Editor’s note:
As an emerging technology, quantum computing brings unique promises in creating a fundamentally different paradigm of information processing. This article reviews introductory concepts and building blocks—quantum logic gates and memory, i.e., qubits, for quantum computing and depicts the challenges as well as prospects associated with design, test, and fabrication of quantum devices.
—Swarup Bhunia, Case Western Reserve University

Quantum computing promises new capabilities for processing information and performing computationally hard tasks. This includes significant algorithmic advances for solving hard problems in computing [1], sensing [2], and communication [3]. The breakthrough examples of Shor’s algorithm for factoring numbers and Grover’s algorithms for unstructured search have fueled a series of more recent advances in computational chemistry, nuclear physics, and optimization research among many others. However, realizing the algorithmic advantages of quantum computing requires hardware devices capable of encoding quantum information, performing quantum logic, and carrying out sequences of complex calculations based on quantum mechanics [4]. For more than 35 years, there has been a broad array of experimental efforts to build quantum computing devices to demonstrate these new ideas. Multiple state-of-the-art engineering efforts have now fabricated functioning quantum processing units (QPUs) capable of carrying out small-scale demonstrations of quantum computing. The QPUs developed by commercial vendors such as IBM, Google, D-Wave, Rigetti, and IonQ are among a growing list of devices that have demonstrated the fundamental elements required for quantum computing [5]. This progress in prototype QPUs has opened up new discussions about how to best utilize these nascent devices [6]–[8].

Quantum computing poses several new challenges to the concepts of design and testing that are unfamiliar to conventional CMOS-based computing devices. For example, a striking fundamental challenge is the inability to interrogate the instantaneous quantum state of these new devices. Such interrogations may be impractically complex within the context of conventional computing, but they are physically impossible within the context of quantum computing due to the no-cloning principles. This physical distinction fundamentally changes how QPUs are designed and their operation tested relative to past practice. This tutorial provides an overview of the principles of operation behind quantum computing devices as well as a summary of the state of the art in QPU development. The continuing development of quantum computing will require expertise from the conventional design and testing community to ensure the integration of these nontraditional devices into existing design workflows and testing infrastructure. There is a wide variety of technologies under consideration for device development, and this article focuses on the current workflows surrounding quantum devices fabricated in semiconducting, superconducting, and...
trapped-ion technologies. We also discuss the design of logical circuits that quantum devices must execute to perform computational work.

While the tutorial captures many of the introductory topics needed to understand the design and testing of quantum devices, several more advanced topics have been omitted due to space constraints. Foremost is the broader theory of quantum computation, which has developed rapidly from early models of quantum Turing machines to a number of different but equally powerful computational models. In addition, we have largely omitted the sophisticated techniques employed to mitigate the occurrence of errors in quantum devices. Quantum error correction is an important aspect of long-term and large-scale quantum computing, which uses redundancy to overcome the loss in information from noisy environments. Finally, our review of quantum computing technologies is intentionally narrowed to three of the leading candidates capable of near-term experimental demonstrations. However, there is a great diversity of experimental quantum physical systems that can be used for encoding and processing quantum information.

Principles of quantum computing

The principles of quantum computing derive from quantum mechanics, a theoretical framework that has accurately modeled the microscopic world for more than 100 years. Quantum computing draws its breakthroughs in computational capabilities from the many unconventional features inherent to quantum mechanics, and we provide a brief overview of these features while others offer more exhaustive explanations [4].

In quantum mechanics, all knowable information about a physical system is represented by a quantum state. A prominent example of a quantum state within the context of quantum computing is the case of a qubit. A qubit, or quantum bit, refers to the quantum state of an isolated two-level quantum mechanical system. Informally, the qubit is the quantum analog of bit that serves as the fundamental unit of information within quantum computing. Methods for storing a qubit of information require the control of a physical two-level system, and we denote those physical systems as quantum register elements that have the ability to store a single qubit of information. We will discuss some of the different physical systems as quantum register elements in the “Devices for quantum computing” section. Logically, the qubit is defined over a basis of binary states labeled as “0” and “1,” respectively, such that an arbitrary state of a qubit may be expressed as the linear combination

$$|\psi\rangle = c_0|0\rangle + c_1|1\rangle.$$  

(1)

The superposition of these orthogonal basis states is fundamental to quantum mechanics. The expansion coefficients are complex-valued numbers normalized as $|c_0|^2 + |c_1|^2 = 1$ and a convenient graphical representation of the qubit is given in spherical coordinates. As shown in Figure 1, the surface of a unit sphere represents all possible qubit values, where the points of $|0\rangle$ and $|1\rangle$ are located at the north and south poles, respectively. While the absolute phase of a quantum state is arbitrary [9], $|\Psi\rangle$ is normalized to unity and must lie on the surface of the sphere. In Figure 1, the amplitudes $c_0$ and $c_1$ represent the projection of the quantum state onto the corresponding basis states and the example qubit $|\Psi\rangle$ has expansion coefficients $c_0 = \cos \theta$ and $c_1 = e^{i\phi} \sin \theta$. This representation of the qubit state on a unit sphere is commonly called the Bloch sphere in quantum mechanics.

![Figure 1. The Bloch sphere with a unit radius provides a geometrical representation of a qubit. The north and south poles of the sphere define the orthonormal basis states $|0\rangle$ and $|1\rangle$, respectively, while the surface defines the set of all possible qubit values. In spherical coordinates, the example qubit $|\Psi\rangle$ has expansion coefficients $c_0 = \cos \theta$ and $c_1 = e^{i\phi} \sin \theta$.](image-url)
More formally, a quantum state is defined as a vector within a Hilbert space, which is a complex-valued vector space supporting an inner product. By convention, the quantum state with label $\Psi$ is expressed using the “ket” notation as $|\Psi\rangle$, whereas the dual vector is expressed as the “bra” $\langle\Psi|$. The inner product between these two vectors is $\langle\Psi|\Psi\rangle$ and normalized to one. An orthonormal basis for an $N$-dimensional Hilbert space satisfies $\langle i|j \rangle = \delta_{ij}$, and an arbitrary quantum state may be represented within a complete basis as

$$|\Psi\rangle = \sum_{j=0}^{N-1} c_j |j\rangle,$$

where $c_j = \langle j|\Psi\rangle$ is the corresponding coefficient. Within a chosen basis, the coefficients of the quantum state are interpreted as probability amplitudes such that the squared magnitude of this amplitude yields the probability to lie along the chosen basis, that is, $p_j = |c_j|^2$. The mathematical theory of quantum mechanics is exceedingly rich and draws from aspects of linear algebra, probability, and complex analysis. Additional details on these aspects points are found, for example, in [9].

The fundamental equation of motion for the quantum state is the Schrodinger equation, a partial differential equation defined as

$$i\hbar \frac{\partial |\Psi(t)\rangle}{\partial t} = \hat{H}(t)|\Psi(t)\rangle,$$

where the time-dependent operator $\hat{H}(t)$ defines the energetic interactions governing the physical system and is referred to as the Hamiltonian. Consequently, the Hamiltonian is important for manipulating the quantum state and its control plays a prominent role in the design and testing of quantum computing technologies. It is important to note that a quantum state cannot be directly observed by physical measurement. Rather, the measurements of a quantum state must be performed relative to a basis set, for example, $|\langle j|\rangle$. The probability to observe the $i$th outcome corresponds to the probability $p_i$ defined above, such that a series of repeated measurements over an ensemble of identically prepared quantum states will generate a distribution of outcomes that approximates the set of probabilities $\{p_i\}$. Thus, the accurate characterization of this distribution can be exceedingly difficult due to a large number of basis states and the infrequent occurrence of measurement outcomes corresponding to low probabilities. A survey of methods for measuring quantum state is provided in [10].

A multiqubit register is an addressable array of $n$ two-level physical systems. The principle of superposition may be extended to the register as the quantum state for the composite physical system is also given by (2). For an $n$-qubit register, the computational basis is expressed in binary notation as

$$|j\rangle = |j_1, j_2 \ldots j_n\rangle = |j_1\rangle \otimes |j_2\rangle \otimes \ldots \otimes |j_n\rangle,$$

where the binary values $j_k$ correspond to the binary expansion of $j$. The dimensionality of the underlying Hilbert space is $N = 2^n$ and any normalized vector represents a valid quantum state. In particular, there are composite quantum states that cannot be expressed as separable products of $n$ single-qubit states. Such states are known as entangled states and they are a hallmark of quantum mechanics and, therefore, quantum computing. For example, consider the quantum state of a two-qubit register as

$$|\Psi\rangle = \frac{1}{\sqrt{2}} (|0_1,0_2\rangle + |1_1,1_2\rangle).$$

Measuring the individual elements of the register will generate binary outcomes 0 or 1 with equal probability. Accordingly, the classical expectation for joint measurement of the register is a uniform distribution of four possible outcomes. However, measurements of this quantum state are always correlated such that both results are either $(0,0)$ or $(1,1)$, where the probability for each of these outcomes is $1/2$. Notably, there is no possibility for observing anticorrelated outcomes for this quantum state, for example, $(0,1)$. The presence of these correlations in the measurement statistics is known as entanglement and the underlying quantum state is said to be entangled. Fundamentally, entanglement is a limitation on the ability to describe states of a register solely by specifying the value of each register element, and entangled states are notable for the ability to violate the local, causal relations predicted by classical mechanics [11].

The no-cloning principle represents a fundamental constraint placed on quantum information processing. The no-cloning principle is a consequence of the linearity of quantum mechanics [12], in which the ability to perfectly clone, aka copy, an arbitrary quantum state is not permitted. In particular, given a quantum register storing an arbitrary state $|\Psi\rangle$, this information cannot be copied into a second register without loss of information. Efforts to optimally approximate the value of the first register, known as quantum cloning [13], can be evaluated by measuring the fidelity defined as

$$f = \langle \Psi_2 | \Psi_1 \rangle^2.$$
where \( |\Psi_2\rangle \) is the value of the second register and \( f \in [0,1] \).

The principles of operation for a quantum computer are based on Schrodinger’s equation in (3), in which the time-dependent Hamiltonian \( \hat{H}(t) \) can be directly controlled through the use of externally applied fields. Depending on the specific technology in place, these controls will consist of electrical, magnetic, or optical fields designed to drive the dynamics toward a specific response. In the “Devices for quantum computing” section, we present examples for devices based on semiconductors, superconductors, and trapped-ion technologies. In some computational models, the time-dependent controls are realized as pulsed fields that act discretely on the quantum register elements. These discrete periods of field interaction are known as gates and the effect of the gate on the quantum register is described by a unitary operator that transforms the stored quantum state. This is known as the gate or circuit model since a diagrammatic sequence of gates acting on registers provides a design for instruction execution.

An alternative computational model applies the time-dependent field as continuous interaction subject to constraints on the rate of change for the overall Hamiltonian. This constraint imposes an adiabatic condition on the dynamics of the quantum system [14], such that the Hamiltonian slowly modifies the interactions between quantum physical subsystems, that is, register elements, relative to the internal energy scales describing those subsystems. As a result, the register state can be driven toward the desired outcome. This is known as the adiabatic model given the constraints on the controls. A device design based on the adiabatic model has been implemented in superconducting technology by the commercial vendor D-Wave Systems, Inc. In the realization of that design, the Hamiltonian control is restricted to a specific functional form, namely, the transverse Ising model, which limits the device operation to computing discrete optimization problems. In addition, the physics of the device is not well modeled by the Schrodinger equation (3) but rather require a more sophisticated model that includes nontrivial interactions with the surrounding quantum physical systems as well as finite-temperature effects [15]. Nonetheless, the device has been observed to correctly compute the solution to a wide variety of discrete optimization problems and has been characterized as having some advantages relative to conventional computing devices. While the remainder of this article will focus on the gate model for quantum computing, we refer the reader interested in adiabatic quantum computing to the recent review by Albash and Lidar [16].

We now summarize the basic criteria that define the expected functionality of quantum computing devices. As first presented by DiVincenzo et al. [17], these criteria represent the minimal behaviors needed to perform general-purpose quantum computing in the presence of likely architectural constraints. First is the ability to address the elements in a scalable register of quantum systems. Scalability implies a manufacturing capability to fabricate and layout as many register elements as needed for a specific computation. Second, these register elements must be capable of being initialized with high fidelity, as the starting quantum state of the computation must be well known to ensure accurate results. Third is the ability to measure register elements in a well-specified basis. As discussed above, measurement samples the statistical distribution encoded by the quantum state according the probabilities \( p_i \) over a given basis set. A measurement sample represents readout from the register of the quantum computer and this value may be subsequently processed.

Fourth, the control over the register must include the ability to apply sequences of gates drawn from a universal set. Universality of the gate set characterizes the potential to perform an arbitrary unitary operation on the quantum state using a sufficiently long series of gates from that set. In particular, it is known that a finite set of gates is sufficient to approximate universality and, moreover, that a finite set of addressable one- and two-qubit gates are sufficient for universality [18]. The latter result, known as the Solovay–Kitaev theorem, provides a constructive method for composing arbitrary gates from a finite, universal gate set. Selection of a universal gate set raises the question of the optimal instruction set architecture for an intended application within a specific device technology [19]. The fifth criterion is that the gate operation times must be much shorter than the characteristic interaction times on which the register couples to other unintended quantum physical systems. These interactions induce decoherence of the stored quantum superposition states, which leads to the loss of information [20], [21]. In order to maintain the stored quantum state with sufficient accuracy, the duration of the gate sequence must be shorter than the characteristic decoherence.
time. Fault-tolerant protocols for gate operations are designed to counter the losses from decoherence and other errors by redundantly encoding information with quantum error correction codes [22].

Two additional functional criteria are necessary for a quantum computer with geometrical constraints on the layout of the quantum register. In particular, layout constraints may impose restrictions on which register elements can be addressed by multiquit gates, for example, nearest neighbors within a 2D rectangular lattice design. Physical layout restrictions may be overcome by moving the stored quantum states between register elements. This is accomplished using the SWAP gate, a unitary operation that exchanges the quantum state between two register elements. In addition, a MOVE operation can support long-distance transport of a stored value, in which the register element itself is displaced. The latter proves useful for distributed quantum registers that may require interconnects, aka communication buses, to SWAP register values. The necessity of these functions depends on the purpose of the quantum computer and especially the limitations of the technology. Presently, all technologies for quantum computing face some constraints on register layout.

Devices for quantum computing

There are many different possible technologies available for building quantum computers, and these are typically classified by how qubits of information are stored [23]. As discussed in the “Principles of quantum computing” section, these devices must meet several functional criteria to carry out reliable quantum computation. In this section, we provide an overview of three technologies that are currently used for developing quantum computing devices and we discuss the progress toward meeting the functional criteria.

Silicon spin qubits

Silicon spin qubits denote a technology implementation by which quantum information is encoded either in the spin states of an electron in a silicon quantum dot, or in the spin state of the electron or nucleus of a single-dopant atom (typically group V donors) in a silicon substrate. In particular, the orientation of the spin in these systems is used to encode the |0⟩ and |1⟩ states. Notably, these silicon devices are fabricated with conventional CMOS techniques and consist of gate electrodes (normally aluminum or polysilicon) that can control the energy landscape in the silicon substrate. These electrodes are appropriately designed and biased such that a single electron is confined in a quantum dot at the interface. Examples of a silicon quantum dot include the MOS device as shown in Figure 2a or the Si/SiGe device as shown in Figure 2b. Similar electrostatic control is used for silicon donor devices like the example shown in Figure 2c of a phosphorus donor implanted inside a silicon substrate. In all of these examples, the electrons are strongly confined such that the lowest electronic orbital energy in the quantum dot or the donor is well isolated from other excited electronic states. The confinement length for the donor electron is ∼1.5 nm in all three dimensions, whereas for the dot electron, these dimensions are ∼10 and ∼2 nm in the lateral and vertical directions, respectively. These characteristic dimensions make silicon qubits the most compact technology as compared to the qubit technologies discussed in later sections.

Addressing silicon spin qubits uses an applied static magnetic field $B_0$ to split the orbital degeneracy of the dot electron at the interface. Due to the Zeeman effect, the orbital for the confined electron is split into the distinct spin states $|↑⟩$ and $|↓⟩$. These spin states encode the computational states |0⟩ and |1⟩, where the energy splitting is given by the Zeeman energy $\gamma_e B_0$ with $\gamma_e \sim 28$ GHz/T, being the gyromagnetic ratio of the electron. For $^{31}$P donors, the electron and nuclear spins are coupled by the hyperfine interaction, $A \sim 117$ MHz [28]. The donor qubits are generally operated under large magnetic fields $B_0 > 1$ T, such that $(\gamma_e + \gamma_n) B_0 \gg \hbar A$, where $\gamma_n \sim 17$ MHz/T is the gyromagnetic ratio of the nucleus. In this limit, the eigen spin states are the tensor products of the electronic spin ($|↑⟩, |↓⟩$) and the nuclear spin ($|↑⟩, |↓⟩$) states. The resulting energies are shown in Figure 2d, where the electron spin qubit splitting depends on the nuclear spin states, and vice versa. Typical energy splittings are of the order of tens of gigahertz and megahertz for the electron and nuclear spins, respectively [29], [30]. The hyperfine interaction $A$ and the electron gyromagnetic ratio $\gamma_e$ depend on the orbital wavefunction of the electron, which can be tuned with electric fields [31], [32]. As a result, the qubit splittings are electrically tunable after the silicon qubit devices are fabricated.

Electron spin qubits are commonly initialized and measured using spin-charge conversion.
techniques [33]. Charge sensors such as quantum point contacts and single-electron transistors (SETs) are located adjacent to the quantum dot (or donor) and are then capacitively coupled to them, cf., Figure 2. The charge sensors are biased appropriately with gate voltages, such that the current passing through them is strongly sensitive to the electrostatic environment in their vicinity. The orbital energy of the electron is then electrically tuned such that the electron can preferentially tunnel to the same or another nearby charge reservoir, depending on its spin. The presence or absence of the electron on the donor or dot can then be detected via a change in current passing through the charge sensors, which aids to readout the electron spin state. The protocol will also initialize the electron spin state in the dot or the donor to $|\downarrow\rangle$ [33].

For spin control, an oscillating (driving) magnetic field is applied to the qubits. The frequency of the oscillating field is chosen to be equivalent to the energy difference between the two spin qubit levels. Based on the principles of magnetic resonance, transitions between the spin states are then achieved at a rate proportional to the amplitude of the driving field [35]. The driving field is pulsed appropriately to obtain a specific rotation of the spin state, for implementing a single qubit gate. A microwave transmission line antenna (see Figure 2a and 2c) is normally used to generate the driving field [36], yielding magnetic field amplitudes of $\sim 0.1$ mT, and single qubit gate times of few microseconds [29] (or millisecond [30]) for the electron (or nucleus). Alternatively, a micromagnet producing a dc magnetic field gradient (Figure 2b) can also be embedded on chip [37]. In the presence of an additional oscillating electric field (from gate voltages), the electron feels an effective oscillating magnetic field, resulting in spin resonance with faster gate times. Note that the frequency of the control field is different for both the electron [electronic spin resonance (ESR) frequencies $\sim$ tens of gigahertz] and the nucleus (NMR frequencies $\sim$ tens of megahertz). The ability to control and readout the electron spin state

Figure 2. (a) Bottom: Scanning electron microscope (SEM) image of a MOS quantum dot device similar to the one where single and two-qubit gates were demonstrated. Top: Cross-sectional schematic of the device illustrating the location of qubits at the Si/SiO$_2$ interface. (Reprinted from [24] with permission from Nature.) (b) Bottom: SEM image of a Si/SiGe double quantum dot device, where two-qubit operations were implemented. Middle: Variation of the static magnetic field along the axis of quantum dots. Top: Cross sectional device schematic highlighting the position of the quantum dots. (Reprinted from [25] with permission from the American Association for the Advancement of Science (AAAS).) (c) SEM image of an ion-implanted $^{31}$P device similar to the one used for demonstrating record spin-coherence times [26], [27]. (d) $^{31}$P donor electron ($|\uparrow\rangle$, $|\downarrow\rangle$) and nuclear ($|\uparrow\rangle$, $|\down\rangle$) spins states [34].

Authorized licensed use limited to: Univ of Calif San Diego. Downloaded on November 30,2023 at 20:23:34 UTC from IEEE Xplore. Restrictions apply.
also allows measurement of the nuclear spin state. As the electron spin resonance frequency is determined by the nuclear spin state (see Figure 2d), probing frequencies at which the electron can be controlled allows readout of the nuclear spin [30].

Since the splittings are dependent on \( A \) and \( \gamma_e \), they can be tuned electrically and it is possible to independently control each donor located within a precisely positioned array [38]. In their idle state, the qubits are electrically detuned from the control field by appropriately tuning \( A \) and \( \gamma_e \). When operations need to be performed on the qubits, they are brought into resonance with the control field, that is, the energy splitting is tuned to the frequency of the control field.

The coupling between two electron spin qubits occurs via the intrinsic exchange interaction between them [38]. The exchange coupling \( J_e \) is primarily determined by the overlap between the two-electron wave functions. \( J_e \) can hence be tuned by either modifying the tunnel barrier between the two electrons or by shifting the relative orbital energies of the two electrons [39]. Both these methods can be realized by appropriately tuning the gate voltages that control the potential landscape in the device. To perform a CNOT gate, the electron spin qubits are operated in a regime where \( J_e \) is smaller than the energy difference between the qubit splittings of the two electrons (often termed as the detuning). In such a regime, each electron spin qubit will have two resonance frequencies, which are determined by the state of the other qubit. Hence, an oscillating control field at one resonant frequency will conditionally rotate the qubit dependent on the state of the other qubit, resulting in a CNOT gate [24], [25]. To perform SWAP, the qubits are initialized in a regime, where the exchange coupling is much smaller than their detuning. The exchange coupling is then increased to a value much larger than their detuning, such that the two qubits exchange information with each other. After an appropriate time that determines the angle of SWAP, the exchange coupling is brought back to a low value.

The spin-orbit coupling is weak for electrons in silicon, resulting in long spin-relaxation times \( T_1 \). The relaxation time has been shown to be dependent on the temperature and magnetic field [40]. Operating the qubits at low temperatures (<1 K) and magnetic fields (<5 T) yield \( T_1 \) exceeding several seconds and even hours. The presence of spin-containing nuclei (such as Si-29) in the lattice, and their fluctuations, can result in decoherence of the electron spins [41]. Hence, isotopic purification of silicon from spin-containing nuclei allows for long-coherence times \( (T_2) \) of milliseconds and seconds for the electron and nuclear spins, respectively [27]. Additional sources of decoherence include charge or electric field noise arising from nearby defects or traps, control signals, gate electrodes, and thermal radiation from the microwave antenna [27].

Although the methods used to address and couple silicon qubits can be integrated with the microelectronics industry, the qubits are very sensitive to atomic details that have not yet been addressed in the industry. These details strongly affect the qubit operation, and hence, it is essential to design devices that minimize their influence on the qubits. First, the exchange coupling between donor electrons is extremely sensitive to the position of donors, necessitating precise donor placement accuracies and/or large exchange coupling tunability [42], [43]. Efforts are underway to demonstrate qubits with single-donor atoms in silicon that are placed precisely with scanning tunneling microscopy [44], as well as to explore alternate means of coupling between the qubits (such as dipolar interactions [45], [46]) that are less sensitive to donor placement inaccuracies. In addition, atomic roughness and step edges at the interface can result in the excited orbital states coming close to the ground orbital state in silicon quantum dots, accelerating relaxation and even resulting in a nonspin-1/2 ground state [40]. The energy separation between the ground and excited orbital states (also referred to as valley splitting) can be tuned with electric field to an extent [47], yet it is always desirable to obtain larger and uniform valley splittings with a smooth interface. Finally, uncontrolled strain in the lattice arises from the thermal mismatch between the gate and substrate materials when the device is cooled from room temperature to milli-Kelvin temperatures [48]. This modifies the potential landscape in the device, altering the position and confinement of the quantum dots, along with introducing accidental dots. Ref. [48] highlights that using gate materials (such as polysilicon rather than aluminum), which have similar thermal expansion coefficients to that of silicon, can aid to reduce the lattice strain.

The exchange interaction between the qubits is short-range (within a few tens of nanometers) can only result in nearest neighbor couplings. To scale
up silicon qubit devices to a large-scale architecture, it is beneficial to have connectivity between qubits that are separated by much larger distances. Methods to couple silicon qubits to a photonic mode spanning ~ centimeter in a microwave resonator have been proposed [45], [49], and recently demonstrated in Si/SiGe quantum dots [50], [51]. Through the photonic mode, two qubits, separated by as far as a centimeter, can be virtually coupled to each other, enhancing the qubit-connectivity significantly. Coupling the spins to the resonator also provides a pathway to readout the spin states [45]. The transmission frequency of the resonator then depends on the spin state of the qubit. Hence, applying a microwave signal to the resonator, and measuring its transmission aids to detect the spin state.

Designing silicon spin qubit devices requires modeling several classical and quantum mechanical parameters with a range of techniques that are adapted from the semiconductor industry [34]. Classical variables that are relevant and need to be solved for include the electrostatic potential landscape, electric fields, electron densities, capacitances, magnetic fields, and strain. The electrostatic parameters in silicon devices can be obtained by solving Poisson’s equation with the finite-element method with traditional technology computer-aided design (TCAD) packages such as Sentaurus TCAD, or a general multiphysics package like COMSOL. Solving Maxwell’s equations with high-frequency electromagnetic solvers (such as CST-Microwave Studio or ANSYS-HFSS) aids estimating the driving magnetic fields generated by the microwave antenna in such devices. Thermal strain while cooling such devices can also be simulated by solving the stress–strain equations with COMSOL [48]. In addition to the classical parameters, it is also essential to solve the electronic structure in silicon qubit devices and estimate the electron orbital energies and wave functions. Effective mass theory and tight-binding techniques have been extensively used for such calculations [40]. The orbital energies and wave functions act as a handle to the hyperfine, exchange, and tunnel couplings, along with the electron gyromagnetic ratio and electron spin relaxation times. These parameters are ultimately fed into a simplified spin Hamiltonian, which is solved with mathematical packages (such as MATLAB, Mathematica, or QuTiP), to simulate the instantaneous spin states and quantum gate fidelities.

Trapped ion qubits

Trapped ion qubits represent an implementation where quantum information is encoded in the electronic energy levels of ions suspended in vacuum. To obtain trapped ions, metals such as Calcium (Ca) or Ytterbium (Yb) are first resistively heated and vaporized with a current passing through them, and then directed to the trap. While loading these ions into the trap, these vaporized neutral atoms are simultaneously photo-ionized, where their outermost electron is removed, resulting in ions that have a single valence electron. As the ions are charged particles, appropriate voltages applied to gate electrodes in their vicinity and resulting electric fields can then confine the ions in the trap. The most common gate electrode configuration for ion trapping is the (rf) Paul trap (Figure 3a), which consists of four electrodes (two with oscillating voltages and two grounded) that induce an effective harmonic potential in the xy plane, and additional two dc gate electrodes to induce harmonic confinement in the \( z \) plane [55]. In the harmonic oscillator potential, there are several eigenstates corresponding to the vibrational modes of the trapped ions. To ensure that thermal effects and fluctuating electromagnetic fields do not cause random excitation of these states and thereby motion of the ions, the ions are laser-cooled to their vibrational ground state [56]. For a small number of ions (~50), the ions will then be arranged in a linear chain along the \( z \)-direction such that overall forces from the external fields cancel out the forces from their Coulomb interaction. Typical ion separation in the trap is ~10 \( \mu \)m.

As mentioned above, a qubit is defined using the energy levels of individual ions in the trap to encode the basis states \(| 0 \rangle \) and \(| 1 \rangle \). Depending on the orbital energy levels used for encoding, there are two popular implementations of trapped-ion qubits: hyperfine and optical. For hyperfine qubits, the states correspond to the hyperfine levels in the atomic s-orbital. For example, as shown in Figure 3b, the ion \(^{171}\text{Yb}^+\) has a nuclear spin of 1/2 and the qubit is encoded using the singlet \(| S \rangle \) and \(| T_0 \rangle \) configurations of the electron and nuclear spins [57]. A small dc magnetic field is applied to separate the \(| T_0 \rangle \) state from other triplet states \(| T_+ \rangle \) and \(| T_- \rangle \). The qubit splitting of 12.6 GHz for \(^{171}\text{Yb}^+\) is determined by the hyperfine interaction between the electron and the nucleus, and insensitive to magnetic field fluctuations up to first order [58]. Alternatively, for the optical qubit
Encoding with trapped ions, the basis corresponds to s-orbital and d-orbital electronic energy levels. As shown for $^{40}$Ca$^+$ in Figure 3c [54], the energy splitting is then $\approx 411$ THz and equivalent to 729 nm. Trapped ion qubits are highly reproducible [59] provided there are no magnetic and electric field inhomogeneities in the trap, which may modify the energy levels through Stark and Zeeman effects, respectively.

Fluorescent techniques are used to visualize the ions, where the qubit states are continuously excited to the p-states with the help of a laser, to induce an electric dipole transition [56]. On such a transition, the ions scatter the photons that are detected by photomultipliers or a charge-coupled device (CCD) camera (see Figure 3a). The required laser frequency is equivalent to the separation between the energy states used for the transition and depends on the choice of the ion.

The hyperfine and optical qubits are initialized with optical pumping. Here, a laser is incident on the ions with an appropriate frequency that can continuously drive the $|1\rangle$ state to the excited p-states. Any spontaneous decay from the excited p-state to ground states apart from $|0\rangle$, are also further driven by the laser [60]. Over a period of time (~microseconds), all the spontaneous emissions result in the qubit state being initialized to $|0\rangle$ [61].

For readout of trapped-ion qubits, the laser is tuned to a frequency that continuously drives one of the basis states (e.g., $|1\rangle$) to an excited p-state. The polarization of the laser and excited state is chosen such that spontaneous emission cannot occur to the other basis state $|0\rangle$, based on spin-selection rules [60]. Hence, if the initial qubit state is $|1\rangle$, the resulting p-state after excitation may spontaneously decay to states apart from $|0\rangle$, which are also continuously excited. Photons from the spontaneous emission are then detected with a CCD camera. If the initial qubit state is $|0\rangle$, the qubit cannot be excited to the p-states by the laser, as its frequency is far away from resonance and there is no output at the CCD camera.

For optical qubits, a stable laser (having $\sim 400$-THz frequencies) with a narrow line-width can drive the transitions between the $|0\rangle$ and $|1\rangle$ states via a quadrupole transition, enabling qubit control [62]. The hyperfine qubits can be controlled with two methods. First, microwave radiation with frequencies (e.g., 12.6 GHz for $^{171}$Yb$^+$) matching the qubit splitting can drive transitions between $|0\rangle$ and $|1\rangle$ states [63]. Microwaves can be generated with a microwave horn that is located several centimeters from the trap. However, as microwaves correspond to centimeters in wavelength, and the ions are separated by micrometers, it is not possible to focus microwaves and address individual qubits in a chain of several ions. Second, stimulated Raman transitions with two laser fields (from a pulsed laser) can be used to control the qubit state [64]. Each laser field excites the qubit states to a virtual level $|e\rangle$ that is well detuned (by $\delta$) from the excited p-states (see Figure 3b). The frequency difference between the two laser fields is chosen to match the qubit splitting. Based on a Raman process, the qubit

---

**Figure 3.** (a) Schematic of a Paul trap used to confine ions in vacuum. Inset: Visualization of ions in the trap with fluorescent techniques. (Reprinted from [52] with permission from *Nature*.) (b) Electronic energy levels of a $^{171}$Yb$^+$ ion illustrating qubit encoding ($|0\rangle$ and $|1\rangle$) with hyperfine energy levels [53]. Transition between qubit states is achieved by a Raman process via excitation to a virtual state $|e\rangle$. (c) Electronic energy levels of a $^{40}$Ca$^+$ ion illustrating qubit encoding with the s- and d-orbital energy levels. (Reprinted from [54] with permission from *Springer*.)
is rotated at a frequency proportional to the product of the individual Rabi frequencies (from $|0\rangle$ to $|e\rangle$ and from $|1\rangle$ to $|e\rangle$ determined by the laser power) and inversely proportional to the detuning $\delta$ from the p-states. This method has the advantage of selectively addressing the qubits, where the laser can be focused individually on each qubit. Typical timescales for single qubit operations are of the order of several microseconds.

The Coulomb interaction between the ions serves to mediate the coupling between the qubits [52]. Based on this interaction, the qubit states are coupled to the vibrational modes of the ion chain. Hence, appropriate laser frequencies can help transferring the qubit states to the vibrational modes. Depending on the vibrational modes of the ion-trap, a subsequent ion in the chain can be rotated with a laser, to demonstrate a CNOT gate. The vibrational modes can also be swapped with the subsequent qubit, resulting in a SWAP gate.

Like silicon spin qubits, trapped ion qubits have extremely long relaxation and coherence times. The relaxation mechanism is via spontaneous decay that approaches several seconds for optical qubits, and several days for hyperfine qubits. The coherence of the qubits is primarily affected by ambient magnetic field fluctuations that modify the qubit energy levels through the Zeeman effect, laser intensity, and frequency fluctuations over time, and coupling of the qubit states to the vibrational degree of freedom during two-qubit operations [65]. The sources of decoherence for the vibrational degree of freedom include unstable trap parameters, the coupling of the electric dipole associated with the motion of ions to thermal radiation in the environment, and ion collisions with the residual background gas. Typical coherence times of the trapped ion qubits due to these effects is of the order of seconds.

The coupling rate between the qubit state and vibrational mode (for two-qubit operations) has been shown to be inversely proportional to the square root of the number of ions in the chain [61]. Hence, increasing the ion number in the chain beyond ~50 slows down the two-qubit operations, where decoherence (heating) of the motional modes and fluctuating electric fields become significant. Architectures for scale-up with a larger number of ions include quantum charge coupled device (QCCD) architectures [66] where individual ions at the edges of a trap are shuttled to nearby traps and made to interact with them, for connecting distant qubits. This would require exquisite control of the shuttling of the atomic ions, as well as the periodically cooling down the excess motion arising from shuttling ions. Although this method could potentially work for a larger number of qubits (~1000), it becomes impractical for scale-up due to the complexity of interconnects, diffraction of optical beams, and extensive hardware requirements. Photonic interfaces have been proposed to connect even larger systems [61]. Here, qubits at the edges of the chain are driven to an excited state with very fast laser pulses so that at most one photon emerges from each qubit after radiative decay. Following selection rules, the radiative decay can lead to entanglement between the photonic and trapped ion qubit. Photons from two separate qubits are mode-matched and interfered on a beamsplitter, which is then detected. A successful detection then yields an entangled state between the two distant ion trap qubits.

The design packages available in the conventional microelectronics industry cannot be directly extended to design trapped ion qubits, as their implementation has very little overlap with that of silicon. Nevertheless, the electric fields available from classical electrostatic solvers (such as COMSOL) can be used to optimize and design the gate electrode configuration and voltages for the trap. As illustrated previously in this section, the electronic orbital levels of single ions (or even a cluster of ions) in the trap, determine the laser frequencies needed for initialization, readout, control, and coupling of the trapped ion qubits. The orbital energies and hyperfine interactions for a variety of trapped ion candidate materials can be determined from ab initio electronic structure calculation techniques such as DFT. A significant aspect of the design also includes the optical setup for the lasers, including its power and focus. These parameters can be obtained with commercial ray-tracing software packages such as Zemax, Code V, or Oslo. The dynamics of the trapped ion qubits upon interaction with a laser can be mapped onto a simplified Hamiltonian, which can then be solved with commercial mathematical packages, such as MATLAB. While there are several analytical expressions and mathematical models for light-matter interactions, a device simulator capable of capturing the nonidealities in realistic trapped ion devices is currently nonexistent.
Superconducting transmon qubits

Transmon qubits encode quantum information in the charge states of superconducting islands connected by Josephson junctions. The superconductors, typically aluminum or niobium, are deposited on a silicon substrate, allowing transmon qubits to be fabricated on a large scale with techniques adapted from the microelectronics industry. A microscopic image of a transmon qubit device is shown in Figure 4a and illustrates that the qubit region spans length scales of tens of micrometers.

The simplified qubit Hamiltonian can be derived as a quantum analog of a classical \( LC \) oscillator, where \( L \) is the Josephson inductance, and \( C \) is the capacitance between the superconductors [68]. The qubit splitting is then given by \( E_{01} \approx \hbar/\sqrt{LC} \), where \( \hbar \) is the reduced Planck constant. \( E_{01} \) is typically \( \sim 5 \) GHz in units of frequency.

The total energy of the system is distributed between the inductor and capacitor, and thereby consists of two parts: 1) Josephson energy \( E_J = \hbar^2/(4e^2L) \) and 2) charging energy \( E_C = e^2/2C \) of the superconductors [69]. As the charge states constitute the qubit, they can heavily be susceptible to electric field noise. The noise sensitivity can be minimized with appropriate distribution of energies \( E_J \) and \( E_C \). Figure 4b plots the energy levels for several values of \( E_J/E_C \), and indicates that large

![Figure 4](image-url)

Figure 4. (a) The transmon qubit consisting of two superconducting islands that are coupled through Josephson junctions and a large interdigitated capacitance. Inset: SEM image of the device in the vicinity of the Josephson junctions. (b) Eigenenergies \( E_m \) (first three levels, \( m = 0, 1, 2 \)) of the superconducting system as a function of the effective offset charge \( n_g \) induced by nearby gate electrodes and environment [67]. Energies are given in units of the transition energy \( E_{01} = E_1 - E_0 \) evaluated at \( n_g = 1/2 \), and are calculated for various values of \( E_J/E_C \). The zero point energy is chosen as the bottom of \( m = 0 \) level. For increasing values of \( E_J/E_C \), \( E_m \) becomes more robust against fluctuations in \( n_g \) arising from environmental noise, whereas the anharmonicity \( (E_1 = E_{01} - E_{12}) \) reduces. \( E_J/E_C \) is chosen between 10 and 50 for transmon qubits in order to obtain robustness with sufficient anharmonicity. (c) Schematic of a transmon qubit capacitively coupled to a superconducting resonator for initialization, readout and control [67]. The capacitances between various entities of the transmon-resonator system are also labeled. (d) Equivalent circuit of a transmon coupled to the resonator [67]. [(b), (c), and (d) are reprinted from [67] with permission from the American Physical Society (APS).]
values of $E_J/E_C$ render the qubits robust against noise. However, this will also lower the difference between qubit splitting and other splittings in the system, often called the anharmonicity. A large anharmonicity is required to ensure that charge states with higher energy levels are not excited while operating the qubit. As a tradeoff, $E_J/E_C$ is normally chosen between 10 and 50 for sufficient robustness of the qubit, along with anharmonicity $E_δ ≈ E_C/2 \sim 100$ MHz [67].

To perform quantum operations, the transmon qubits are commonly placed adjacent to a superconducting resonator (Figure 4c) and is capacitively coupled to it (Figure 4d) [67], [70], [71]. Here, the qubit-resonator system is designed to be in the dispersive regime, where the detuning ($Δ \sim 100$ MHz) between the qubit and the photonic mode of the resonator is much larger than the coupling ($g \sim 10$ MHz) between them. In this regime, the shift in the resonator transmission frequency from its fundamental mode frequency is given by $\pm g^2/Δ$, where the sign (+ or −) depends on the qubit state [70]. By applying microwave pulses to the resonator, and measuring its transmission, the qubit state can hence be readout.

Resonant microwave pulses can be used to control the qubits, as the qubit splitting is $\sim 5$ GHz. Qubit control timescales are a few hundreds of nanoseconds depending on the quantum gate operation and are much faster than that of trapped ion and silicon spin qubits. Measurement of the qubit and its subsequent control also aids in deterministic initialization of the qubit state.

Two qubits, which are significantly detuned from the resonator, can be coupled to each other via the resonator. The coupling rate between the qubits is given by $g_1 g_2 / (Δ_1 + 1/Δ_2)$, where $g_1$ and $g_2$ are their individual coupling strengths to the resonator, and $Δ_1$ and $Δ_2$ are their detunings with respect to the resonator [72]. However, the effective coupling rates ($\sim$ megahertz) between the qubits will still be smaller than the detunings ($\sim$ 300 MHz) between them, caused by differences in the qubit splittings during manufacturing. As a result, the resonance frequency of each qubit will be determined by the state of the other qubit, similar to the electron/nuclear spin qubit splittings shown in Figure 2d. This enables conditional rotation of one qubit, dependent on the state of the other qubit, and hence a CNOT gate.

Alternatively, direct capacitive coupling between two adjacent transmon qubits can also be leveraged for demonstrating CNOT gates. However, using only direct capacitive coupling between the qubits leads to significant cross talk when they are incorporated in a large-scale architecture.

Compared to silicon and trapped-ion qubits, the relaxation and coherence times of superconducting qubits are short. The main sources of decoherence arise from the coupling of the qubits to additional two-level systems present in the bulk or interfaces of the device, nonequilibrium quasi-particles generated from stray infrared light, and radiation to additional modes present in device [73], [74]. The relaxation rate has also been shown to be exponentially dependent on the temperature, due to the qubit interaction with thermal photons [67]. As a result, extremely low temperatures, $\sim 20$ mK, are necessary for the high-fidelity operation of qubits. Different device designs and operation regimes during the last decade have resulted in improvements in the relaxation and coherence times by several orders of magnitude. Dephasing times currently is of the order of $\sim 100$ µs.

The Josephson energy is strongly determined by the critical current across the junction, which, in turn, is dependent on the superconducting energy gap and the normal resistance ($R_n$) of the Josephson junction when it is operated above the critical temperature [75]. $R_n$ is determined by the thickness (few nm) of the Josephson junction and can be variable across different devices. This results in nonuniform qubit splittings across devices, with an inhomogeneity of $\sim$300 MHz. Another significant challenge is the large size (several tens of micrometers) of superconducting qubits, limiting the number of qubits that can be coupled to each other via a single resonator, which spans about a centimeter. Scaling up the current demonstrations to a large-scale architecture with millions of well-connected qubits operating at extremely low temperature will benefit strongly by a reduction in the size of the qubits [76].

While a standalone tool for designing superconducting qubits is nonexistent, parameters such as the capacitance (for determining $E_C$) and inductance (for determining $E_J$) can be estimated with classical electrostatic and electromagnetic packages such as FastCap and FastHenry, respectively. Microwave software such as TXLINE (in AWR Microwave Office) has been used to design and estimate the characteristic
impedance of the superconducting resonator that aids to readout, control, and couple the qubits. In addition, the electromagnetic fields experienced by the superconducting qubits can be obtained by solving Maxwell’s equations with high-frequency electromagnetic simulators, such as ANSYS-HFSS. As for silicon and trapped-ion qubits, the qubit dynamics can also be obtained by solving the simplified Hamiltonian with mathematical packages.

To conclude this section, we reiterate that different implementations are unique with their qubit type, and methods for qubit readout, control, and interaction. We summarize our description of the different qubit technologies in Table 1.

Testing and characterization of quantum devices

In spite of the great progress in fabrication and control of qubits, today’s quantum computing devices are far noisier and error-prone than conventional digital circuits. Bit error probabilities of $10^{-3}$ to $10^{-2}$ per qubit per operation (or per clock cycle) are typical. Even with continued progress in qubit technologies, it is unlikely that the errors incurred by physical qubits will ever become negligible. Thus understanding and mitigating fault processes in qubit devices is a critical aspect of quantum computer development. Correspondingly, the experimental testing of qubit devices primarily concerns the accuracy and reliability of hardware operation rather than the correctness of the circuit logic.

Qubit device testing may be divided into two broad categories: characterization, wherein the goal is to obtain a detailed model of a device’s fault modes; and benchmarking, wherein the goal is to determine a few high-level performance metrics. Characterization is the more costly type of testing but can provide important insights leading to fault mitigation strategies or improved devices. For simply assessing the performance of a device, benchmarking is more practical.

Benchmarking: Metrics and techniques

The most basic performance metric is the probability that the device outputs the correct state. In the context of quantum mechanics, this corresponds to the inner product (or overlap) between the output state and the intended state, which is called the fidelity. The infidelity, defined as 1 minus the fidelity, quantifies the amount of error in the output state. Another common way of quantifying the output error is in terms of the geometric distance between the output state and the target state in the complex vector space.

If a qubit device is used to output a specific quantum state, for example, some reference state or resource state, the fidelity of the output with respect to this known state can be estimated by measuring random subsets of qubits along various directions of the Bloch sphere [77], [78]. In such cases, the experimental cost scales favorably with the register size. However, a qubit device would be used to perform a wide variety of computations each with a different output state, and these output states presumably cannot be computed by any conventional means. In this case, one desires experimental metrics that allow one to estimate or bound the fidelity of the device output for any computation it performs. The state-of-the-art approach for this purpose is randomized benchmarking (RB) [79]. RB is a technique for assessing how much, on average, each operation decreases the output fidelity. Essentially, RB involves measuring the final fidelity of a qubit for random operation sequences of varying lengths. For

<table>
<thead>
<tr>
<th>Feature</th>
<th>Silicon Qubits</th>
<th>Trapped-Ion Qubits</th>
<th>Transmon Qubits</th>
</tr>
</thead>
<tbody>
<tr>
<td>Qubit</td>
<td>electron/nuclear spin states</td>
<td>electron energy levels of ions in vacuum</td>
<td>charge states in superconducting islands</td>
</tr>
<tr>
<td>Qubit energy</td>
<td>tens of gigahertz (electron spin),</td>
<td>$\sim 12$ GHz ($^{171}$Yb$^+$), $\sim 411$ THz ($^{40}$Ca$^-$)</td>
<td>$\sim 5$ GHz</td>
</tr>
<tr>
<td>splitting</td>
<td>tens of megahertz (nuclear spin)</td>
<td>optical pumping</td>
<td>resonator</td>
</tr>
<tr>
<td>Readout</td>
<td>spin-charge conversion</td>
<td>optical fields</td>
<td>microwave fields</td>
</tr>
<tr>
<td>Control</td>
<td>microwave fields</td>
<td>trap vibrational modes</td>
<td>resonator</td>
</tr>
<tr>
<td>Coupling</td>
<td>exchange-interaction</td>
<td>reproducible, long coherence times</td>
<td>solid-state, fast operation</td>
</tr>
<tr>
<td>Advantages</td>
<td>compact, solid-state, long coherence times</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
weak uncorrelated errors, the fidelity decays exponentially as a function of sequence length. The RB decay constant is broadly interpreted as the average error per gate, an obviously useful performance metric. Extensions of RB have been devised to yield operation-specific error metrics [80], [81], to incorporate multiqubit operations [82], to include qubit loss [83], and to assess cross-talk [84]. While RB remains a very popular benchmarking method, its underlying fault model is not universal; hence, RB in its current form may not be entirely valid or accurate as engineering efforts continue to make the simple fault modes assumed by RB less and less prominent [85]. Additionally, it has been noted that relating RB decay constants to operation fidelities is subtly problematic [86].

Characterization via quantum tomography

An alternative to benchmarking is to thoroughly characterize the fault modes of the device. Since the output state of a quantum circuit is exponentially large in the number of qubits, characterization of a quantum circuit as a whole is generally infeasible. The established strategy is to characterize each operation of a qubit device as completely as possible so that the result of any given sequence of operations can (in principle) be predicted accurately. The general name for this strategy is quantum tomography, a name derived from the medical imaging technique in which a 3D image of a subject is reconstructed from a set of 2D projections. In a similar manner, quantum tomography reconstructs a quantum state or operation from multiple measurements, each of which reveals a particular projection of the state. This reconstruction is based on the fact that a quantum state is uniquely specified by the probability distributions for certain characteristic quantities of a physical system. (For a spin qubit, the characteristic quantities are the projection of the spin along three independent spatial directions.) State tomography is the determination of the quantum state via statistical estimation of these characteristic distributions. Tomographic methods can also be used to characterize qubit operations. A qubit operation can be thought of as a linear transformation of the characteristic probability distributions. Quantum process tomography is the determination of the transformation matrix by characterizing the output state for each possible input state, or more precisely, for a set of linearly independent states that span the state space.

Quantum tomography as just described requires well-calibrated measurements, whereas qubit measurements are among the device operations that need to be characterized. This problem is overcome with Gate Set Tomography [87], [88], the state-of-the-art method for detailed characterization of qubit devices. Gate set tomography involves tomographic measurements of many different sequences of device operations. These sequences, which range in length up to hundreds or thousands of operations, are carefully chosen to reveal all possible types of qubit errors. The data are then fit to a highly nonlinear model using a sophisticated procedure, yielding a self-consistent model of all of a device’s operations, including the measurement operations themselves. Gate Set Tomography has been used to characterize and significantly improve the control of trapped-ion qubits [89].

Other approaches

In addition to RB and Gate Set Tomography, a number of other testing approaches have been developed. Some of these remain theoretical proposals, while others have had at least limited experimental demonstrations.

One approach is to test a quantum device utilizing another quantum device, either as a reference or as a resource to perform more powerful quantum-based tests [90]. This line of approach stands to greatly reduce the cost of quantum device characterization, but it requires the availability of well-characterized quantum circuits that are similarly difficult to certify.

Another approach is to exploit prior knowledge to reduce the cost of conventional benchmarking and tomographic methods. For example, adaptive testing based on Bayesian principles can significantly accelerate both RB [91] and tomography [92], [93]. In the case that the state or operation in question has some known characteristics (e.g., it has low rank or belongs to a certain symmetry class), specialized testing methods that are more efficient are applicable [94], [95]. Related to this, the technique of compressive sensing has been adapted to the quantum domain and applied to the characterization of quantum states [96].

Other forms of testing may be categorized as model fitting, for example, determining particular parameters of qubit dynamics, or assessing particular properties of the device output (e.g., purity or
One recently developed approach to characterizing the quality of many qubit devices is to measure the distribution of output states produced by executing random quantum circuits [97]. This reveals the extent to which the device can create and maintain superpositions of computational states, a key facet of the “quantumness” of quantum computation. Finally, there is now a rapidly growing interest in the use of machine learning techniques for characterizing quantum systems. Instead of attempting to match the experimental data to an intrinsically quantum model that is likely to be intractable, researchers have begun to use neural nets to learn the behavior of quantum systems from experimental data [98]–[101]. The learning process implicitly creates a tractable model of the quantum system.

Quantum circuit design and synthesis

Quantum circuits provide representation for how register elements may be modified by a sequence of gates to implement basic computation. As summarized in the “Principles of quantum computing” section, gates represent quantum mechanical operators that address one or more register elements and by design, the gates are reversible and represented by unitary matrices [102]. However, the available gates are often restricted to well-defined subsets of available operators from which a quantum circuit specification must be constructed. Fixed-point arithmetic circuits can be used for solving complex elementary functions including evaluation of Taylor series [103], [104]. In this section, we review the design of quantum circuits with an emphasis on arithmetic operations, such as addition, subtraction, and multiplication, which are required in the implementations of many quantum algorithms [102], [105]. We also review the steps required for the synthesis of quantum circuits into technology-specific implementations.

The design of quantum arithmetic circuits based on Clifford+T gates has caught the attention of researchers [105]–[108]. Figure 5 presents the quantum gates in the Clifford+T gate set with their matrix and graphic representations. The Clifford+T quantum gate set can be used to realize multiqubit logic gates such as the Toffoli and Fredkin gates previously presented in [109] and [110]. These multiqubit gates will prove useful for describing the implementation of quantum circuits presented in this article.

<table>
<thead>
<tr>
<th>CLIFFORD+T GATE SET</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hadamard Gate</td>
</tr>
<tr>
<td>T Gate</td>
</tr>
<tr>
<td>Inverse T Gate</td>
</tr>
<tr>
<td>Phase Gate</td>
</tr>
<tr>
<td>Inverse Phase Gate</td>
</tr>
<tr>
<td>NOT Gate</td>
</tr>
<tr>
<td>(CNOT) Gate</td>
</tr>
</tbody>
</table>

![Figure 5. The Clifford+T gate set is a universal basis for expressing quantum circuits.](image)

- CNOT gate: The CNOT gate belongs to the set of Clifford+T gates, cf., Figure 5, and the two inputs produce two outputs according to the logical mapping \( A, B \rightarrow A, A \oplus B \).
- Toffoli gate: Figure 6 presents the circuit diagram and matrix representations of the Toffoli gate, while Figure 7 shows an example of how this three-qubit gate may be implemented as a series of two-qubit Clifford gates and the T gate. Notably, the Toffoli gate itself is universal for reversible logic and the three inputs produce three outputs according to the logical mapping \( A, B, C \rightarrow A, B, A \cdot B \oplus C \).
- Fredkin gate: Figure 8 presents the circuit diagram for the three-qubit Fredkin gate and its matrix representation. The Fredkin gate is universal.
for reversible logic and, as shown in Figure 9, it can also be realized as a sequence of two-qubit Clifford gates and T gates. The Fredkin gate maps three inputs to three outputs as 
\[ A, B, C \rightarrow A, \overline{A} \cdot B + A \cdot C, A \cdot B + \overline{A} \cdot C. \]

Recent proposals for the realizations of reversible logic gates and quantum circuits have focused on the fault tolerant Clifford+T gate set due to its demonstrated tolerance to noise errors \[111\], \[112\]. Potential fault-tolerant implementations of these gates could play an important role in mitigating the noise observed in current quantum computing devices \[111\]–\[113\].

While fault-tolerant implementations can help to tolerate limited amounts of noise \[114\], \[115\], it is important to note that the overhead associated with the implementation of fault-tolerant protocols can be significant \[112\], \[113\]. Therefore, an important concern for designing quantum circuits is to account for the resource overhead associated with each gate. For example, fault-tolerant T gates are well known to incur a significant increase in resources, thereby making T-count and T-depth important performance measures for fault-tolerant quantum circuit design \[114\], \[116\].

The number of qubits in a quantum circuit is a resource measure of interest because of the limited number of qubits available on existing quantum computers \[117\], \[118\]. We now define the T-count, T-depth, and qubit cost resource measures.

- Qubit cost: The total number of qubits required to design the quantum circuit.
- T-count: The total number of T gates used in the quantum circuit.
- T-depth: The number of T gate layers in the circuit, where a layer consists of quantum operations that can be performed simultaneously.

Quantum operators are reversible, and therefore, a quantum circuit must establish a one-to-one mapping between the input and output states. Ensuring a one-to-one mapping between input and output states may require circuit overhead that includes the use of ancillae qubits and garbage outputs. For example, any constant input required by the quantum circuit may be encoded using ancillae qubits. Garbage output refers to output that may not be a useful part of the result but is necessary for the quantum circuit to preserve a one-to-one mapping. The inputs regenerated at the circuit output are not considered garbage outputs \[110\]. An ideal quantum circuit would be garbageless in nature, and efforts to minimize the circuit overhead from ancillae and garbage outputs are active research areas.

When a quantum circuit has garbage outputs, the garbage outputs can be removed by using Bennett’s garbage removal scheme \[119\]. Figure 10 illustrates Bennett’s garbage removal scheme. Let \( U \) represent an arbitrary quantum circuit that performs \( f(x_1, x_2, \ldots, x_{n-1}, x_n) \) and let \( U^{-1} \) represent its logical inverse.

Bennett’s garbage removal scheme is three-step process. After \( U \) is applied, all desired outputs are copied to ancillae with CNOT gates. Then, \( U^{-1} \) is applied to the qubits of the original circuit \( U \). Thus, at the end of the computation, the garbage outputs have been restored to their initial values.
Quantum arithmetic circuits

The quantum logic gates presented in the previous section can be combined to create quantum circuits that implement quantum algorithms. As a demonstration of these ideas, we present a series of quantum circuits designed for arithmetic operations such as addition, subtraction, and multiplication. We draw these examples from recent results in the literature [106]–[108], [120].

Quantum circuit for addition

We show an example of a quantum ripple carry addition circuit with no input carry presented in [107]. Consider the addition of two n-bit numbers $a$ and $b$ stored at quantum registers $|A\rangle$ and $|B\rangle$, respectively. Furthermore, let quantum register location $|A_0\rangle$ be initialized with $z = 0$. At the end of the computation, quantum register $|A\rangle$ will keep the values $a$ and $b$, respectively. At the end of the computation, the quantum register locations $|P_{0:2\cdot n-1}\rangle$ will have the product of $a$ and $b$. The quantum register location $|P_{2\cdot n}\rangle$ will be restored to the value 0.

$$s_i = \begin{cases} a_i \oplus b_i \oplus c_i & \text{if } 0 \leq i \leq n-1 \\ c_i & \text{if } i = n \end{cases}$$

where $c_i$ is the carry bit and is defined as

$$c_i = \begin{cases} 0 & \text{if } i = 0 \\ a_{i-1} \cdot b_{i-1} \oplus b_{i-1} \cdot c_{i-1} \oplus a_{i-1} \cdot c_{i-1} & \text{if } 1 \leq i \leq n. \end{cases}$$

Quantum circuit for multiplication

We present an example of a quantum integer multiplication circuit that is presented in [108]. The quantum circuit is based on a novel design of a quantum conditional addition (Ctrl-Add) circuit with no input carry and the Toffoli gate array. The quantum multiplication circuit implements the shift and add multiplication algorithm. As a result, the circuit will require a total of $n$ Ctrl-Add circuits and Toffoli gate arrays. The Ctrl-Add circuits and Toffoli gate arrays are placed such that the shift operations are accomplished with no additional gates.

Consider the multiplication of two $n$ bit numbers $a$ and $b$ stored in quantum registers $|A\rangle$ and $|B\rangle$, respectively. Furthermore, consider a quantum register $|P\rangle$ of size $2 \cdot n + 1$ initialized to $z = 0$. At the end of the computation, the quantum registers $|A\rangle$ and $|B\rangle$ keep the values $a$ and $b$, respectively. At the end of the computation, the quantum register locations $|P_{0:2\cdot n-1}\rangle$ will have the product of $a$ and $b$. The quantum register location $|P_{2\cdot n}\rangle$ will be restored to the value 0.

Figure 11 illustrates the complete addition circuit for the case of two 4-bit inputs $a$ and $b$.

The carry bits $c_i$ are produced based on the inputs $a_{i-1}$, $b_{i-1}$ and the carry bit $c_{i-1}$ from the previous stage. Each generated carry bit $c_i$ is stored at the quantum register location $|A_i\rangle$ that initially stored the value $a_i$ for $0 \leq i \leq n - 1$. After the generated carry bits are used in further computation, each quantum register location $|A_i\rangle$ is restored to the value $a_i$ while each quantum register location $|B_i\rangle$ stores the sum bit $s_i$ for $0 \leq i \leq n - 1$. The restoration of $|A_i\rangle$ to the value $a_i$ eliminates all garbage outputs and transforming $|B_i\rangle$ to the sum $s_i$ cuts the ancillae cost to 1.
Figure 12 illustrates the quantum integer multiplication circuit for the case of two 4-bit inputs \(a\) and \(b\).

Figure 13 illustrates the quantum Ctrl-Add circuit used in the quantum multiplier for the case of two 4-bit inputs \(a\) and \(b\). The operation of the quantum Ctrl-Add circuit is conditioned on the value of the qubit (\(ctrl\)). When \(|ctrl = 1\rangle\), the circuit performs addition. The sum of \(a\) and \(b\) will appear on the quantum register that originally holds the value \(b\) at the end of the computation. The quantum register that originally holds the value \(a\) will be restored to the value \(a\). When \(|ctrl = 0\rangle\), the quantum registers that initially hold the values \(a\) and \(b\) will be restored to the values \(a\) and \(b\) at the end of the computation.

Application of quantum arithmetic circuits in Taylor series

In this section, we present an application of the quantum multiplication and quantum addition circuits presented in the previous sections. For this example, we consider a quantum circuit implementation of the Taylor series expansion. Taylor series are used in scientific computing applications, such as \(\sin(x)\), \(\ln(x)\), and \(e^x\). These functions have been used in additional quantum algorithms besides those for Hamiltonian evolution such as algorithms for Pell’s equation and the principal ideal problem [123], [124]. The value of a given function \(f(x)\) near a point \(c\) is estimated by computing the Taylor series expansion shown below

\[
f(x) \approx \sum_{i=0}^{\infty} \frac{f^{(i)}(c)}{i!} \cdot (x-c)^i.
\]

We will consider an example for a Taylor series limited to the first three terms, that is,

\[
f(x) \approx f(c) + f'(c) \cdot (x-c) + \frac{f''(c)}{2!} \cdot (x-c)^2,
\]

in which we compute this approximation for \(f(x)\) centered at value \(c\). Let \(c\) and \(x\) be \(n\)-bit values stored in quantum registers \(|x\rangle\) and \(|c\rangle\), respectively. Furthermore, let \(f(c)\), \(f'(c)\), and \(f''(c)/2!\) be represented as \(n\)-bit numbers stored at quantum registers \(|f(c)\rangle\), \(|f'(c)\rangle\), and \(|f''(c)/2\rangle\), respectively. Finally, consider quantum registers \(|Y_1\rangle, |Y_2\rangle, |Y_3\rangle,\) and \(|Y_4\rangle\) that contain ancillae set to 0. At the end of the computation, quantum register \(|Y_4\rangle\) will have the first three terms of the Taylor series expansion. The quantum registers \(|c\rangle, |x\rangle, |Y_1\rangle, |Y_2\rangle, |Y_3\rangle,\) and \(|Y_4\rangle\) will be restored to the initial values.

The quantum Taylor series circuit is built from the quantum addition circuit, the quantum subtraction circuit, the quantum multiplication circuit and the logical reverse of the quantum multiplication circuit. Figure 14 shows the graphical representation of components used in the Taylor series circuit. We will use a quantum subtraction circuit presented in [120] that is based on the ripple carry adder shown in this article.
Figure 15 illustrates an example of the quantum subtraction circuit based on the design in [120]. The quantum circuit shown calculates $\overline{b} + a$ where $\overline{b} + a = b - a$. Given two 2's complement inputs, the quantum subtractor will not experience overflow. Therefore, the circuitry used to calculate the sum bit $s_n$ is removed from the quantum adder because the circuitry is not needed to calculate $\overline{b} + a$. The steps to design the quantum Taylor series circuit are explained below. Figure 16 illustrates Steps 1 and 2.

1. **Step 1**: Calculate $f(x) = f(c) + f'(c) \cdot (x - c) + \frac{f''(c)}{2} \cdot (x - c)^2$. We use the quantum multiplication circuit, quantum addition circuit, and quantum subtraction circuit in this step. The result of the quantum subtraction circuit $x - c$ is copied to ancillae using an array of $n$ CNOT gates.

2. **Step 2**: Remove garbage output. At the end of Step 1, three quantum registers ($|Y_1\rangle$, $|Y_2\rangle$, and $|Y_3\rangle$) that initially held ancillae are transformed to $f(c) + f'(c) \cdot (x - c), (x - c)^2$ and $(x - c)$. Furthermore, at the end of the computation, quantum register $|x\rangle$ that initially held the value $x$ has been transformed to the value $x - c$. These outputs are garbage outputs. We use the logical reverse of the quantum multiplication circuit, the quantum adder, the quantum subtraction circuit, and an array of CNOT gates to remove these garbage outputs.

Quantum circuit synthesis

Quantum circuit synthesis is the generation of a quantum circuit derived from a given input definition and any necessary constraints. While this process may be either manual or automated, a typical generic quantum circuit synthesis design flow is shown in...
Figure 17. The first step of presynthesis optimization includes the addition of ancilla lines and output ordering to satisfy reversibility. The synthesis step then transforms any irreversible operation into a reversible operation—this step may be performed either optimally or heuristically. There are many different optimal methods for quantum circuit synthesis including those proposed in [126]–[128]. However, many circuits grow nonlinearly with input size and they can quickly become too large for optimal synthesis techniques. Heuristic algorithms, such as decision diagrams (see [127] and [129]) and the search of circuit databases (see [125]), are widely used in these limits. Such heuristic synthesis methods are often suboptimal and therefore further improvements may be achieved by local optimization methods. In general, synthesis methods can optimize resource metrics by reducing gate count, reducing the number of ancilla qubits, reducing the overall circuit depth, and improving locality. Postsynthesis optimization techniques such as template matching [125], [127] can also be employed to further achieve better resource usage. Finally, during the technology mapping stage, the quantum circuit is mapped (decomposed) into the intrinsic gates available within the target technology. The growing literature on the topic of quantum circuit synthesis for various levels of abstraction hints at many important avenues of research [125]–[127], [129]–[133].

*We have summarized* the basic features and requirements for quantum computing devices. This includes the fundamental criteria that a quantum computing device must implement as well as the principles of operation for performing computation within the circuit model. We have reviewed the state of the art in three specific technologies currently being developed for quantum computing devices. Silicon spins, trapped ions, and superconducting transmons represent three of the leading approaches for quantum computing but these devices still face fundamental research challenges. Therefore, methods to accurately characterize and benchmark the behavior of quantum computing devices play an important role in design and testing. We have emphasized the necessity of statistical analysis to infer the operation of quantum devices. We have also discussed the design of quantum circuits for the case of arithmetic operations, which represent an important use case for future quantum computing devices. These circuits were designed to minimize the occurrence of a specific instruction, the T gate, due to the expected complexity of fault-tolerant implementation. These designs are expected to play a critical role in future device operation as tradeoffs in gate and device complexity become more sophisticated. While we have not discussed logic test for quantum circuits, there are existing methods for fault detection in classical reversible circuits that might be applicable to quantum circuits [134]–[137]. As an example,
quantum circuits with noisy gates would require novel fault-models and novel methods for test-vector generation [138]. Therefore, we anticipate that the fault-testing of quantum circuits will be an important future research area.

The design and testing of early quantum computing devices face many near-term challenges. We have emphasized a small subset of the technologies currently under investigation for developing quantum computing devices. However, there are many more approaches to be considered, each with their own nuanced physics. This suggests that variations in the physics of each quantum computing technology may lead to different implementations for design and testing. Comparison across technologies will require standard calibration techniques that have yet to be developed. In addition, methods for quantifying well-defined metrics will be important for evaluating device performance. Current testing is focused on meeting the minimal criteria for functionality in the regime of noisy, error-prone, and faulty devices. Finally, we note that the current state of quantum computing remains focused on relatively small scale devices. Future devices, or networks of devices, are likely to include quantum registers with millions of elements and sequences with millions of highly parallelized instructions. Those devices and circuits will require more sophisticated methods for design and testing.

Acknowledgments

We thank Dr. Y.-M. Shue for valuable insights into superconducting fabrication tools and Dr. R. Pooser for guidance on quantum device development. This material is based on the work supported by the U.S. Department of Energy, Office of Science, Office of Advanced Scientific Computing Research, and Oak Ridge National Laboratory Directed Research and Development. This article has been authored by UT-Battelle, LLC, under Contract DE-AC0500OR22725 with the U.S. Department of Energy. The U.S. Government and the publisher, by accepting the article for publication, acknowledge that the U.S. Government retains a nonexclusive, paid-up, irrevocable, worldwide license to publish or reproduce the published form of this manuscript, or allow others to do so, for the U.S. Government purposes. The Department of Energy will provide public access to these results of federally sponsored research in accordance with the DOE Public Access Plan (http://energy.gov/downloads/doe-public-accessplan).

References


**Travis S. Humble** is a Distinguished Scientist with the Oak Ridge National Laboratory, Oak Ridge, TN, and holds a Joint Faculty Appointment at the Bredesen Center for Interdisciplinary Research
IEEE Design&Test and Graduate Education, University of Tennessee, Knoxville. His research interests include the design and development of quantum computing systems for scientific discovery and innovation. Humble has a PhD from the University of Oregon, Eugene, OR (2005).

Himanshu Thapliyal is an Assistant Professor and Endowed Robley D. Evans Faculty Fellow with the Department of Electrical and Computer Engineering, University of Kentucky, Lexington, KY. His research interests include quantum circuits, emerging technologies, and hardware security. Thapliyal has a PhD from the University of South Florida, Tampa, FL (2011).

Edgard Muñoz-Coreas is currently pursuing the PhD degree in electrical engineering from the University of Kentucky, Lexington, KY. His research interests include the circuit design of quantum computing and emerging technologies.

Fahd A. Mohiyaddin is a Research and Development Engineer with the Interuniversity Microelectronics Center (IMEC), Leuven, Belgium. His research interests include silicon quantum computation, nanoelectronic design, and semiconductor modeling at the atomic scale. Mohiyaddin has a PhD in electrical engineering from the University of New South Wales (UNSW Sydney), Sydney, NSW, Australia.

Ryan S. Bennink is a Senior Scientist with the Oak Ridge National Laboratory, Oak Ridge, TN. His research interests include the verification and validation of quantum information systems. Bennink has a PhD from the University of Rochester, New York, NY (2004).

Direct questions and comments about this article to Travis Humble, Oak Ridge National Laboratory, Oak Ridge, TN 37830 USA; humblets@ornl.gov.