutions, respectively. Then we can express $|\psi\rangle$ as follows:

$$|\psi\rangle = \sqrt{\frac{N-M}{N}} |\phi\rangle + \sqrt{\frac{M}{N}} |\varphi\rangle.$$

The Grover operator is to perform two reflections, one about the vector $|\phi\rangle$ and another about the vector $|\psi\rangle$. The two reflections together are a rotation with angle θ in the two-dimensional space spanned by $|\phi\rangle$ and $|\varphi\rangle$, where

$$\cos(\theta/2) = \sqrt{\frac{N-M}{N}}.$$

After the rotation, the initial state $|\psi\rangle = \cos(\theta/2)|\phi\rangle + \sin(\theta/2)|\varphi\rangle$ becomes state

$$\cos(3\theta/2)|\phi\rangle + \sin(3\theta/2)|\varphi\rangle.$$

Thus each application of the Grover operator is a rotation with angle θ. The initial state $|\psi\rangle$ has angle $\pi/2 - \theta/2$ with $|\varphi\rangle$; after the first rotation, the resulted state has angle $\pi/2 - 3\theta/2$ with $|\varphi\rangle$; and in general after the rth rotation, the resulted state has angle $\pi/2 - (2r+1)\theta/2$ with $|\varphi\rangle$. Repeatedly applying the Grover operator, we rotate the state vector near $|\varphi\rangle$. With the initial state $|\psi\rangle = \cos(\theta/2)|\phi\rangle + \sin(\theta/2)|\varphi\rangle$, we need to rotate through $\arccos\sqrt{M/N}$ radians to transform the state vector to $|\varphi\rangle$. After $R = \arccos(\sqrt{M/N})/\theta = O(\sqrt{N/M})$ times of applications of the Grover operator, we rotate the state vector $|\psi\rangle$ to within an angle $\theta/2$ of $|\varphi\rangle$. Performing measurements of the state yields a solution to the search problem with probability at least $\cos^2(\theta/2) \geq 1 - M/N$.

The number of iterations R depends on M, the number of solutions. Since $R \leq \pi/(2\theta)$ and $\theta/2 \geq \sin(\theta/2) = \sqrt{M/N}$, $R \leq (\pi/4)\sqrt{N/M}$. Typically, $M \ll N$, $\theta \approx \sin\theta \approx 2\sqrt{M/N}$, thus $R \approx (\pi/4) \cdot \sqrt{N/M}$. We estimate the number of solutions by quantum counting, which is to combine the Grover operator with the phase estimation method. Under the basis $|\phi\rangle$ and $|\varphi\rangle$ the Grover operator has eigenvalues $e^{i\theta}$ and $e^{i(2\pi-\theta)}$. Applying the phase estimation method we can estimate the eigenvalues and thus θ with prescribed precision and probability, which in turn yields M. The combination of the quantum counting and search procedure will find a solution of the search problem with certain probability. Repeating the quantum search algorithm will boost the probability and enable us to obtain a solution to the search problem.

Quantum walk and quantum Markov chain are currently being investigated for new quantum search algorithms and quantum speed-up of Markov chain based algorithms (Aharonov and Ta-Shma, 2003, Childs et al., 2003; Childs, 2010; Tulsi, 2008; Shenvi, Kempe and Whaley, 2003 and Szegedy, 2004). In Section 5 we show that the quantum search algorithm can also be viewed as a quantum simulation procedure.

5. QUANTUM SIMULATION

Quantum simulation is to intentionally and artificially mimic a natural quantum dynamics, which is hard to access, and analyze, by a computer-generated quantum system, which is easy to manipulate and investigate. It provides scientific means for simulating complex biological, chemical or physical systems in order to study and understand certain scientific phenomena and evaluate hard-to-obtain quantities in the systems. Examples in modern scientific studies include the estimation of dielectric constant, proton mass, and precise energy of molecular hydrogen, the study of superconductivity, the test of novel nano-materials, and the design of new biomolecules.

To simulate a quantum system we need to solve the Schrödinger equation (1) which governs the dynamic evolution of the system. For a typical Hamiltonian with real particles the Schrödinger equation usually consists of elliptical differential equations, each of which can be easily simulated by a classical computer. However, the real challenge in simulating a quantum system is to solve the exponential number of such differential equations. For a quantum system of b qubits, its states

have 2^b amplitudes. To simulate the dynamic behavior of b qubits evolving according to the Schrödinger equation, we need to solve a system of 2^b differential equations. Due to the exponential growth in the number of differential equations, the simulation of general quantum systems by classical computers is very inefficient. Classical simulation of quantum systems is feasible for the cases where insightful approximations are available to dramatically reduce the effective number of equations involved. Quantum computers may excel in simulating physically interesting and important quantum systems for which efficient simulation by classical computers may not be available.

5.1 Simulate a Quantum System

The key of quantum simulation is to solve the Schrödinger equation (1) which has solution

$$(14) \qquad |\psi(t)\rangle = e^{-i\mathbf{H}t}|\psi(0)\rangle, \quad i = \sqrt{-1}.$$

Numerical evaluation of $e^{-i\mathbf{H}t}$ is needed. The Hamiltonian $\mathbf{H}$ is usually exponentially large and extremely difficult to exponentiate. The common approach in numerical analysis is to use the first-order linear approximation, $1 - i\mathbf{H}\delta$, of $e^{-i\mathbf{H}(t+\delta)} - e^{-i\mathbf{H}t}$, which often yields unsatisfactory numerical solutions.

Many classes of Hamiltonians have sparse representations. For such sparse Hamiltonians we can find efficient evaluation of the solutions (14) with high-order approximation. For example, the Hamiltonians in many physical systems involve only local interactions, which originate from the fact that most interactions fall off with increasing distance in location or increasing difference in energy. In the local Hamiltonian case, the Hamiltonian of a quantum system with α particles in a d-dimensional space has the form

$$(15) \qquad \mathbf{H} = 2\sum_{\ell=1}^{L} \mathbf{H}_\ell,$$

where L is a polynomial in $\alpha + d$, and each $\mathbf{H}_\ell$ acts on a small subsystem of size free from α and d. For example, the terms $\mathbf{H}_\ell$ are typically two-body interactions and one-body Hamiltonians. Hence $e^{-i\mathbf{H}_\ell\delta}$ are easy to approximate, although $e^{-i\mathbf{H}\delta}$ is very hard to evaluate. Since $\mathbf{H}_\ell$ and $\mathbf{H}_k$ are non-commuting, $e^{-i\mathbf{H}\delta} \neq e^{-i\mathbf{H}_1\delta} \cdots e^{-i\mathbf{H}_L\delta}$. Applying a modification of the Trotter formula (Kato, 1978; Trotter, 1959; Yu, 2001) we obtain

$$e^{-i\mathbf{H}\delta} = \mathbf{U}_\delta + O(\delta^2),$$

where

$$(16) \quad \mathbf{U}_\delta = [e^{-i\mathbf{H}_1\delta} \cdots e^{-i\mathbf{H}_L\delta}][e^{-i\mathbf{H}_L\delta} \cdots e^{-i\mathbf{H}_1\delta}].$$

Thus we can approximate $e^{-i\mathbf{H}\delta}$ by $\mathbf{U}_\delta$ which needs to evaluate only each $e^{-i\mathbf{H}_\ell\delta}$.

Assume that the quantum system starts at $t = 0$ with initial state $|\psi(0)\rangle$ and ends at final time $t = 1$. For an integer m, set $t_j = j/m$, $j = 0, 1, \ldots, m$. The quantum simulation is to apply approximation $\mathbf{U}_\delta$ of $e^{-2i\mathbf{H}\delta}$ to evaluate (14) at t_j iteratively and generate approximate solutions to $|\psi(t_j)\rangle$. Denote by $|\tilde{\psi}(t_j)\rangle$ the state at t_j obtained from the quantum simulation as an approximation of the true state $|\psi(t_j)\rangle$ at t_j. Then for $j = 1, \ldots, m$,

$$(17) \quad \begin{aligned} |\psi(t_j)\rangle &= e^{-2i\mathbf{H}\delta}|\psi(t_{j-1})\rangle = e^{-2i\mathbf{H}j\delta}|\psi(t_0)\rangle, \\ |\tilde{\psi}(t_j)\rangle &= \mathbf{U}_\delta|\tilde{\psi}(t_{j-1})\rangle = \mathbf{U}_\delta^j|\psi(t_0)\rangle. \end{aligned}$$

While classical computers are inefficient in simulating general quantum systems, quantum computers can efficiently carry out the quantum simulation procedure and provide an exponential speedup for the quantum simulation on classical computers. In spite of the inefficiency, classical computers are currently being used to simulate quantum systems in biochemistry and material science. Quantum simulation will be among the important applications of quantum computers. See Abrams and Lloyd (1997), Aspuru-Guzik, Dutoi and Head-Gordon (2005), Bennett et al. (2002), Berry et al. (2007), Freedman, Kitaev and Wang (2002), Jané et al. (2003), Boghosian and Taylor (1998), Lloyd (1996), Nielsen and Chuang (2000), and Zalka (1998).

5.2 Recast Quantum Search Algorithm as Quantum Simulation

Grover's search algorithm discussed in Section 4.5 is an important finding in quantum computation. It can be heuristically sketched as a quantum simulation by writing down an explicit Hamiltonian $\mathbf{H}$ such that a quantum system evolves from its initial state $|\psi\rangle$ to $|x\rangle$ after some specified time, where x is a solution of the search problem. Of course the Hamiltonian $\mathbf{H}$ depends on the initial state $|\psi\rangle$ and solution x. Suppose that $|y\rangle$ is another state such that $|x\rangle$ and $|y\rangle$ form an computational basis, and $|\psi\rangle = \alpha|x\rangle + \beta|y\rangle$ for real α and β with $\alpha^2 + \beta^2 = 1$. Define Hamiltonian

$$\mathbf{H} = |x\rangle\langle x| + |\psi\rangle\langle\psi| = \mathbf{I} + \alpha(\beta\boldsymbol{\sigma}_x + \alpha\boldsymbol{\sigma}_z),$$

where $\boldsymbol{\sigma}_x$ and $\boldsymbol{\sigma}_z$ are Pauli matrices defined in (7). Then

$$\begin{aligned} &\exp(-i\mathbf{H}t)|\psi\rangle \\ &= e^{-it}[\cos(\alpha t)|\psi\rangle - i\sin(\alpha t)(\beta\boldsymbol{\sigma}_x + \alpha\boldsymbol{\sigma}_z)|\psi\rangle] \\ &= e^{-it}[\cos(\alpha t)|\psi\rangle - i\sin(\alpha t)|x\rangle]. \end{aligned}$$

Measuring the system at time $t = \pi/(2\alpha)$ yields the solution state $|x\rangle$.

5.3 Quantum Monte Carlo Simulation

Quantum theory is intrinsically stochastic and quantum measurement outcome is random. As many naturally occurring quantum systems involve a large number of interacting particles, due to the computational complexity we are forced to utilize Monte Carlo techniques in the simulations of such quantum systems. The combination of Monte Carlo methods with quantum simulation makes it possible to obtain reliable quantifications of quantum phenomena and estimates of quantum quantities. Such combination procedures are often referred to as quantum Monte Carlo simulation (Nightingale and Umrigar, 1999; Rousseau, 2008). Consider the problem of estimating the following quantity:

$$(18) \qquad \theta = \mathrm{Tr}(\mathbf{X}\boldsymbol{\rho}) = E(X),$$

where $\mathbf{X}$ is an observable, X is its measurement result, and $\boldsymbol{\rho}$ is the state of the quantum system under which we perform the measurements and evaluate the quantity θ.

As $\boldsymbol{\rho}$ is the true final state of the quantum system, we denote by $\tilde{\boldsymbol{\rho}}$ the final state of the quantum system obtained via quantum simulation. The quantum systems are prepared in initial state $\boldsymbol{\rho}_0$, and we use the quantum simulation procedure described above to simulate the evolutions of the systems from initial state $\boldsymbol{\rho}_0$ to final state $\tilde{\boldsymbol{\rho}}$ according to Schrödinger's equation (14) with Hamiltonian H given by (15). We repeatedly perform the measurements of such n identically simulated quantum systems at the final state and obtain measurement results $X_1, \ldots, X_n$. We estimate θ defined in (18) by

$$(19) \qquad \hat{\theta} = \frac{1}{n} \sum_{j=1}^{n} X_j.$$

The target θ given by (18) is defined under the true state $\boldsymbol{\rho}$, while the simulated quantum system is under approximate final state $\tilde{\boldsymbol{\rho}}$ which is close to $\boldsymbol{\rho}$. The measurement results $X_1, \ldots, X_n$ are obtained via quantum simulation from the quantum systems in the simulated state $\tilde{\boldsymbol{\rho}}$. Therefore, the Monte Carlo quantum estimator $\hat{\theta}$ in (19) involves both bias and variance. Wang (2011) studied the quantum simulation procedure and investigated the bias and variance of $\hat{\theta}$. The derived bias and variance results can be used to design optimal strategy for the best utilization of computational resources to obtain the quantum Monte Carlo estimator.

6. QUANTUM INFORMATION

Classical information theory is centered on Shannon's two coding theorems on noiseless and noisy channels. The noiseless channel coding theorem quantifies the number of classical bits required to store information for transmission by Shannon entropy, while the noisy channel coding theorem quantifies the amount of information that can be reliably transmitted through a noisy channel by an error-correction coding scheme. The quantum analogs of Shannon entropy and Shannon noiseless coding theorem are von Neumann entropy and Schumacher's noiseless channel coding theorem, respectively. The von Neumann entropy is defined to be $S(\boldsymbol{\rho}) = -\mathrm{tr}(\boldsymbol{\rho} \log \boldsymbol{\rho})$. Schumacher's noiseless channel coding theorem quantifies quantum resources required to compress quantum states by von Neumann entropy (Schumacher, 1995). Analogous to Shannon's noisy channel coding theorem, a theorem known as Holevo–Schumacher–Westmoreland theorem can be used to compute the product quantum state capacity for some noisy channels (Holevo, 1998; Schumacher and Westmoreland, 1997). However, communications over noisy quantum channels are much less understood than the classical counterpart. It is an unsolved problem to determine quantum channel capacity or the amount of quantum information that can be reliably transmitted over noisy quantum channels. See Hayashi (2006) and Nielsen and Chuang (2000).

In spite of the above similarity, there are intrinsic differences between classical information and quantum information. Classical information can be distinguished and copied. For example, we can identify different letters and produce an identical version of a digital image for back-up. However, quantum mechanics does not allow unknown quantum states to be distinguished or copied exactly. For example, we cannot reliably distinguish between quantum states $|0\rangle$ and $(|0\rangle + |1\rangle)/\sqrt{2}$. If we perform measurement for quantum state $|0\rangle$, the measurement result will be 0 with probability 1, while measuring quantum state $(|0\rangle + |1\rangle)/\sqrt{2}$ yields measurements 0 or 1 with equal probability. A measurement result of 0 cannot tell the identity of the quantum state being measured. A theorem known as a no-cloning theorem states that unknown quantum states cannot be copied exactly (Wootters and Zurek, 1982; Nielsen and Chuang, 2000).

As we discussed in Section 3.3, quantum entanglement plays a crucial role in strange quantum effects such as quantum teleportation, violation of Bell's inequality, and superdense coding (Hayashi, 2006;

Nielsen and Chuang, 2000). Entanglement is a new type of resource that differs vastly from the traditional resources in classical information theory. We are far from having a general theory to understand quantum entanglement but encouraging progress made so far reveals the amazing property and intriguing structure of entangled states and remarkable connections between noisy quantum channels and entanglement transformation. Consider quantum error-correction for reliable quantum computation and quantum information processing. Quantum error-correction is employed in quantum computation and quantum communication to protect quantum information from loss due to quantum noise and other errors like faulty quantum gates. Classical information uses redundancy to achieve error-correction, but the no-cloning theorem presents an obstacle to copying quantum information and formulating a theory of quantum error-correction based on simple redundancy. Again quantum entanglement comes to the rescue. It is forbidden to copy qubits but we can spread the information of one qubit onto a highly entangled state of several qubits. Shor (1995) first discovered the method of formulating a quantum error-correction code by storing the information of one qubit onto a highly entangled state of nine qubits. Over time several quantum error-correction codes are proposed (Calderbank and Shor, 1996; Cory et al., 1998; Steane, 1996a, 1996b). These quantum error-correction codes can protect quantum information against quantum noise, and thus quantum noise likely poses no fundamental barrier to the performance of large-scale quantum computing and quantum information processing.

Here is how quantum error-correction codes work. We consider the single qubit case. First assume that a qubit $\alpha_0|0\rangle + \alpha_1|1\rangle$ is passed through a bit flip channel which flips the state of a qubit from $|0\rangle$ to $|1\rangle$ and from $|1\rangle$ to $|0\rangle$, each with probability p, and leaves each of states $|0\rangle$ and $|1\rangle$ untouched with probability $1 - p$. We describe a bit flip code that protects the qubit against quantum noise from the bit flip channel.

We encode states $|0\rangle$ and $|1\rangle$ in three qubits, with $|0\rangle$ encoded as $|000\rangle$ and $|1\rangle$ as $|111\rangle$. Thus the qubit state $\alpha_0|0\rangle + \alpha_1|1\rangle$ is encoded in three qubits as $\alpha_0|000\rangle + \alpha_1|111\rangle$. We pass each of the three qubits through an independent copy of the bit flip channel, and assume that at most one qubit is flipped. The following simple two-step error-correction procedure can be used to recover the correct quantum state.

Step 1. Perform a measurement on a specially constructed observable and call the measurement result

an error syndrome. The error syndrome can inform us what error, if any, occurred on the quantum state. The error syndrome can inform us what error, if any, occurred on the quantum state. The observable has eigenvalues 0, 1, 2 and 3, with corresponding projection operators,

$$\mathbf{Q}_0 = |000\rangle\langle000| + |111\rangle\langle111| \quad \text{no error,}$$

$$\mathbf{Q}_1 = |100\rangle\langle100| + |011\rangle\langle011|$$

bit flip on the first qubit,

$$\mathbf{Q}_2 = |010\rangle\langle010| + |101\rangle\langle101|$$

bit flip on the second qubit,

$$\mathbf{Q}_3 = |001\rangle\langle001| + |110\rangle\langle110|$$

bit flip on the third qubit.

If one of three qubits has one or no bit flip, the error syndrome will be one of 0, 1, 2 and 3, with 0 corresponding to no flip, and 1, 2 and 3 to a bit flip on the first, second and third qubit, respectively. For example, if the first qubit is flipped, the corrupted state is $|\psi\rangle = \alpha_0|100\rangle + \alpha_1|011\rangle$. Since $\langle\psi|\mathbf{Q}_1|\psi\rangle = 1$ and $\langle\psi|\mathbf{Q}_j|\psi\rangle = 0$ for $j \neq 1$, in this case the error syndrome is 1. Although performing measurements usually causes change to the quantum state, the speciality of the constructed observable is that syndrome measurement does not perturb the quantum state: it is easy to check that the state is $|\psi\rangle$ both before and after the syndrome measurement. While the syndrome provides information about what flip error has occurred, it does not contain any information about the state being protected, that is, it does not allow us to deduce anything about the amplitudes α_0 and α_1. Such a special property is the generic feature of syndrome measurement.

Step 2. The error type supplied by the error syndrome can inform us what procedure to use to recover the original state. For example, error syndrome 1 indicates a bit flip on the first qubit, and a flip on the first qubit again will perfectly recover the original state $\alpha_0|000\rangle + \alpha_1|111\rangle$. The error syndrome 0 implies no error and doing nothing, and error syndromes 1, 2 and 3 correspond to a bit flip again on the first, second, third qubit, respectively. The procedure will recover the original state with perfect accuracy, if there is at most one bit flip in the encoded three qubits. The probability that more than one bit flipped is $p^3 + 3p^2(1 - p) = 3p^2 - 2p^3$, which is much smaller than the error probability p of making no-correction for the typical bit flip channel. Thus the encoding and decoding scheme makes the storage and transmission of the qubit more reliable.

Next we consider a more interesting noisy quantum channel: a phase flip channel which, with probability

p, changes a qubit state $\alpha_0|0\rangle + \alpha_1|1\rangle$ to $\alpha_0|0\rangle - \alpha_1|1\rangle$, and with probability $1 - p$, leaves alone the qubit. The following scheme is to turn the phase flip channel into a bit flip channel. Let $|+\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$ and $|-\rangle = (|0\rangle - |1\rangle)/\sqrt{2}$ be a qubit basis. The phase flip channel leaves alone states $|+\rangle$ and $|-\rangle$ with probability $1 - p$ and changes $|+\rangle$ to $|-\rangle$ and vice versa with probability p. In other words, the phase flip channel with respect to the basis $|+\rangle$ and $|-\rangle$ acts just like a bit flip channel with respect to the basis $|0\rangle$ and $|1\rangle$. Thus we encode $|0\rangle$ as $|+++\rangle$ and $|1\rangle$ as $|---\rangle$ for protection against phase flip errors. The operations for encoding, error-detection and recovery are the same as for the bit flip channel but with respect to the $|+\rangle$ and $|-\rangle$ basis instead of the $|0\rangle$ and $|1\rangle$ basis.

Last we describe Shor error-correction code. It is a combination of the three-qubit phase flip and bit flip codes. First use the phase flip code to encode states $|0\rangle$ and $|1\rangle$ in three qubits, with $|0\rangle$ encoded as $|+++\rangle$ and $|1\rangle$ as $|---\rangle$; next, use the three-qubit bit flip code to encode each of these qubits, with $|+\rangle$ encoded as $(|000\rangle + |111\rangle)/\sqrt{2}$ and $|-\rangle$ encoded as $(|000\rangle - |111\rangle)/\sqrt{2}$. The resulted nine-qubit code has codeworks as follows:

$$|0\rangle \rightarrow \frac{|000\rangle + |111\rangle}{\sqrt{2}} \frac{|000\rangle + |111\rangle}{\sqrt{2}} \frac{|000\rangle + |111\rangle}{\sqrt{2}},$$

$$|1\rangle \rightarrow \frac{|000\rangle - |111\rangle}{\sqrt{2}} \frac{|000\rangle - |111\rangle}{\sqrt{2}} \frac{|000\rangle - |111\rangle}{\sqrt{2}}.$$

With the mixture of both phase flip and bit flip codes, the Shor error-correction code can protect against bit flip errors, phase flip errors, as well as a combined bit and phase flip errors on any single qubit. In fact it has been shown that this simple quantum error-correction code can protect against the effects of any completely arbitrary errors on a single qubit (Shor, 1995).

7. CONCLUDING REMARKS

Quantum information science gains enormous attention in computer science, mathematics, physical sciences and engineering, and several interdisciplinary subfields are developing under the umbrella of quantum information. This paper reviews quantum computation and quantum information from a statistical perspective. We introduce concepts like qubits, quantum gates and quantum circuits in quantum computation and discuss quantum entanglement, quantum parallelism and quantum error-correction in quantum computation and quantum information. We present major quantum algorithms and show their advantages over the available classical algorithms. We illustrate quantum simulation procedure and Monte Carlo methods in quantum simulation. As classical computation and simulation are ubiquitous nowadays in statistics, we expect quantum computation and quantum simulation will have a paramount role to play in modern statistics. This paper exposes the topics to statisticians and encourages more statisticians to work in the fields. There are many statistical issues in theoretical research as well as experimental work in quantum computation, quantum simulation and quantum information. For example, as measurement data collected in quantum experiments require more and more sophisticated statistical methods for better estimation, simulation and understanding, it is imperative to develop good quantum statistics methods and quantum simulation procedures and study interrelationship and mutual impact between quantum estimation and quantum simulation. Since quantum computation is intrinsically random, and quantum simulation employs Monte Carlo techniques, as we point out in Section 4.3 and Wang (2011), it is important to provide sound statistical methods for analyzing quantum algorithms and quantum simulation in general and study high-order approximations to exponentiate Hamiltonians and the efficiency of the resulted quantum simulation procedures in particular. On the other hand, quantum computation and quantum simulation have great potential to revolutionize computational statistics. Below are a few cases in point.

1. The "random numbers" generated by classical computers are pseudo-random numbers in the sense that they are produced by deterministic procedures and can be exactly repeated and perfectly predicted given the deterministic schemes and the initial seeds. On the contrary, superposition states enable quantum computers to produce genuine random numbers. For example, measuring $(|0\rangle + |1\rangle)/\sqrt{2}$ yields 0 and 1 with equal probability. In general we generate b-bit binary random numbers $x = x_1 \cdots x_b$, $x_j = 0, 1$ as follows. Apply b Hadamard gates to b qubits of $|0 \cdots 0\rangle$ to obtain

$$\frac{|0\rangle + |1\rangle}{\sqrt{2}} \cdots \frac{|0\rangle + |1\rangle}{\sqrt{2}} = \frac{1}{\sqrt{2^b}} \sum_x |x\rangle,$$
$$x = x_1 \cdots x_b, x_j = 0, 1,$$

where the sum is over all possible 2^b values of x, and then measure the obtained qubits and yield b-bit binary random numbers $x = x_1 \cdots x_b$ with equal probability. Quantum theory guarantees that such

random numbers are genuinely random. Thus quantum computers are able to generate genuine random numbers and perform true Monte Carlo simulation. It is exciting to design general quantum random number generator and study quantum Monte Carlo simulation. Perhaps we may need to re-examine Monte Carlo simulation studies conducted by classical computers.

2. It is interesting to investigate the potential of quantum computation and quantum simulation for computational statistics. We expect that quantum computers may be much faster than classical computers for computing some statistical problems. Moreover, quantum computers may be able to carry out some computational statistical tasks that are prohibitive by classical computers. Specific examples are as follows: (a) We may use the basic ideas of Grover's search algorithm to develop fast quantum algorithms for implementing some statistical procedures. For example, finding the median of a huge data set is to search for a numerical value that separates the top and bottom halves of the data, and quantum algorithms can offer quadratical speedup for calculating median and trimmed mean. (b) With genuine random number generator and faster mean evaluation, quantum computers may offer significant advantages over classical computers for Monte Carlo integration. For example, Monte Carlo integration in high dimensions may be exponentially or quadratically faster on quantum computers than on classical computers. (c) It might be possible for quantum computers to carry out some prohibiting statistical computing tasks like the Bayesian computation discussed in Section 3. Some preliminary research along these lines may be found in Nayak and Wu (1999) and Heinrich (2003).

3. As quantum computation and quantum simulation are ideal for simulating interacting particle systems like the Ising model, it is fascinating to explore the interplay between quantum simulation and Markov chain Monte Carlo methodology and the quantum potential to speed up Markov chain based algorithms. In fact, it has been shown that quantum walk based algorithms can offer quadratical speedup for certain Markov chain based algorithms (Magniez et al., 2011; Richter, 2007; Szegedy, 2004, and Wocjan and Abeyesinghe, 2008).

Finally we point out that quantum computers are wonderful but it is difficult to build quantum computers with present technology. To build a quantum computer the physical apparatus must satisfy requirements that the quantum system realized qubits needs to be well isolated in order to retain its quantum properties and at the same time the quantum system has to be accessible so that the qubits can be manipulated to perform computations and measure output results. The two opposing requirements are determined by the strength of coupling of the quantum system to the external entities. The coupling causes quantum decoherence. Decoherence refers to the loss of coherence between the components of a quantum system or quantum superposition from the interaction of the quantum system with its environment. It is very crucial but challenging to control a quantum system of qubits and correct the effects of decoherence in quantum computation and quantum information. Quantum computing has witnessed great advances in recent years, and quantum computers of a handful of qubits and basic quantum communication devices have been built in research laboratories (see Barz et al., 2012; Clarke and Wilhelm, 2008; DiCarlo et al., 2009; Johnson et al., 2011; Lee et al., 2011; Neumann et al., 2008), but there are technological hurdles in the development of a quantum computer of large capacity. History shows that scientific innovations and technological surprises are a never-ending saga. It is anticipated that quantum computers with a few dozen of qubits will be built in near future. As we have discussed in Section 3.1, such a quantum computer has capacity of a classical supercomputer. We are very optimistic that someday quantum computers will be available for statisticians to crunch numbers. For the time being, instead of waiting in the sidelines for that to happen, statisticians should get into the field of play. It is time for us to dive into this frontier research and work with scientists and engineers to speed up the arrival of practical quantum computers. As a last note, in 2011 a Canadian company called D-Wave has sold the claimed first commercial quantum computer of 128 qubits to the Lockheed-Martin corporation, despite the D-Wave's quantum system being criticized as a black box. Large scale quantum computers may be years away, but quantum computing is already here as a scientific endeavor to provoke deep thoughts and integrate profound questions in physics and computer science.

ACKNOWLEDGMENTS

REFERENCES

ABRAMS, D. S. and LLOYD, S. (1997). Simulation of many-body Fermi systems on a quantum computer. *Phys. Rev. Lett.* **79** 2586–2589.

AHARONOV, D. and TA-SHMA, A. (2003). Adiabatic quantum state generation and statistical zero knowledge. In *Proceedings of the Thirty-Fifth Annual ACM Symposium on Theory of Computing* 20–29. ACM, New York (electronic). MR2121066

ARTILES, L. M., GILL, R. D. and GUŢĂ, M. I. (2005). An invitation to quantum tomography. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **67** 109–134. MR2136642

ASPECT, A., GRANGIER, P. and ROGER, G. (1981). Experimental tests of realistic local theories via Bell's theorem. *Phys. Rev. Lett.* **47** 460–463.

ASPECT, A., GRANGIER, P. and ROGER, G. (1982a). Experimental realization of Einstein–Podolsky–Rosen–Bohm Gedankenexperiment: A new violation of Bell's inequalities. *Phys. Rev. Lett.* **49** 91–94.

ASPECT, A., GRANGIER, P. and ROGER, G. (1982b). Experimental test of Bell's inequalities using time-varying analyzers. *Phys. Rev. Lett.* **49** 1804–1807.

ASPURU-GUZIK, A. D., DUTOI, P. J. L. and HEAD-GORDON, M. (2005). Simulated quantum computation of molecular energies. *Science* **309** 1704.

BARNDORFF-NIELSEN, O. E., GILL, R. D. and JUPP, P. E. (2003). On quantum statistical inference. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **65** 775–816. MR2017871

BARZ, S., KASHEFI, E., BROADBENT, A., FITZSIMONS, J.F., ZEILINGER, A. and WALTHER, P. (2012). Demonstration of blind quantum computing. *Science* **335** 303–308. MR2919052

BELL, J. (1964). On the Einstein Podolsky Rosen paradox. *Physics* **1** 195–200.

BENNETT, C. H., CIRAC, J. I., LEIFER, M. S., LEUNG, D. W., LINDEN, N., POPESCU, S. and VIDAL, G. (2002). Optimal simulation of two-qubit Hamiltonians using general local operations. *Phys. Rev. A (3)* **66** 012305. MR1929513

BERRY, D. W., AHOKAS, G., CLEVE, R. and SANDERS, B. C. (2007). Efficient quantum algorithms for simulating sparse Hamiltonians. *Comm. Math. Phys.* **270** 359–371. MR2276450

BOGHOSIAN, B. M. and TAYLOR, W. IV (1998). Simulating quantum mechanics on a quantum computer. *Phys. D* **120** 30–42. MR1679863

BOHM, D. (1951). *Quantum Theory*. Prentice-Hall, Englewood Cliffs, NJ.

BUTUCEA, C., GUŢĂ, M. and ARTILES, L. (2007). Minimax and adaptive estimation of the Wigner function in quantum homodyne tomography with noisy data. *Ann. Statist.* **35** 465–494. MR2336856

CALDERBANK, A. R. and SHOR, P. W. (1996). Good quantum error-correcting codes exist. *Phys. Rev. A* **54** 1098–1105.

CHILDS, A. M. (2010). On the relationship between continuous- and discrete-time quantum walk. *Comm. Math. Phys.* **294** 581–603. MR2579466

CHILDS, A. M., CLEVE, R., DEOTTO, E., FARHI, E., GUTMANN, S. and SPIELMAN, D. A. (2003). Exponential algorithmic speedup by quantum walk. In *Proc. 35th ACM Symposium on Theory of Computing* 59–68. ACM Press, New York.

CLARKE, J. and WILHELM, F. K. (2008). Superconducting quantum bits. *Nature* **453** 1031–1042.

CLAUSER, J. F., HORNE, M. A., SHIMONY, A. and HOLT, R. A. (1969). Proposed experiment to test local hidden-variable theories. *Phys. Rev. Lett.* **23** 880–884.

CORY, D. G., MASS, W., PRICE, M., KNILL, E., LAFLAMME, R., ZUREK, W. H., HAVEL, T. F. and SOMAROO, S. S. (1998). Experimental quantum error correction. *Phys. Rev. Lett.* **81** 2152–2155.

CRANDALL, R. and POMERANCE, C. (2001). *Prime Numbers: A Computational Perspective*. Springer, New York. MR1821158

DEUTSCH, D. (1985). Quantum theory, the Church–Turing principle and the universal quantum computer. *Proc. R. Soc. Lond. Ser. A* **400** 97–117. MR0801665

DICARLO, L., CHOW, J. M., GAMBETTA, J. M., BISHOP, L. S., JOHNSON, B. R., SCHUSTER, D. I., MAJER, J., BLAIS, A., FRUNZIO, L., GIRVIN, S. M. and SCHOELKOPF, R. J. (2009). Demonstration of two-qubit algorithms with a superconducting quantum processor. *Nature* **460** 240–244.

DIVINCENZO, D. P. (1995). Quantum computation. *Science* **270** 255–261. MR1355956

EINSTEIN, A., PODOLSKY, B. and ROSEN, N. (1935). Can quantum-mechanical description of physical reality be considered complete? *Phys. Rev.* **47** 777–780.

FEYNMAN, R. P. (1981/82). Simulating physics with computers. *Internat. J. Theoret. Phys.* **21** 467–488. MR0658311

FREEDMAN, M. H., KITAEV, A. and WANG, Z. (2002). Simulation of topological field theories by quantum computers. *Comm. Math. Phys.* **227** 587–603. MR1910832

GRIFFITHS, D. J. (2004). *Introduction to Quantum Mechanics*, 2nd ed. Benjamin Cummings, San Francisco, CA.

GROVER, L. K. (1996). A fast quantum mechanical algorithm for database search. In *Proceedings of the Twenty-Eighth Annual ACM Symposium on the Theory of Computing (Philadelphia, PA, 1996)* 212–219. ACM, New York. MR1427516

GROVER, L. K. (1997). Quantum mechanics helps in searching for a needle in a haystack. *Phys. Rev. Lett.* **79** 325–328.

HAYASHI, M. (2006). *Quantum Information: An Introduction*. Springer, Berlin. Translated from the 2003 Japanese original. MR2228302

HEINRICH, S. (2003). From Monte Carlo to quantum computation. *Math. Comput. Simulation* **62** 219–230. MR1988372

HOLEVO, A. S. (1982). *Probabilistic and Statistical Aspects of Quantum Theory. North-Holland Series in Statistics and Probability* **1**. North-Holland, Amsterdam. Translated from the Russian by the author. MR0681693

HOLEVO, A. S. (1998). The capacity of the quantum channel with general signal states. *IEEE Trans. Inform. Theory* **44** 269–273. MR1486663

HUBER, P. J. and RONCHETTI, E. M. (2009). *Robust Statistics*, 2nd ed. Wiley, Hoboken, NJ. MR2488795

JANÉ, E., VIDAL, G., DÜR, W., ZOLLER, P. and CIRAC, J. I. (2003). Simulation of quantum dynamics with quantum optical systems. *Quantum Inf. Comput.* **3** 15–37. MR1965173

JOHNSON, M. W., AMIN, M. H. S., GILDERT, S., LANTING, T., HAMZE, F., DICKSON, N., HARRIS, R., BERKLEY, A. J., JOHANSSON, J., BUNYK, P., CHAPPLE, E. M., ENDERUD, C., HILTON, J. P., KARIMI, K., LADIZINSKY, E., LADIZINSKY, N., OH, T., PERMINOV, I., RICH, C., THOM, M. C., TOLKACHEVA, E., TRUNCIK, C. J. S., UCHAIKIN, S., WANG, J., WILSON, B., ROSE, G. et al. (2011). Quantum annealing with manufactured spins. *Nature* **473** 194–198.

KATO, T. (1978). Trotter's product formula for an arbitrary pair of self-adjoint contraction semigroups. In *Topics in Functional Analysis* (*Essays Dedicated to M. G. Kreĭn on the Occasion of His* 70*th Birthday*). *Adv. in Math. Suppl. Stud.* **3** 185–195. Academic Press, New York. MR0538020

KIEFER, C. (2004). On the interpretation of quantum theory—from Copenhagen to the present day. In *Time*, *Quantum and Information* 291–299. Springer, Berlin. MR2180154

LEE, N., BENICHI, H., TAKENO, Y., TAKEDA, S., WEBB, J., HUNTINGTON, E. and FURUSAWA, A. (2011). Teleportation of nonclassical wave packets of light. *Science* **332** 330–333.

LLOYD, S. (1996). Universal quantum simulators. *Science* **273** 1073–1078. MR1407944

MAGNIEZ, F., NAYAK, A., ROLAND, J. and SANTHA, M. (2011). Search via quantum walk. *SIAM J. Comput.* **40** 142–164. MR2783206

MARIANTONI, M., WANG, H., YAMAMOTO, T., NEELEY, M., BIALCZAK1, R. C., CHEN, Y., LENANDER, M., LUCERO, E., O'CONNELL, A. D., SANK, D., WEIDES, M., WENNER, J., YIN, Y., ZHAO, J., KOROTKOV, A. N., CLELAND, A. N. and MARTINIS, J. M. (2011). Implementing the quantum von Neumann architecture with superconducting circuits. *Science* **7** 1208517.

MENEZES, A., VAN OORSCHOT, P. C. and VANSTONE, S. A. (1996). *Handbook of Applied Cryptography*. CRC Press, New York.

NAYAK, A. and WU, F. (1999). The quantum query complexity of approximating the median and related statistics. In *Annual ACM Symposium on Theory of Computing* (*Atlanta*, *GA*, 1999) 384–393. ACM, New York (electronic). MR1798059

NEUMANN, P., MIZUOCHI, N., REMPP, F., HEMMER, P., WATANABE, H., YAMASAKI, S., JACQUES, V., GAEBEL, T., JELEZKO, F. and WRACHTRUP, J. (2008). Multipartite entanglement among single spins in diamond. *Science* **320** 1326–1329.

NIELSEN, M. A. and CHUANG, I. L. (2000). *Quantum Computation and Quantum Information*. Cambridge Univ. Press, Cambridge. MR1796805

NIGHTINGALE, M. P. and UMRIGAR, C. J., eds. (1999). *Quantum Monte Carlo Methods in Physics and Chemistry*. *NATO Science Series C*: *Mathematical and Physical Sciences* **525**. Kluwer Academic, Dordrecht. MR1712250

NUSSBAUM, M. and SZKOŁA, A. (2009). The Chernoff lower bound for symmetric quantum hypothesis testing. *Ann. Statist.* **37** 1040–1057. MR2502660

PARTHASARATHY, K. R. (1992). *An Introduction to Quantum Stochastic Calculus. Monographs in Mathematics* **85**. Birkhäuser, Basel. MR1164866

RICHTER, P. C. (2007). Quantum speedup of classical mixing processes. *Phys. Rev. A* **76** 042306.

RIVEST, R. L., SHAMIR, A. and ADLEMAN, L. (1978). A method for obtaining digital signatures and public-key cryptosystems. *Comm. ACM* **21** 120–126. MR0700103

ROUSSEAU, V. G. (2008). Stochastic Green function algorithm. *Phys. Rev. E* (3) **77** 056705. MR2495497

SAKURAI, J. J. and NAPOLITANO, J. (2010). *Modern Quantum Mechanics*, 2nd ed. Addison-Wesley, Reading, MA.

SAYRIN, C., DOTSENKO, I., ZHOU, X., PEAUDECERF, B., RYBARCZYK, T., GLEYZES, S., ROUCHON, P., MIRRAHIMI, M., AMINI, H., BRUNE, M., RAIMOND, J. M. and HAROCHE, S. (2011). Real-time quantum feedback prepares and stabilizes photon number states. *Nature* **477** 10376.

SCHUMACHER, B. (1995). Quantum coding. *Phys. Rev. A* (3) **51** 2738–2747. MR1328824

SCHUMACHER, B. and WESTMORELAND, M. D. (1997). Sending classical information via noisy quantum channels. *Phys. Rev. A* **56** 131138.

SHANKAR, R. (1994). *Principles of Quantum Mechanics*, 2nd ed. Plenum Press, New York. MR1343488

SHENVI, N., KEMPE, J. and WHALEY, K. B. (2003). Quantum random-walk search algorithm. *Phys. Rev. A* **67** 052307.

SHOR, P. W. (1994). Algorithms for quantum computation: Discrete logarithms and factoring. In 35*th Annual Symposium on Foundations of Computer Science* (*Santa Fe*, *NM*, 1994) 124–134. IEEE Comput. Soc. Press, Los Alamitos, CA. MR1489242

SHOR, P. W. (1995). Scheme for reducing decoherence in quantum computer memory. *Phys. Rev. A* **52** 2493–2496.

SHOR, P. W. (1997). Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer. *SIAM J. Comput.* **26** 1484–1509. MR1471990

STEANE, A. M. (1996a). Error correcting codes in quantum theory. *Phys. Rev. Lett.* **77** 793–797. MR1398854

STEANE, A. (1996b). Multiple-particle interference and quantum error correction. *Proc. R. Soc. Lond. Ser. A* **452** 2551–2577. MR1421749

SZEGEDY, M. (2004). Quantum speed-up of Markov chain based algorithms. In *Proc. 45th IEEE Symposium on Foundations of Computer Science* 32–41. IEEE Computer Society Press, Los Alamitos, CA.

TROTTER, H. F. (1959). On the product of semi-groups of operators. *Proc. Amer. Math. Soc.* **10** 545–551. MR0108732

TSIRELSON, B. S. (1980). Quantum generalizations of Bell's inequality. *Lett. Math. Phys.* **4** 93–100. MR0577178

TULSI, A. (2008). Faster quantum-walk algorithm for the two-dimensional spatial search. *Phys. Rev. A* **78** 012310.

VIDAKOVIC, B. (1999). *Statistical Modeling by Wavelets*. Wiley, New York. MR1681904

VON NEUMANN, J. (1955). *Mathematical Foundations of Quantum Mechanics*. Princeton Univ. Press, Princeton, NJ. Translated by Robert T. Beyer. MR0066944

WANG, Y. Z. (1994). Quantum Gaussian processes. *Acta Math. Appl. Sin.* (*Engl. Ser.*) **10** 315–327. MR1310171

WANG, Y. (2011). Quantum Monte Carlo simulation. *Ann. Appl. Stat.* **5** 669–683. MR2840170

WOCJAN, P. and ABEYESINGHE, A. (2008). Speed-up via quantum sampling. arXiv:0804.4259v3[quant-ph].

WOOTTERS, W. K. and ZUREK, W. H. (1982). A single quantum cannot be cloned. *Nature* **299** 802–803.

YU, A. B. (2001). Campbell-Hausdorff formula. In *Encyclopaedia of Mathematics* (M. Hazewinkel, ed.). Kluwer Academic, Dordrecht.

ZALKA, C. (1998). Simulating a quantum systems on a quantum computer. *R. Soc. Lond. Philos. Trans. Ser. A Math. Phys. Eng. Sci.* **454** 313–322.