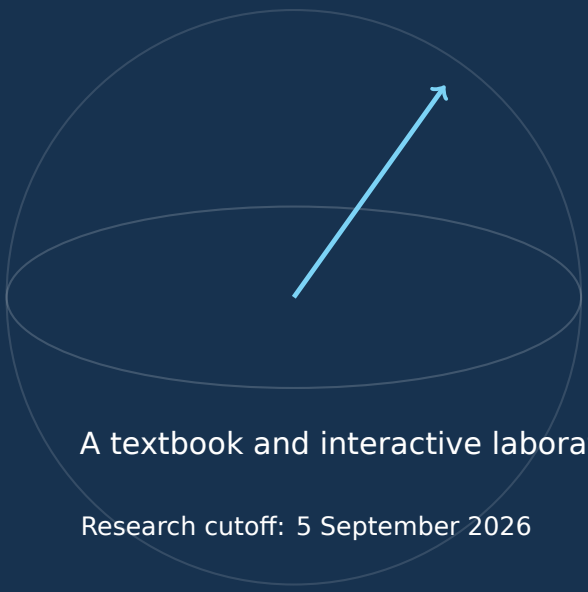
FOUNDATIONS · ALGORITHMS · RELIABLE SYSTEMS

Introduction to

Quantum Computing

2026



A textbook and interactive laboratory

Research cutoff: 5 September 2026

Introduction to Quantum Computing — 2026

An independently authored introductory course in quantum information and computation, with a synchronized interactive web edition.

Research cutoff: 5 September 2026.

Edition: 1.0, September 2026.

The text, mathematics, exercises, and laboratory specifications share one source manuscript.

Prepared with OpenAI Codex for this project. This work has reproducible mathematical and software checks; it has not undergone independent external peer review. It is not an authorized revision of Thomas G. Wong's textbook. The supplied 2022 book is acknowledged as a curriculum reference.

All diagrams are original. Research references identify their publication status and link to primary sources. Contemporary claims should be read with the stated research boundary, experimental scope, and assumptions.

INTERACTIVE EDITION

quantum-computing-2026.bensont.chatgpt.site

Laboratory links require access to the accompanying site. The PDF remains a self-contained course: derivations, worked examples, exercises, and solutions are included here.

Contents

How to use this book

This book develops quantum computing from probability and complex numbers to algorithms, noisy devices, and fault-tolerant systems. It assumes ordinary algebra, a little trigonometry, and a willingness to calculate. Programming experience and a prior course in quantum mechanics are helpful but unnecessary. The mathematical ideas that do real work are introduced where they first become useful.

The central question is operational: **what can we prepare, transform, measure, and reliably infer?** A list of amplitudes is not a list of readable answers. An entangled state is not a communication channel by itself. A quantum algorithm is not a speedup until its input, output, accuracy, and classical comparison have been specified. These distinctions connect the mathematics to honest claims about technology.

The two editions

The PDF is a continuous course with worked calculations, exercises, and solutions. The web edition contains the same chapter text and numbered exercises. Laboratory panels appear where experimentation helps. Each laboratory has a permanent identifier, and the PDF links to its location. Changing an input recomputes the mathematical model; the displays are not prerecorded demonstrations.

Begin by predicting a result, then run an experiment, then explain any discrepancy. For finite-shot experiments, distinguish disagreement caused by sampling from a wrong model. Record the preparation, basis, shot count, and random seed. A screenshot without those details is rarely a reproducible experiment.

The laboratories run classical simulations in the browser. They do not send circuits to quantum hardware. Some calculate exact state vectors, some evolve density matrices, and some use an explicitly stated reduced model. Their limits are part of the lesson: a small exact simulation, a repetition-code memory, and a hardware resource estimate answer different questions.

Routes through the material

First encounter. Read Chapters 1–8 in order. Work the projections in Chapter 3 by hand, including complex conjugation. Use the circuit laboratory before attempting the algorithm chapters.

Algorithms. Chapters 9–15 move from oracle interference to period finding, Hamiltonian simulation, and variational methods. The oracle examples expose mechanisms; small black-box problems are not industrial applications.

Reliable systems. Chapters 16–21 develop simulation, coding, fault tolerance, resource estimates, and physical hardware. Chapter 22 connects information protocols to networking and cryptography. Chapters 23–25 cover software, research evidence, and integrated projects.

A two-semester course can cover the text with laboratories and projects. A one-semester course can emphasize Chapters 1–13, 17–19, 21, and 24. This is a suggested teaching arrangement, not a promise that every beginner will master all chapters at the same speed.

Notation and arithmetic

The computational basis is $\{|0\rangle, |1\rangle\}$. The X basis is

$$|+\rangle = \frac{|0\rangle + |1\rangle}{\sqrt{2}}, \qquad |-\rangle = \frac{|0\rangle - |1\rangle}{\sqrt{2}}.$$

The Y basis is

$$|+i\rangle = \frac{|0\rangle + i|1\rangle}{\sqrt{2}}, \qquad |-i\rangle = \frac{|0\rangle - i|1\rangle}{\sqrt{2}}.$$

A bra is the conjugate transpose of a ket. In particular, $\langle +i| = (\langle 0| - i\langle 1|)/\sqrt{2}$. This minus sign is not optional. The symbol i satisfies $i^2 = -1$; indices j, k are integers. Information logarithms are base two unless stated otherwise.

Throughout the book and browser simulator, q_0 is the **leftmost, most significant** qubit:

$$|q_0 q_1 \cdots q_{n-1}\rangle, \qquad x = \sum_{j=0}^{n-1} q_j 2^{n-1-j}.$$

Many SDKs, including Qiskit, give qubit 0 the least significant weight in a state-vector index. Chapter 23 gives the conversion explicitly. Neither convention is more quantum; silently mixing them produces wrong answers.

Exact fractions and radicals are retained when they reveal cancellation. Laboratory decimals are rounded for display; calculations use double precision. A displayed zero may be smaller than the display precision. Numerical checks use tolerances and do not substitute for proofs.

Evidence and the research boundary

Research cutoff: 5 September 2026. Dated statements describe sources available by this boundary. An issue year, online publication date, preprint date, and correction date can differ. Chapter 24 records these distinctions. The evidence selection is a researched introduction, not a systematic census of every laboratory or paper.

Label	Meaning
Foundation	Mathematical result or established physical framework, with assumptions stated.
Experiment	Measurements on physical hardware, with task and metric identified.
Theory / estimate	Analysis or simulation under assumptions; not a built machine.
Preprint	Public research not treated here as a verified journal result.
Company claim	A vendor's stated capability or goal, distinguished from independent validation.
Open question	A question without a settled answer in the scope considered.

Peer review strengthens scrutiny; it does not remove uncertainty. A logical memory experiment does not establish a general-purpose fault-tolerant computer. Postselection can improve conditional fidelity while discarding many runs. A benchmark advantage does not automatically imply useful economic advantage.

Provenance and independent authorship

Thomas G. Wong's Introduction to Classical and Quantum Computing (2022) is the supplied curriculum reference [1]. Its accessible progression from classical information to qubits, linear algebra, multiple qubits, protocols, and algorithms motivated the entry level. This book uses independently written explanations, newly organized chapters, new exercises, original diagrams, and contemporary research. It is a separate educational work, not an authorized revision or publication by Wong. No substantial prose or figures from the source book are reproduced.

The emphasis differs: density operators and channels are foundational; error correction connects to logical operations and resource budgets; software follows explicit circuit semantics and current interfaces; and dated hardware evidence is separated from timeless theory.

The text and software were prepared with OpenAI Codex for this project. Mathematical identities and implemented models have reproducible checks. This production process is not an independent external peer review. References lead to underlying research, particularly where the introduction stops short of graduate-level detail.

Working with exercises

Every chapter includes problems and solutions or guidance. In the web edition, try a problem before revealing its solution. The measurement-practice laboratory generates new numerical questions, checks a probability, and reveals its projection in stages. A correct decimal alone does not demonstrate correct conjugation.

Keep a notebook of mistakes: normalization, conjugation, tensor order, gate order, omitted conditioning, wrong noise convention, and inappropriate complexity comparison. These categories become a debugging method as circuits grow.

CHAPTER 1

Information, probability, and computation

Learning goals. Encode information in bits, calculate conditional probabilities, distinguish reversible computation, and compare algorithms using stated costs.

1.1 Bits and representations

A bit labels one of two distinguishable alternatives. The labels do not specify the implementation: voltage, magnetization, or optical path may carry a bit if the alternatives can be prepared and read reliably. An n-bit string labels 2^n possibilities because every new position doubles the count.

The string 10110 represents

$$1 \cdot 2^4 + 0 \cdot 2^3 + 1 \cdot 2^2 + 1 \cdot 2 + 0 = 16 + 4 + 2 = 22.$$

It could instead label a color or database record. Encoding is a convention shared by the algorithm and its users. Confusing a number with its representation causes complexity errors. An integer N requires about $\log_2 N$ bits; a procedure using N operations can be exponential in the input length.

NOT exchanges 0 and 1. XOR, written $\oplus$, is addition modulo two: $0 \oplus 0 = 1 \oplus 1 = 0$ and $0 \oplus 1 = 1 \oplus 0 = 1$. AND gives 1 exactly when both inputs are 1. A half adder outputs sum $s = a \oplus b$ and carry $c = ab$. For $a = b = 1$, $(c, s) = (1, 0)$ is binary 10, decimal two.

1.2 Probability and shots

A distribution assigns nonnegative values $p(x)$ summing to one. For a noisy bit, $p(1) = p$ and $p(0) = 1 - p$. A probability does not mean the bit has a fractional value; it predicts outcome frequencies across repeated preparations.

For independent trials, the probability of a particular sequence multiplies. Exactly k ones in N trials has probability

$$\Pr(K = k) = \binom{N}{k} p^k (1 - p)^{N-k}.$$

The binomial coefficient counts the sequences with that many ones. The sample proportion $\hat{p} = K/N$ has mean p and variance $p(1-p)/N$. Its standard deviation decreases as $1/\sqrt{N}$. Four times the shots approximately halves statistical uncertainty.

Worked example. For $p = 1/4$ and $N = 400$,

$$\mathbb{E}[K] = 100, \quad \mathrm{Var}(K) = 75,$$
$$\mathrm{sd}(\hat{p}) = \sqrt{\frac{(1/4)(3/4)}{400}} = \frac{\sqrt{3}}{80} \approx 0.02165.$$

An observed proportion 0.27 is unsurprising. The calculation assumes identically distributed, independent shots. Drift or correlated noise can invalidate the assumption without changing the arithmetic.

Conditional probability is

$$\Pr(A \mid B) = \frac{\Pr(A \cap B)}{\Pr(B)}, \qquad \Pr(B) > 0.$$

If 60 of 100 runs are accepted and 54 accepted runs are correct, conditional accuracy is $54/60 = 0.9$. The probability of an accepted correct result per attempted run is $54/100 = 0.54$. Both numbers matter when evaluating postselection.

Writing the joint probability in two ways gives Bayes' rule:

$$\Pr(A \mid B)\Pr(B) = \Pr(B \mid A)\Pr(A),$$
$$\Pr(A \mid B) = \frac{\Pr(B \mid A)\Pr(A)}{\Pr(B)}.$$

A decoder uses this idea to infer which error likely produced an observed syndrome. The most likely error need not involve the fewest flipped bits if the prior probabilities differ.

1.3 Entropy and redundancy

Shannon entropy is

$$H(X) = -\sum_{x} p(x) \log_2 p(x),$$

with $0 \log_2 0 = 0$ by continuity. A fair bit has entropy one; a known bit has entropy zero. This measures uncertainty in a distribution, not memory size or energy [2].

Repeat a bit three times as 000 or 111. Under independent flips of probability p, majority voting fails when two or three bits flip:

$$p_{\text{fail}} = \binom{3}{2} p^2(1-p) + p^3 = 3p^2 - 2p^3.$$

At $p = 0.1$ the failure probability is 0.028. Protection costs extra physical bits and assumes reliable encoding and decoding. A common disturbance flipping all three bits defeats the independent-noise model. Quantum codes will preserve arbitrary amplitudes rather than copies of an unknown state, but the need to specify noise remains.

1.4 Reversibility and workspace

A reversible transformation is one-to-one. NOT is reversible. AND alone is not: 00, 01, and 10 produce the same output. A function can be embedded in a reversible map by preserving its input:

$$U_f : (x, y) \longmapsto (x, y \oplus f(x)).$$

Applying it twice restores y, because $f(x) \oplus f(x) = 0$. When the target starts at zero, it contains $f(x)$ afterward. The retained input prevents information loss.

CNOT maps (a, b) to $(a, b \oplus a)$. Toffoli maps (a, b, c) to $(a, b, c \oplus ab)$. Toffoli computes AND into a zero target while retaining the inputs. Its action permutes eight basis strings, so its matrix is also a quantum unitary.

Intermediate workspace can remain entangled with an answer. The compute–use–uncompute pattern calculates a function, uses it coherently, then reverses the calculation to clear temporary data. This is not housekeeping: distinguishable workspace can destroy interference between alternatives [3].

Laboratory L01 — Reversible logic and repetition. Explore CNOT and Toffoli truth tables. Change the flip probability and compare unencoded and three-bit failure probabilities. Identify the independence assumption.

1.5 Comparing algorithms

Costs include gates, sequential depth, oracle calls, shots, memory, and elapsed time. Low depth can require many parallel qubits. Few queries can hide an expensive oracle. An accurate estimate may require many executions of the entire circuit.

$T(n) = O(n^2)$ means constants C, n_0 exist with $T(n) \le Cn^2$ for $n \ge n_0$. It is an asymptotic upper bound, not an exact runtime. $\Theta(n^2)$ includes a matching lower bound. Small-input overhead and constants remain relevant in practice.

P contains decision problems with polynomial-time deterministic classical algorithms. BPP permits bounded-error classical randomization. BQP permits bounded-error quantum computation with uniform polynomial-size circuits. NP means that yes-instances have efficiently verifiable certificates; it does not stand for “not polynomial.” Whether BPP equals BQP is open. No theorem says quantum computers efficiently solve all NP-complete problems.

Factoring has a polynomial-time quantum algorithm but is not known to be NP-complete. Grover gives a quadratic query improvement in unstructured search, not an automatic exponential improvement in every optimization task. A useful comparison specifies the problem, encoding, accuracy, and best applicable classical method.

1.6 Exercises

1.1. Convert 45 to six-bit binary and count the possible strings.

Solution / guidance

$45 = 32 + 8 + 4 + 1$, giving 101101. Six bits label $2^6 = 64$ strings.

1.2. For outcome probability 0.8, find the standard deviation of the observed proportion in 100 and 1,600 shots.

Solution / guidance

$\sqrt{0.16/100} = 0.04$ and $\sqrt{0.16/1600} = 0.01$. Sixteen times the shots gives four times smaller standard deviation.

1.3. Factor $3p^2 - 2p^3 - p$. When does repetition help?

Solution / guidance

It is $-p(2p-1)(p-1)$, negative for $0 < p < 1/2$. It vanishes at $0, 1/2, 1$. A known channel with $p > 1/2$ can first have its outputs relabeled.

1.4. Explain why AND alone cannot be unitary and Toffoli can.

Solution / guidance

AND merges orthogonal inputs and changes dimension. A unitary preserves inner products and dimension. Toffoli is a bijective basis permutation with orthonormal matrix columns.

1.5. An algorithm calls an oracle $\sqrt{N}$ times; each call costs N elementary operations. What work is hidden by a query count?

Solution / guidance

The calls alone cost $N^{3/2}$ elementary operations. An oracle-call bound does not include implementation, preparation, or other algorithmic work unless explicitly stated.

CHAPTER 2

Complex amplitudes and quantum states

Learning goals. Manipulate complex numbers, normalize a state, calculate inner products, and distinguish coherent superpositions from classical mixtures.

2.1 Why amplitudes precede probabilities

Two indistinguishable paths can contribute to one detector outcome. Changing their relative phase can change the detection rate without changing either path's separate intensity. Classical addition of path probabilities cannot describe that interference. Quantum theory adds amplitudes for coherent alternatives and takes a modulus squared to obtain a probability.

This does not mean a computer calculates every answer and lets us read whichever we want. An algorithm must arrange physical transformations so the desired measurement statistics reveal useful information.

A complex number is $z = a + ib$, where $i^2 = -1$. Its conjugate is $z^* = a - ib$, and

$$|z|^2 = z^* z = (a - ib)(a + ib) = a^2 + iab - iab - i^2 b^2 = a^2 + b^2.$$

The mixed terms cancel. Squaring z instead gives $a^2 - b^2 + 2iab$, generally not a real probability.

Worked arithmetic.

$$(2 - i)(1 + 3i) = 2 + 6i - i - 3i^2 = 5 + 5i,$$
$$|2 - i|^2 = (2 + i)(2 - i) = 4 - 2i + 2i - i^2 = 5.$$

Polar form is $z = re^{i\phi} = r(\cos\phi + i\sin\phi)$. Multiplication multiplies moduli and adds phases. In particular, multiplying by $e^{i\phi}$ rotates a complex-plane arrow without changing its length.

2.2 Vectors and kets

The computational basis is

$$|0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad |1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

A pure qubit is represented by

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle = \begin{pmatrix} \alpha \\ \beta \end{pmatrix}, \qquad |\alpha|^2 + |\beta|^2 = 1.$$

A ket names a vector. Addition and scalar multiplication act componentwise. Normalization ensures that the two computational measurement probabilities sum to one.

For a nonzero vector $|v\rangle = a|0\rangle + b|1\rangle$,

$$|\psi\rangle = \frac{|v\rangle}{\sqrt{|a|^2 + |b|^2}}.$$

The zero vector cannot be normalized: it assigns probability zero to every outcome. Software should reject it, not silently substitute a state.

Worked example. Let $|v\rangle = (1+i)|0\rangle + 2|1\rangle$. Expanding the norm gives

$$\begin{aligned} \langle v|v\rangle &= (1-i)(1+i)\langle 0|0\rangle + 2(1-i)\langle 0|1\rangle \\ &\quad + 2(1+i)\langle 1|0\rangle + 4\langle 1|1\rangle \\ &= (1-i)(1+i) + 4 \\ &= 1 + i - i - i^2 + 4 = 6. \end{aligned}$$

We used $\langle 0|0\rangle = \langle 1|1\rangle = 1$ and the cross overlaps zero. Hence the normalized vector is $((1+i)|0\rangle + 2|1\rangle)/\sqrt{6}$. Its Z probabilities are $1/3, 2/3$. The cross terms vanished because the basis vectors are orthogonal, not because imaginary coefficients can be ignored.

2.3 Bras and inner products

The bra of a ket is its conjugate transpose:

$$\langle\psi| = \begin{pmatrix} \alpha^* & \beta^* \end{pmatrix}.$$

For $|\phi\rangle = \gamma|0\rangle + \delta|1\rangle$,

$$\langle\phi|\psi\rangle = \gamma^*\alpha + \delta^*\beta.$$

The inner product is linear in its right argument and conjugate-linear in its left. Thus $\langle c\phi|\psi\rangle = c^*\langle\phi|\psi\rangle$, and $\langle\phi|\psi\rangle = \langle\psi|\phi\rangle^*$.

Orthogonality means inner product zero. A suitable measurement can perfectly distinguish orthogonal states. A single copy of a nonorthogonal state cannot be distinguished perfectly from another nonorthogonal alternative with zero error.

For the X basis,

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

Both states have equal Z probabilities, but they are distinct orthogonal states. Z is simply the wrong measurement for distinguishing them.

2.4 Interference and phase

Two amplitudes $1/2$ and $1/2$ sum to one; two amplitudes $1/2$ and $-1/2$ sum to zero. Their squared sums are one and zero. Such coherent addition must be implemented by a physical transformation; it is not permission to combine probabilities arbitrarily.

Multiplying an entire ket by a global phase changes no measurement probability:

$$|\langle m|e^{i\chi}\psi\rangle|^2 = |e^{i\chi}|^2|\langle m|\psi\rangle|^2 = |\langle m|\psi\rangle|^2.$$

A phase attached to only one component is relative and can become visible after a basis change. All states $(|0\rangle + e^{i\phi}|1\rangle)/\sqrt{2}$ have equal Z probabilities; only $\phi = 0$ gives $|+\rangle$. They occupy different positions around the Bloch sphere's equator.

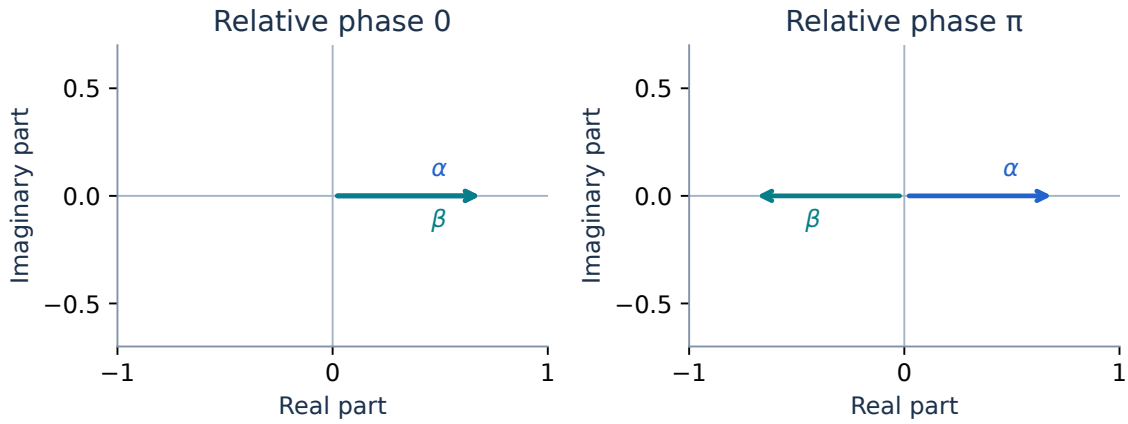


Figure 2.1: Two equal magnitudes with different relative phases. The arrows coincide on the left; on the right they point in opposite directions. A basis change makes this phase difference observable.

2.5 A mixture is not a ket sum

Procedure A prepares $|+\rangle$ every time. Procedure B uses a fair classical coin to prepare either $|0\rangle$ or $|1\rangle$. Both yield balanced Z results. But A yields X result $+$ with certainty, while B yields $+$ and $-$ equally often.

This is an operational distinction, not an interpretation dispute. Chapter 6 represents B by averaging density operators. Adding kets with classical probabilities would introduce coherence that the preparation did not create.

A state is also not identical to a complete description of the apparatus. The idealized qubit excludes leakage levels, fields controlling the system, and environmental degrees of freedom. We later incorporate their effect using channels. The simple model is useful because its assumptions can be stated and tested.

Laboratory L02 — State preparation. Edit both complex amplitudes. Inspect the raw norm, normalized vector, complex arrows, and probabilities. Compare coefficients $(1, 1)$, $(1, -1)$, and $(1, i)$. Use different bases to distinguish them.

2.6 Exercises

2.1. Find the conjugate, modulus squared, and reciprocal of $3 + 4i$.

Solution / guidance

They are $3 - 4i$, 25, and $(3 - 4i)/25$. Multiply numerator and denominator by the conjugate to obtain the reciprocal.

2.2. Normalize $(2 - i)|0\rangle + (1 + 2i)|1\rangle$ and calculate Z probabilities.

Solution / guidance

Each coefficient has modulus squared 5. Divide by $\sqrt{10}$; both probabilities are $1/2$. Keep the complex phases when changing basis.

2.3. Show that $-\beta^*|0\rangle + \alpha^*|1\rangle$ is a normalized orthogonal partner of normalized $\alpha|0\rangle + \beta|1\rangle$.

Solution / guidance

Its norm squared is $|\beta|^2 + |\alpha|^2 = 1$. Its overlap with the original ket is $-\beta\alpha + \alpha\beta = 0$.

2.4. Compare $(i|0\rangle - |1\rangle)/\sqrt{2}$ with $|+i\rangle$, and compare $|+i\rangle$ with $|-i\rangle$.

Solution / guidance

The first is $i|+i\rangle$, so only global phase differs. The Y basis pair has overlap $(1 + (-i)(-i))/2 = 0$ and is distinguishable in Y.

2.5. Derive $|(1 + e^{i\phi})/2|^2$.

Solution / guidance

Multiply by the conjugate to obtain $(2 + e^{i\phi} + e^{-i\phi})/4 = (1 + \cos\phi)/2$. It ranges from zero at $\phi = \pi$ to one at $\phi = 0$, modulo 2π.

Further study. An operational, rigorous development of quantum states and measurements is given by Watrous [4].

CHAPTER 3

Measurement and changes of basis

Learning goals. Project a state onto an orthonormal basis, retain the conjugation and cancellations, calculate conditional states, and connect probabilities to repeated experiments.

3.1 A basis specifies a question

An orthonormal basis $\{|b_0\rangle, |b_1\rangle\}$ satisfies $\langle b_j|b_k\rangle = \delta_{jk}$. Projective measurement in this basis gives

$$p(b_j) = |\langle b_j|\psi\rangle|^2.$$

For an ideal rank-one measurement, outcome b_j leaves the system in $|b_j\rangle$ up to global phase. The prior state determines probabilities; the observed result determines the conditioned state afterward.

Writing $|\psi\rangle = c_0|b_0\rangle + c_1|b_1\rangle$ and multiplying by $\langle b_j|$ gives $c_j = \langle b_j|\psi\rangle$. A basis change calculates new coordinates of the same vector. Physically applying a gate changes the vector; a gate followed by Z measurement can implement a different measurement basis.

3.2 One state, three full calculations

Our recurring example is

$$|\psi\rangle = \frac{2|0\rangle + 3i|1\rangle}{\sqrt{13}}.$$

Its norm squared is

$$\frac{(2\langle 0| - 3i\langle 1|)(2|0\rangle + 3i|1\rangle)}{13} = \frac{4 + 6i(0) - 6i(0) - 9i^2}{13} = \frac{4+9}{13} = 1.$$

3.2.1 Z basis

$$\begin{aligned}\langle 0|\psi\rangle &= \frac{2\langle 0|0\rangle + 3i\langle 0|1\rangle}{\sqrt{13}} \\ &= \frac{2(1) + 3i(0)}{\sqrt{13}} = \frac{2}{\sqrt{13}}, \\ p(0) &= \left(\frac{2}{\sqrt{13}}\right)^{*}\left(\frac{2}{\sqrt{13}}\right) = \frac{4}{13}.\end{aligned}$$

Similarly,

$$\begin{aligned}\langle 1|\psi\rangle &= \frac{2(0) + 3i(1)}{\sqrt{13}} = \frac{3i}{\sqrt{13}}, \\ p(1) &= \frac{(-3i)(3i)}{13} = \frac{-9i^2}{13} = \frac{9}{13}.\end{aligned}$$

3.2.2 X basis

$$\begin{aligned}\langle +|\psi\rangle &= \frac{(\langle 0| + \langle 1|)(2|0\rangle + 3i|1\rangle)}{\sqrt{26}} \\ &= \frac{2\langle 0|0\rangle + 3i\langle 0|1\rangle + 2\langle 1|0\rangle + 3i\langle 1|1\rangle}{\sqrt{26}} \\ &= \frac{2 + 0 + 0 + 3i}{\sqrt{26}} = \frac{2 + 3i}{\sqrt{26}}.\end{aligned}$$

The squared modulus is

$$p(+) = \frac{(2 - 3i)(2 + 3i)}{26} = \frac{4 + 6i - 6i - 9i^2}{26} = \frac{13}{26} = \frac{1}{2}.$$

For the other outcome,

$$\begin{aligned}\langle -|\psi\rangle &= \frac{(\langle 0| - \langle 1|)(2|0\rangle + 3i|1\rangle)}{\sqrt{26}} \\ &= \frac{2 + 0 - 0 - 3i}{\sqrt{26}} = \frac{2 - 3i}{\sqrt{26}}, \\ p(-) &= \frac{(2 + 3i)(2 - 3i)}{26} = \frac{1}{2}.\end{aligned}$$

3.2.3 Y basis

Forming the bra conjugates the imaginary coefficient:

$$\langle +i| = \frac{\langle 0| - i\langle 1|}{\sqrt{2}}, \qquad \langle -i| = \frac{\langle 0| + i\langle 1|}{\sqrt{2}}.$$

Now expand, including the terms that vanish:

$$\begin{aligned}
\langle +i|\psi\rangle &= \frac{(\langle 0| - i\langle 1|)(2|0\rangle + 3i|1\rangle)}{\sqrt{26}} \\
&= \frac{2\langle 0|0\rangle + 3i\langle 0|1\rangle - 2i\langle 1|0\rangle + (-i)(3i)\langle 1|1\rangle}{\sqrt{26}} \\
&= \frac{2 + 0 - 0 - 3i^2}{\sqrt{26}} = \frac{5}{\sqrt{26}}, \\
p(+i) &= \frac{25}{26}.
\end{aligned}$$

For the negative outcome,

$$\begin{aligned}
\langle -i|\psi\rangle &= \frac{(\langle 0| + i\langle 1|)(2|0\rangle + 3i|1\rangle)}{\sqrt{26}} \\
&= \frac{2 + 0 + 0 + 3i^2}{\sqrt{26}} = -\frac{1}{\sqrt{26}}, \\
p(-i) &= \left(-\frac{1}{\sqrt{26}}\right)^{*} \left(-\frac{1}{\sqrt{26}}\right) = \frac{1}{26}.
\end{aligned}$$

A negative amplitude is allowed; a negative probability is not. Each basis has its own normalized distribution:

Basis	First outcome	Second outcome
Z	$p(0) = 4/13$	$p(1) = 9/13$
X	$p(+) = 1/2$	$p(-) = 1/2$
Y	$p(+i) = 25/26$	$p(-i) = 1/26$

Do not add all six probabilities. These are three different experimental choices on freshly prepared copies.

Laboratory L03 — Projection workbench. Start with coefficients 2 and $3i$. Select Z, X, or Y; reveal the bra, expansion, cancellations, amplitude, and modulus squared. Change the coefficients and sample a chosen number of shots.

3.3 General formulas and a sign check

For normalized $\alpha|0\rangle + \beta|1\rangle$,

$$p(0) = |\alpha|^2, \quad p(1) = |\beta|^2,$$
$$p(\pm) = \frac{|\alpha \pm \beta|^2}{2} = \frac{1}{2} \pm \mathrm{Re}(\alpha^*\beta),$$
$$p(\pm i) = \frac{|\alpha \mp i\beta|^2}{2} = \frac{1}{2} \pm \mathrm{Im}(\alpha^*\beta).$$

Check the Y sign rather than memorizing it:

$$|\alpha - i\beta|^2 = (\alpha^* + i\beta^*)(\alpha - i\beta) = |\alpha|^2 + |\beta|^2 + i(\beta^*\alpha - \alpha^*\beta).$$

If $\alpha^*\beta = u + iv$, the final term is $i((u - iv) - (u + iv)) = i(-2iv) = 2v$. Divide by two to obtain $1/2 + v$. Real coefficients give balanced Y probabilities; an imaginary relative coefficient need not.

3.4 Consecutive measurements

Prepare $|+\rangle$, measure Z, then measure X. Z gives 0 or 1 equally often. Either conditioned state has balanced X probabilities. Therefore

$$\Pr(\text{final } +) = \Pr(0)\Pr(+ \mid 0) + \Pr(1)\Pr(+ \mid 1) = \frac{1}{2}\frac{1}{2} + \frac{1}{2}\frac{1}{2} = \frac{1}{2}.$$

Without the intermediate Z measurement, the final X outcome would be $+$ with certainty. Ignoring a measurement record does not undo the physical interaction.

For a projector P_m, possibly of higher rank,

$$p(m) = \langle\psi|P_m|\psi\rangle, \qquad |\psi_m\rangle = \frac{P_m|\psi\rangle}{\sqrt{p(m)}}.$$

Only outcomes with $p(m) > 0$ can be conditioned upon. A projector onto a subspace can retain coherence inside that subspace. Parity checks in quantum error correction exploit this fact.

3.5 Observables and expectation values

If outcomes have numerical values a_m, define $A = \sum_m a_m P_m$. Then

$$\langle A \rangle = \sum_m a_m p(m) = \langle \psi | A | \psi \rangle.$$

For Pauli Z, eigenvalues $+1, -1$ correspond to $|0\rangle, |1\rangle$. The ket label 0 is not an eigenvalue zero. Our example gives $\langle Z \rangle = -5/13$, $\langle X \rangle = 0$, and $\langle Y \rangle = 12/13$.

Variance is $\langle A^2 \rangle - \langle A \rangle^2$. Each Pauli squares to identity, so $\mathrm{Var}(Z) = 1 - \langle Z \rangle^2$. Incompatible observables lack a shared complete eigenbasis. Their uncertainty is not merely a poorly calibrated detector.

Laboratory L04 — Measurement practice. Generate an integer-coefficient state and requested outcome. Enter a fraction or decimal; request a hint or reveal the derivation step by step.

3.6 Exercises

3.1. Calculate all three-basis probabilities for $(3|0\rangle + 4|1\rangle)/5$.

Solution / guidance

Z: $9/25, 16/25$. X amplitudes: $7/(5\sqrt{2}), -1/(5\sqrt{2})$, giving $49/50, 1/50$. Y amplitudes: $(3 \mp 4i)/(5\sqrt{2})$, giving $1/2, 1/2$.

3.2. Replace $3i$ by $-3i$ in the recurring example. What changes?

Solution / guidance

Z and X probabilities remain unchanged. The imaginary part of $\alpha^*\beta$ changes sign, exchanging Y probabilities: $p(+i) = 1/26$, $p(-i) = 25/26$.

3.3. A Z measurement returns 1. What happens on an immediate repeated Z measurement, and if X measurement is inserted?

Solution / guidance

The conditioned state is $|1\rangle$, so direct repetition returns 1 with certainty. Inserting X leaves $|+\rangle$ or $|-\rangle$; either gives final Z result 1 with probability $1/2$.

3.4. Verify orthonormality of $|b_0\rangle = \cos\eta|0\rangle + \sin\eta|1\rangle$ and $|b_1\rangle = -\sin\eta|0\rangle + \cos\eta|1\rangle$. Find $p(b_0)$ for input $|0\rangle$.

Solution / guidance

Both norms are one and the overlap is $-\cos\eta\sin\eta + \sin\eta\cos\eta = 0$. The projection is $\cos\eta$, giving probability $\cos^2\eta$.

3.5. Find the shot-noise standard deviation of the observed $+i$ fraction for the recurring example in 2,600 shots.

Solution / guidance

$\sqrt{(25/26)(1/26)/2600} \approx 0.00377$. This describes independent sampling, not preparation or measurement bias.

CHAPTER 4

Gates, observables, and the Bloch sphere

Learning goals. Multiply gate matrices, distinguish active transformations from coordinates, derive Bloch coordinates, and understand rotations and non-commuting operations.

4.1 Matrices act on vectors

A matrix is determined by what it does to basis vectors. Its first column is the output for $|0\rangle$; its second is the output for $|1\rangle$. Linearity then fixes the output for every superposition:

$$U(\alpha|0\rangle + \beta|1\rangle) = \alpha U|0\rangle + \beta U|1\rangle.$$

For $U = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ the result is $(a\alpha + b\beta)|0\rangle + (c\alpha + d\beta)|1\rangle$. Every row forms one output amplitude. A valid closed-system gate preserves all inner products, so it satisfies $U^\dagger U = I$. Such a matrix is **unitary**, with inverse $U^\dagger$.

The Pauli gates and Hadamard are

$$X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad Y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad Z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

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

X swaps the amplitudes, Z changes their relative sign, and Y combines a swap with phases. H maps $|0\rangle$ to $|+\rangle$ and $|1\rangle$ to $|-\rangle$. These matrices are Hermitian as well as unitary; not every gate is Hermitian.

The phase gates are

$$S = \begin{pmatrix} 1 & 0 \\ 0 & i \end{pmatrix}, \qquad T = \begin{pmatrix} 1 & 0 \\ 0 & e^{i\pi/4} \end{pmatrix}, \qquad S = T^2.$$

S has inverse $S^\dagger = \mathrm{diag}(1, -i)$, not S itself.

4.2 Gate order is matrix order in reverse time

If U acts first and V second, the output is $VU|\psi\rangle$. The rightmost matrix acts first. A circuit is conventionally read left to right, which is why translating circuits into matrix products requires care.

Worked example. Apply H to $\alpha|0\rangle + \beta|1\rangle$:

$$H|\psi\rangle = \frac{(\alpha+\beta)|0\rangle + (\alpha-\beta)|1\rangle}{\sqrt{2}}.$$

Applying H again gives

$$H^2|\psi\rangle = \frac{(\alpha+\beta+\alpha-\beta)|0\rangle + (\alpha+\beta-\alpha+\beta)|1\rangle}{2} = \alpha|0\rangle + \beta|1\rangle.$$

The cancellations show $H^2 = I$. They also explain interference: the first H creates alternatives, and the second recombines them.

By contrast, $XZ = -ZX$. For instance, $XZ|0\rangle = |1\rangle$ but $ZX|0\rangle = -|1\rangle$. The minus sign is a global phase when these are complete single-qubit operations. In controlled versions, a phase confined to one control branch becomes relative and can matter. Never discard a phase before checking its role in a larger circuit.

4.3 Implementing a measurement basis

To measure X with a Z detector, apply H first, because $\langle 0|H = \langle +|$ and $\langle 1|H = \langle -|$. To measure Y, apply $S^\dagger$ and then H, so the total matrix is $HS^\dagger$:

$$\langle 0|HS^\dagger = \frac{\langle 0| - i\langle 1|}{\sqrt{2}} = \langle +i|.$$

Using SH instead produces a different operation. Writing the bra of the actual measurement is an effective way to verify a basis-rotation circuit.

4.4 Bloch coordinates from amplitudes

Remove a global phase so the first nonzero coefficient is real. Every pure qubit can then be written

$$|\psi\rangle = \cos\frac{\theta}{2}|0\rangle + e^{i\phi}\sin\frac{\theta}{2}|1\rangle, \quad 0 \le \theta \le \pi.$$

Define $x = \langle X\rangle$, $y = \langle Y\rangle$, and $z = \langle Z\rangle$. Matrix multiplication gives

$$x = \alpha^*\beta + \beta^*\alpha = 2\,\mathrm{Re}(\alpha^*\beta),$$
$$y = -i\alpha^*\beta + i\beta^*\alpha = 2\,\mathrm{Im}(\alpha^*\beta),$$
$$z = |\alpha|^2 - |\beta|^2.$$

Substituting the angular form yields

$$(x, y, z) = (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta).$$

Its length is one. The north and south poles are $|0\rangle, |1\rangle$; the positive x and y directions are $|+\rangle, |+i\rangle$. The sphere represents pure qubit states modulo global phase, not the physical path or shape of a particle.

For $(2|0\rangle + 3i|1\rangle)/\sqrt{13}$,

$$(x, y, z) = (0, 12/13, -5/13), \quad x^2 + y^2 + z^2 = (144 + 25)/169 = 1.$$

Thus equal X probabilities do not place the state at the sphere's center. They only imply its x coordinate is zero.

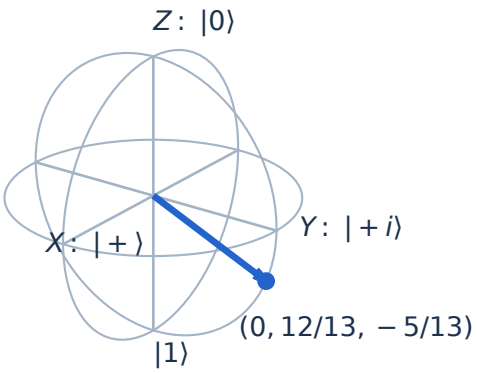


Figure 4.1: Bloch coordinates and the three measurement axes.

4.5 Rotations and generators

For a unit vector n, let $A = n_x X + n_y Y + n_z Z$. Since $A^2 = I$, the exponential series separates into even and odd powers:

$$e^{-i\theta A/2} = I\cos(\theta/2) - iA\sin(\theta/2).$$

This is $R_n(\theta)$, rotating the Bloch vector by angle θ about n. For example,

$$R_y(\theta) = \begin{pmatrix} \cos(\theta/2) & -\sin(\theta/2) \\ \sin(\theta/2) & \cos(\theta/2) \end{pmatrix},$$

$$R_z(\theta) = \mathrm{diag}(e^{-i\theta/2}, e^{i\theta/2}).$$

The half-angle is necessary: a 2π rotation gives $-I$, physically the same pure state but a different vector representative. H corresponds to a half-turn about $(x+z)/\sqrt{2}$, up to global phase; it is not simply a quarter-turn about y.

4.6 Observables and uncertainty

An observable is Hermitian: $A^\dagger = A$. Its eigenvalues are real, and its orthonormal eigenvectors specify a projective measurement. A gate is unitary. Some matrices, including the Paulis, serve both roles, but "apply Z" and "measure Z" are different physical instructions.

The commutator is $[A, B] = AB - BA$. Since $[X, Y] = 2iZ$, the uncertainty relation gives $\Delta X \Delta Y \geq |\langle Z \rangle|$. In $|0\rangle$, both X and Y outcomes are random, so $\Delta X = \Delta Y = 1$, saturating the bound. This is a statement about measurement distributions across identically prepared copies.

Laboratory L05 — Bloch sphere and gates. Adjust θ, ϕ, rotate the view, and apply H, X, Y, Z, S, or a rotation. Read all three expectations and basis probabilities. Check the recurring example's coordinates.

4.7 Exercises

4.1. Verify $HZH = X$ using the matrix columns or its action on both basis states.

Solution / guidance

$H|0\rangle = |+\rangle$, $Z|+\rangle = |-\rangle$, $H|-\rangle = |1\rangle$. The other input maps to $|0\rangle$. These outputs are exactly X's columns.

4.2. Calculate Z probabilities after $HSH|0\rangle$.

Solution / guidance

The intermediate state is $(|0\rangle + i|1\rangle)/\sqrt{2}$. The final amplitudes are $(1+i)/2, (1-i)/2$. Each has squared modulus $2/4 = 1/2$.

4.3. Find the Bloch coordinates of $(|0\rangle + e^{i\pi/3}|1\rangle)/\sqrt{2}$.

Solution / guidance

$\theta = \pi/2, \phi = \pi/3$, so $(x, y, z) = (1/2, \sqrt{3}/2, 0)$.

4.4. Derive Z probabilities after $R_y(\theta)|0\rangle$ and verify normalization.

Solution / guidance

The state is $\cos(\theta/2)|0\rangle + \sin(\theta/2)|1\rangle$. The probabilities are $\cos^2(\theta/2), \sin^2(\theta/2)$ and sum to one.

4.5. Why can a unitary not reset both $|0\rangle$ and $|1\rangle$ to $|0\rangle$?

Solution / guidance

The input overlap is zero; the proposed output overlap is one. Reset is possible as an open-system process that exports information, not as a unitary on the qubit alone.

CHAPTER 5

Registers, circuits, and entanglement

Learning goals. Expand tensor products, apply controlled gates, inspect circuit evolution, and distinguish product states from entangled states.

5.1 Tensor products build joint state spaces

For two qubits,

$$(\alpha|0\rangle + \beta|1\rangle) \otimes (\gamma|0\rangle + \delta|1\rangle) = \alpha\gamma|00\rangle + \alpha\delta|01\rangle + \beta\gamma|10\rangle + \beta\delta|11\rangle.$$

The joint amplitudes multiply for a product preparation. Our vector order is $|00\rangle, |01\rangle, |10\rangle, |11\rangle$, with q_0 on the left. An arbitrary two-qubit pure state is

$$|\psi\rangle = a|00\rangle + b|01\rangle + c|10\rangle + d|11\rangle,$$

with $|a|^2 + |b|^2 + |c|^2 + |d|^2 = 1$. Not every such state factors.

For matrices, $A \otimes B$ replaces each entry a_{jk} of A by the block $a_{jk}B$. Consequently $A \otimes I$ acts on the left qubit and $I \otimes A$ on the right. The dimension doubles per qubit; a general n-qubit pure state has 2^n complex amplitudes. This is the memory cost of a particular classical representation, not an accessible list of classical data inside a quantum register.

5.2 Controlled operations

A controlled-U with the left qubit controlling is

$$C(U) = |0\rangle\langle 0| \otimes I + |1\rangle\langle 1| \otimes U.$$

For CNOT,

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

A control in superposition does not perform a classical random choice. Linearity applies the correct transformation to each branch while retaining coherence.

Worked Bell preparation. Begin in $|00\rangle$:

$$|00\rangle \xrightarrow{H\otimes I} \frac{|00\rangle + |10\rangle}{\sqrt{2}} \xrightarrow{\text{CNOT}} \frac{|00\rangle + |11\rangle}{\sqrt{2}} = |\Phi^+\rangle.$$

The circuit has two nonzero final amplitudes. Z measurements always agree, with 00 and 11 each occurring half the time.

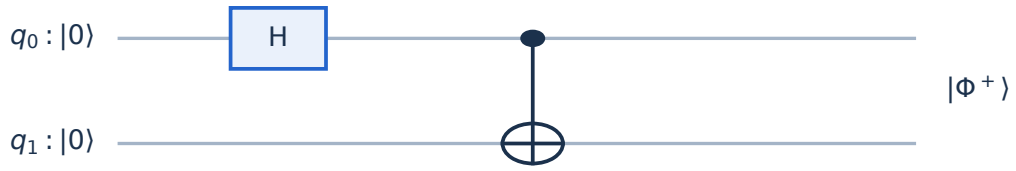


Figure 5.1: Bell-state preparation, with the top wire as the most significant qubit.

5.3 How to recognize entanglement

If the state factors, its coefficients obey $a = \alpha\gamma, b = \alpha\delta, c = \beta\gamma, d = \beta\delta$, hence $ad - bc = 0$. Conversely a nonzero 2×2 coefficient matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ with determinant zero has rank one and factors as an outer product. Therefore a normalized two-qubit pure state is a product state exactly when $ad - bc = 0$. For $|\Phi^+\rangle$, $ad - bc = 1/2$, so it is entangled. A useful pure-state measure is concurrence

$$C = 2|ad - bc|,$$

which ranges from zero for products to one for maximally entangled pairs. This formula does not apply unchanged to a mixed two-qubit state.

The family

$$|\psi(\theta, \phi)\rangle = \cos\theta|00\rangle + e^{i\phi}\sin\theta|11\rangle$$

is a product at $\theta = 0, \pi/2$ and maximally entangled at $\theta = \pi/4$. Its Z statistics do not depend on ϕ, but correlations in other bases do.

5.4 Measuring part of a register

For $a|00\rangle+b|01\rangle+c|10\rangle+d|11\rangle$, measuring the first qubit gives 0 with probability $|a|^2+|b|^2$. The conditioned state is

$$\frac{a|00\rangle+b|01\rangle}{\sqrt{|a|^2+|b|^2}}.$$

We sum probabilities over unresolved, orthogonal alternatives. We do not add $a+b$ before squaring because the second-qubit states remain distinguishable.

Worked example. Consider

$$|\chi\rangle=\frac{|00\rangle+2i|01\rangle+2|10\rangle}{3}.$$

The norm squared is $(1+4+4)/9=1$. Measuring the first qubit yields $p(0)=5/9$ and $p(1)=4/9$. Conditioned on 0, the state is $(|00\rangle+2i|01\rangle)/\sqrt{5}$; conditioned on 1 it is $|10\rangle$. The conditional probability of right-qubit result 1 in the first branch is $4/5$, so the joint probability is $(5/9)(4/5)=4/9$.

5.5 Circuit semantics

A wire represents a subsystem, not a trajectory. A box is a specified gate; a vertical line connects controls and targets; measurement produces a classical record. Disjoint gates commute, but operations on a common subsystem may not.

The laboratory lets you append gates, remove or reorder operations, step through state-vector evolution, and sample a final computational measurement. A mid-circuit measurement creates a trajectory conditioned on a sampled record. Repeating such a circuit must resample the measurement on each shot. Sampling a single collapsed trajectory repeatedly would give the wrong unconditional distribution.

Classical feed-forward means a recorded bit chooses a later operation. In teleportation, this is essential. Coherently controlling a gate with a qubit and classically controlling it with a measurement result are related but distinct implementations.

Laboratory L06 — Circuit workbench. Build circuits with up to eight qubits. Apply single-qubit, controlled, swap, and measurement operations. Step through the actual amplitudes; change shots and seed; export the circuit. Prepare a Bell state, undo it, then insert a measurement before the undo operation.

5.6 Entanglement does not guarantee computational difficulty

Some highly entangled states have compact descriptions. Bell pairs, GHZ states, and many stabilizer states can be simulated efficiently using specialized classical representations. Entanglement is important, but “entangled” is not synonymous with “beyond classical simulation.” Chapter 16 compares representations and the structures that make each effective.

The same caution applies to qubit counts. A thousand independent qubits may be easier to simulate than a much smaller generic entangled circuit. A useful resource statement must identify state structure, allowed gates, and the required output.

5.7 Exercises

5.1. Expand $|+\rangle \otimes |-i\rangle$.

Solution / guidance

The amplitudes in our order are $(1, -i, 1, -i)/2$. The squared moduli are all $1/4$; the determinant of the coefficient matrix is zero.

5.2. Is $(|00\rangle + |01\rangle + |10\rangle - |11\rangle)/2$ entangled?

Solution / guidance

$ad - bc = -1/4 - 1/4 = -1/2$, so concurrence is one. Equal measurement probabilities alone do not distinguish it from a product state.

5.3. Calculate CNOT applied to $|+\rangle|+\rangle$.

Solution / guidance

CNOT permutes the four equal amplitudes, leaving the state unchanged. A two-qubit entangling gate need not entangle every input.

5.4. Starting with $|\Phi^+\rangle$, apply CNOT and then H on the first qubit.

Solution / guidance

CNOT gives $(|00\rangle + |10\rangle)/\sqrt{2} = |+\rangle|0\rangle$; H then gives $|00\rangle$. The Bell-preparation circuit is reversed.

5.5. Explain why a left-qubit measurement probability is $|a|^2 + |b|^2$, not $|a+b|^2$.

Solution / guidance

The projector is $|0\rangle\langle 0| \otimes I$. The surviving vector is $a|00\rangle + b|01\rangle$, whose norm has cross terms proportional to $\langle 0|1\rangle = 0$ on the right subsystem.

CHAPTER 6

Density matrices and subsystems

Learning goals. Represent mixtures, calculate partial traces, test physical density matrices, and use entropy and fidelity with explicit conventions.

6.1 The density operator

For a pure state, define $\rho = |\psi\rangle\langle\psi|$. If $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$, then

$$\rho = \begin{pmatrix} |\alpha|^2 & \alpha\beta^* \\ \beta\alpha^* & |\beta|^2 \end{pmatrix}.$$

For a preparation choosing states $|\psi_j\rangle$ with classical probabilities p_j,

$$\rho = \sum_j p_j |\psi_j\rangle\langle\psi_j|.$$

This is an average of outer products, not an average ket. A valid density operator is Hermitian, positive semidefinite, and trace one. Positivity means $\langle v|\rho|v\rangle \geq 0$ for all vectors. The eigenvalues are probabilities in an eigenstate mixture.

The Born rule and unitary evolution become

$$p(m) = \mathrm{Tr}(P_m \rho), \qquad \rho \longmapsto U\rho U^\dagger.$$

The trace identity $\mathrm{Tr}(AB) = \mathrm{Tr}(BA)$ explains why the formula agrees with $\langle\psi|P_m|\psi\rangle$ for a pure state.

6.2 Coherence is basis dependent

Compare

$$\rho_+ = |+\rangle\langle+| = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, \qquad \rho_{\mathrm{mix}} = \frac{1}{2}|0\rangle\langle 0| + \frac{1}{2}|1\rangle\langle 1| = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Both have Z probabilities $1/2, 1/2$. But $\mathrm{Tr}(|+\rangle\langle+|\rho_+) = 1$, while the same expression for the mixture is $1/2$. The off-diagonal terms in the Z representation encode coherence between the Z alternatives.

The mixture has many ensemble decompositions: $I/2 = (|+\rangle\langle+| + |-\rangle\langle-|)/2$ as well. No local measurement can reveal which of these preparation stories occurred if only the density operator is available. A density operator summarizes the statistics, not a unique hidden list of constituent pure states [4].

6.3 Bloch ball and purity

Every qubit density operator has the form

$$\rho = \frac{I + xX + yY + zZ}{2} = \frac{1}{2}\begin{pmatrix} 1+z & x-iy \\ x+iy & 1-z \end{pmatrix}.$$

Its eigenvalues are $(1 \pm r)/2$, where $r = \sqrt{x^2+y^2+z^2}$. Positivity requires $r \leq 1$. Pure states lie on the sphere; mixed states lie inside the ball. The center is $I/2$, which is not a pure state with zero amplitudes.

Purity is

$$\mathrm{Tr}(\rho^2) = \frac{1+r^2}{2}.$$

It equals one exactly for pure states. For a general D-dimensional system it lies between $1/D$ and one. Purity alone does not characterize every aspect of a state's usefulness.

6.4 Partial trace: forgetting access, not measuring a hidden value

Let $|\Psi\rangle = \sum_{a,b} c_{ab}|a\rangle_A|b\rangle_B$. The state available to A is

$$\rho_A = \mathrm{Tr}_B\,|\Psi\rangle\langle\Psi|, \qquad (\rho_A)_{aa'} = \sum_b c_{ab}c^*_{a'b}.$$

Derive this by expanding the joint outer product:

$$\rho_{AB} = \sum_{a,b,a',b'} c_{ab}c^*_{a'b'}|a\rangle\langle a'| \otimes |b\rangle\langle b'|.$$

The trace of $|b\rangle\langle b'|$ is $\delta_{bb'}$, which eliminates terms with different environment indices. We sum matching indices, not all entries of the matrix.

For a Bell pair,

$$\rho_{AB} = \tfrac{1}{2}(|00\rangle\langle 00| + |00\rangle\langle 11| + |11\rangle\langle 00| + |11\rangle\langle 11|).$$

The two cross terms vanish under Tr_B because $\mathrm{Tr}(|0\rangle\langle 1|) = 0$. Hence $\rho_A = I/2$. The joint state is pure, but each subsystem is mixed. A local mixed state need not reflect ordinary ignorance about an independently existing local pure state.

Worked partial trace. For $|\chi\rangle = (|00\rangle + 2i|01\rangle + 2|10\rangle)/3$,

$$\rho_A = \frac{1}{9}\begin{pmatrix} 5 & 2 \\ 2 & 4 \end{pmatrix}, \qquad \rho_B = \frac{1}{9}\begin{pmatrix} 5 & -2i \\ 2i & 4 \end{pmatrix}.$$

For example $(\rho_A)_{01} = c_{00}c_{10}^* + c_{01}c_{11}^* = (1/3)(2/3) + (2i/3)(0) = 2/9$. Their purities both equal $(25+4+4+16)/81 = 49/81$, as required for the reductions of a bipartite pure state.

Laboratory L07 — Entanglement and partial trace. Edit all four complex amplitudes of two qubits. Inspect both reduced matrices, their Bloch vectors, purity, entropy, and pure-state concurrence. Compare a Bell state with a product having the same Z distribution on each subsystem.

6.5 Schmidt coefficients and entropy

The singular-value decomposition of the coefficient matrix gives the Schmidt form

$$|\Psi\rangle = \sum_j \sqrt{\lambda_j}|u_j\rangle_A|v_j\rangle_B, \qquad \lambda_j \geq 0, \quad \sum_j \lambda_j = 1.$$

The λ_j are the nonzero eigenvalues of both reduced density matrices. One nonzero Schmidt coefficient means a product state. More than one means entanglement for a pure joint state.

Von Neumann entropy is $S(\rho) = -\mathrm{Tr}(\rho \log_2 \rho) = -\sum_j \lambda_j \log_2 \lambda_j$. For a bipartite pure state, $S(\rho_A) = S(\rho_B)$ quantifies entanglement. For a mixed joint state, local entropy alone is not an entanglement measure. A classically correlated mixture of 00 and 11 also has locally maximally mixed states.

6.6 Distinguishability and generalized measurements

Our fidelity convention is the **squared** convention:

$$F(\rho, \sigma) = \left(\mathrm{Tr}\sqrt{\sqrt{\rho}\,\sigma\sqrt{\rho}}\right)^2.$$

For two pure states it is $|\langle\psi|\phi\rangle|^2$. Some papers use its square root; compare definitions before comparing numbers. Trace distance is $D(\rho,\sigma)=\frac{1}{2}\|\rho-\sigma\|_1$, where the trace norm sums singular values. It bounds the difference in probability assigned to any measurement event.

A generalized measurement, or POVM, consists of positive operators E_m summing to I, with $p(m)=\mathrm{Tr}(E_m\rho)$. Effects specify outcome probabilities, but not the full postmeasurement state. For that one needs measurement operators M_m or an instrument, with $E_m=M_m^\dagger M_m$ and conditioned state $M_m\rho M_m^\dagger/p(m)$. Several instruments can realize the same POVM.

6.7 Exercises

6.1. Find ρ for the recurring single-qubit example.

Solution / guidance

$\rho = \frac{1}{13}\begin{pmatrix} 4 & -6i \\ 6i & 9 \end{pmatrix}$. Its determinant is $(36-36)/169 = 0$, trace one, and purity one.

6.2. Is $\begin{pmatrix} 1/2 & 3/4 \\ 3/4 & 1/2 \end{pmatrix}$ physical?

Solution / guidance

No. It is Hermitian with trace one, but eigenvalues are $5/4, -1/4$. Positivity is a separate requirement.

6.3. Reduce $\sqrt{3/4}|00\rangle + \sqrt{1/4}|11\rangle$.

Solution / guidance

Each reduction is $\mathrm{diag}(3/4, 1/4)$, purity $5/8$, and entropy $h_2(1/4) \approx 0.8113$ bits. The off-diagonal joint terms vanish under the trace.

6.4. Explain why a locally maximally mixed state does not prove joint entanglement.

Solution / guidance

Both the Bell state and the separable mixture $(|00\rangle\langle 00| + |11\rangle\langle 11|)/2$ have reduction $I/2$. Joint coherences or correlations in additional bases are needed.

6.5. Calculate fidelity of $|0\rangle$ with $|+\rangle$, and of $|0\rangle$ with $I/2$.

Solution / guidance

Both equal $1/2$ in the squared convention. Fidelity with a pure target is $\langle 0|\rho|0\rangle$.

CHAPTER 7

Bell correlations, teleportation, and dense coding

Learning goals. Derive the CHSH bound, distinguish correlation from signalling, and track the resources and corrections in two communication protocols.

7.1 No-cloning and no-signalling

Suppose a unitary cloned arbitrary states: $U|\psi\rangle|0\rangle = |\psi\rangle|\psi\rangle$ and likewise for $|\phi\rangle$. Preserving the input overlap requires

$$\langle\phi|\psi\rangle = (\langle\phi|\psi\rangle)^2.$$

An overlap strictly between zero and one cannot satisfy this. Thus an unknown arbitrary state cannot be copied perfectly. Orthogonal labels can be copied; the theorem does not prohibit ordinary classical copying [5].

Entanglement also cannot transmit a chosen bit without communication. For a local trace-preserving channel on B with Kraus operators K_j,

$$\rho'_A = \mathrm{Tr}_B \sum_j (I \otimes K_j)\rho_{AB}(I \otimes K_j^\dagger) = \rho_A,$$

because $\sum_j K_j^\dagger K_j = I$. A conditioned state at A may depend on B's result, but A needs B's classical record to identify that subensemble.

7.2 CHSH from four local answers

Alice chooses setting a or a', Bob chooses b or b', and each outputs ± 1. In a local hidden-variable model, fix the hidden variable and write predetermined responses $A, A', B, B' \in \{\pm 1\}$. Then

$$AB + AB' + A'B - A'B' = A(B + B') + A'(B - B').$$

Either $B = B'$ or $B = -B'$. Exactly one parenthesis is zero and the other is ± 2, so the expression is always ± 2. Averaging gives

$$|S| = |E(a,b) + E(a,b') + E(a',b) - E(a',b')| \leq 2.$$

Stochastic local responses can be included in the hidden variable, so the bound remains. The assumptions include independent setting choices and an appropriate separation of local responses; a classical simulation programmed with quantum probabilities is not an experimental test of these assumptions [6, 7].

For $|\Phi^+\rangle$ and observables $A(\theta) = \cos\theta Z + \sin\theta X$, one obtains $E(a,b) = \cos(a-b)$. Choose $a = 0, a' = \pi/2, b = \pi/4, b' = -\pi/4$:

$$S = \frac{1}{\sqrt{2}} + \frac{1}{\sqrt{2}} + \frac{1}{\sqrt{2}} - \left(-\frac{1}{\sqrt{2}}\right) = 2\sqrt{2}.$$

The joint probabilities are $p(s,t \mid a,b) = (1 + st\cos(a-b))/4$. Summing over t gives $p(s) = 1/2$, independent of Bob's setting. Strong correlations coexist with no signalling.

Laboratory L08 — CHSH experiment. Change all four measurement angles, visibility, shots per setting, and seed. Compare exact and sampled correlations and S. Visibility models $\rho = v|\Phi^+\rangle\langle\Phi^+| + (1-v)I/4$, not a hardware calibration.

7.3 Teleportation: move a state using shared entanglement

Alice owns input $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$ and one half of a Bell pair; Bob owns the other half. Order the wires as input, Alice's Bell half, Bob's half. The initial state is

$$\frac{1}{\sqrt{2}}(\alpha|000\rangle + \alpha|011\rangle + \beta|100\rangle + \beta|111\rangle).$$

Alice applies CNOT from the input to her Bell half, then H to the input. Grouping by Alice's two measurement bits gives

$$\begin{aligned}\frac{1}{2}\big[&|00\rangle(\alpha|0\rangle + \beta|1\rangle) + |01\rangle(\alpha|1\rangle + \beta|0\rangle)\\ &+|10\rangle(\alpha|0\rangle - \beta|1\rangle) + |11\rangle(\alpha|1\rangle - \beta|0\rangle)\big].\end{aligned}$$

Each bracket has norm one, so every two-bit outcome has probability $1/4$, independent of the input. Bob's conditioned state is $X^b Z^a|\psi\rangle$, where a is Alice's first bit and b her second. He corrects by applying X when $b = 1$, then Z when $a = 1$: the total correction is $Z^a X^b$.

Alice's bits ab	Bob before correction	Operations in time order
00	$\|\psi\rangle$	None
01	$X\|\psi\rangle$	X
10	$Z\|\psi\rangle$	Z
11	$XZ\|\psi\rangle$	X, then Z

For branch 11, $(ZX)(XZ) = I$. Reversing correction order only adds a branch-global minus sign, but consistent conventions make the derivation transparent.

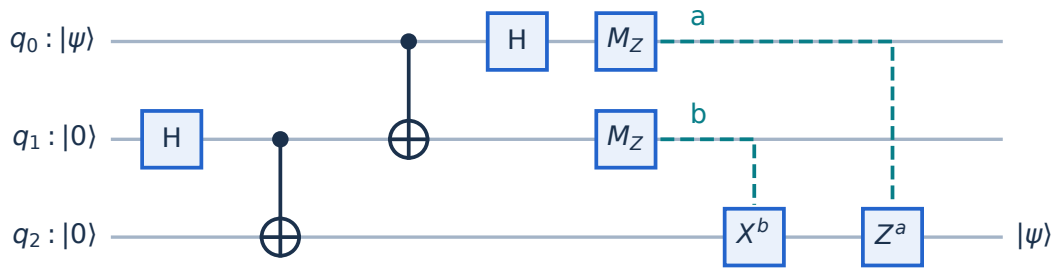


Figure 7.1: Teleportation with two classical feed-forward bits.

No copy remains at Alice. The protocol consumes shared entanglement and two classical bits. Bob cannot use it for instantaneous messaging: before receiving a and b, his averaged state is $I/2$. The original protocol is due to Bennett and collaborators [8].

Laboratory L09 — Teleportation. Prepare an arbitrary input, inspect all four branches, disable either correction, and compare the output fidelity with the original state.

7.4 Superdense coding: the complementary resource trade

Alice and Bob first share $|\Phi^+\rangle$. Alice encodes two classical bits a,b by applying X when $b = 1$, then Z when $a = 1$ to her qubit. This gives four orthogonal Bell states up to phase. She then physically sends her qubit to Bob.

Bob applies CNOT from Alice's qubit to his, then H to Alice's qubit. A Z measurement returns ab exactly. One transmitted qubit carries two selected classical bits **given an already shared Bell pair**. The prior entanglement is a consumed resource; ignoring its distribution would misstate the communication cost [9].

Laboratory L10 — Dense coding. Select a two-bit message, inspect the encoded Bell state, and run the decoding circuit. Compare the sender's reduced state for all four messages.

7.5 Exercises

7.1. For the noisy Bell model, determine the visibility needed to violate CHSH with optimal angles.

Solution / guidance

$S = 2\sqrt{2}v$. Violation requires $v > 1/\sqrt{2} \approx 0.7071$. Failure to violate this particular inequality does not by itself prove separability.

7.2. Why is Alice's teleportation record independent of α, β?

Solution / guidance

Each branch is a unitary image of the normalized input multiplied by $1/2$, hence has probability $1/4$. Two classical bits cannot reveal the continuous input amplitudes.

7.3. Teleport $|+i\rangle$ and condition on ab=01. Show the correction.

Solution / guidance

Bob has $X|+i\rangle = (i|0\rangle + |1\rangle)/\sqrt{2}$. Applying X again returns $|+i\rangle$, since $X^2 = I$.

7.4. Dense-code message 11. Find the encoded state and final output.

Solution / guidance

X on Alice produces $(|10\rangle + |01\rangle)/\sqrt{2}$. Z gives $(-|10\rangle + |01\rangle)/\sqrt{2}$. CNOT and H produce $|11\rangle$ up to an irrelevant overall phase.

7.5. Does an arbitrarily large sample of one local half of Bell pairs reveal the distant party's basis choice?

Solution / guidance

No. Every local copy has density matrix $I/2$ regardless of a trace-preserving distant operation. Joint records reveal correlations only after communication.

CHAPTER 8

Noise channels, measurement data, and mitigation

Learning goals. Evolve density matrices with Kraus operators, distinguish noise conventions, estimate observables, and identify mitigation's bias and sampling costs.

8.1 Open-system dynamics

A closed joint system evolves unitarily, but a subsystem interacting with an unobserved environment generally does not. For an initially uncorrelated system and environment, tracing out the environment gives a completely positive trace-preserving map:

$$\mathcal{E}(\rho) = \sum_j K_j \rho K_j^\dagger, \qquad \sum_j K_j^\dagger K_j = I.$$

The second equation preserves the trace. Complete positivity means positivity remains even when the system is part of a larger entangled state. Merely mapping isolated positive matrices to positive matrices is not enough.

Kraus representations are not unique. The operators can be interpreted through an environment measurement in a chosen dilation, but an arbitrary Kraus index should not be treated as a directly observed physical error. Conditioning on an actual record and averaging over unobserved records are different operations.

8.2 Three channels with explicit conventions

The phase-flip channel is

$$\mathcal{E}_Z(\rho) = (1-p)\rho + pZ\rho Z.$$

It leaves populations unchanged and multiplies off-diagonal entries by $1-2p$. Thus complete dephasing occurs at $p = 1/2$, not at $p = 1$. At $p = 1$ it is a deterministic Z gate. A different convention writes coherence decay as $1-\lambda$; these parameters must not be interchanged.

We define depolarization by

$$\mathcal{D}_p(\rho) = (1-p)\rho + pI/2, \quad 0 \le p \le 1.$$

It shrinks the Bloch vector by $1-p$. Equivalently,

$$\mathcal{D}_p(\rho) = (1-3p/4)\rho + \frac{p}{4}(X\rho X + Y\rho Y + Z\rho Z).$$

Here the probability of a nonidentity Pauli in this representation is $3p/4$. Some packages call that total probability p instead. Always inspect the implemented equation.

Amplitude damping models relaxation from $|1\rangle$ toward $|0\rangle$:

$$K_0 = \begin{pmatrix} 1 & 0 \\ 0 & \sqrt{1-\gamma} \end{pmatrix}, \qquad K_1 = \begin{pmatrix} 0 & \sqrt{\gamma} \\ 0 & 0 \end{pmatrix}.$$

For $\rho = \begin{pmatrix} a & c \\ c^* & b \end{pmatrix}$,

$$\mathcal{A}_\gamma(\rho) = \begin{pmatrix} a+\gamma b & \sqrt{1-\gamma}\, c \\ \sqrt{1-\gamma}\, c^* & (1-\gamma)b \end{pmatrix}.$$

The trace remains $a + b = 1$. The map is not unital: it sends $I/2$ toward a state biased to $|0\rangle$. In an exponential relaxation model, $\gamma = 1 - e^{-t/T_1}$. Coherence time T_2 also includes dephasing; under the usual independent Markovian model, $1/T_2 = 1/(2T_1) + 1/T_\phi$. These are model-specific time constants, not interchangeable gate fidelities.

Worked example. Apply damping with $\gamma = 1/4$ to $|+\rangle$:

$$\rho' = \begin{pmatrix} 5/8 & \sqrt{3}/4 \\ \sqrt{3}/4 & 3/8 \end{pmatrix}.$$

Its Bloch vector is $(\sqrt{3}/2, 0, 1/4)$, with squared length $13/16$ and purity $29/32$. Both the population shift and coherence reduction matter.

Laboratory L11 — Density-matrix noise. Choose a pure input, optionally mix it with $I/2$, then apply bit flip, phase flip, depolarization, or amplitude damping. Inspect the full output matrix, Bloch vector, eigenvalues, and purity.

8.3 Noise is not always an independent coin flip

A coherent overrotation applies $R_z(\epsilon)$ every time. After L repetitions, it is $R_z(L\epsilon)$, so errors may add coherently. An independent stochastic phase-flip channel combines differently. Crosstalk, leakage out of the qubit subspace, slow drift, loss, and correlated events require richer models.

Replacing coherent noise by a Pauli model can simplify simulation but change the prediction. Pauli twirling is a physical or mathematical averaging operation with specific assumptions, not proof that the original noise was already stochastic.

8.4 Estimating observables and states

A Pauli measurement yields values ± 1. With N independent shots and sample mean $\hat{E}$,

$$\mathrm{Var}(\hat{E}) = \frac{1 - E^2}{N}.$$

For $H = \sum_j h_j P_j$, estimate each Pauli expectation and sum. If independently sampled, the energy estimator variance is $\sum_j h_j^2 \,\mathrm{Var}(\hat{E}_j)$. Commuting terms may share measurement settings, in which case covariance must be included.

Single-qubit tomography estimates x,y,z from separate basis samples and forms $\hat{\rho} = (I + \hat{x}X + \hat{y}Y + \hat{z}Z)/2$. Finite data can yield a vector outside the Bloch ball, making the naive estimate nonphysical. Constrained estimation can enforce positivity, but introduces statistical choices that should be reported. Full generic n-qubit tomography requires exponentially many parameters; special assumptions can reduce the task.

8.5 Error mitigation is an estimator strategy

Error correction protects encoded quantum information during a computation. Error mitigation combines noisy experiments and classical inference to estimate an ideal quantity. Mitigation need not output a protected quantum state and may incur severe sampling costs [10].

For zero-noise extrapolation, assume an observable varies smoothly with a controllable effective noise strength:

$$E(\lambda) = E_0 + c_1\lambda + c_2\lambda^2 + \cdots.$$

Measurements at λ and 2λ give

$$2E(\lambda) - E(2\lambda) = E_0 - 2c_2\lambda^2 + \cdots.$$

The leading linear bias cancels. But for independent estimates,

$$\mathrm{Var}(2\hat{E}_1 - \hat{E}_2) = 4\,\mathrm{Var}(\hat{E}_1) + \mathrm{Var}(\hat{E}_2).$$

With comparable input variances, this is five times one estimate's variance. A quadratic combination $3E(\lambda) - 3E(2\lambda) + E(3\lambda)$ has still larger coefficient-squared cost. Extrapolations can lie outside $[-1, 1]$; clipping changes the estimator and its bias.

Probabilistic error cancellation represents an inverse noise operation with signed weights. The absolute-weight sum drives sampling overhead, often growing rapidly with circuit size. Readout mitigation inverts or fits a calibrated classical assignment model and can amplify statistical errors when that model is ill-conditioned. None provides a free route to arbitrary-depth ideal computation.

Laboratory L12 — Bias and variance. Use the stated model $E(\lambda) = e^{-\lambda}$, sample three noise levels, and compare raw, linear, and quadratic estimates. Change noise, shots, and seed. The model's ideal value is known so bias can be measured.

8.6 Exercises

8.1. What does the phase-flip channel do to $|+\rangle$ at $p = 0, 1/2, 1$?

Solution / guidance

It yields $|+\rangle\langle+|$, $I/2$, and $|-\rangle\langle-|$. Purity decreases and then increases; the parameter is a flip probability, not monotone decoherence strength over the whole interval.

8.2. Verify $K_0^\dagger K_0 + K_1^\dagger K_1 = I$ for amplitude damping.

Solution / guidance

The two terms are $\mathrm{diag}(1, 1-\gamma)$ and $\mathrm{diag}(0, \gamma)$, whose sum is identity.

8.3. Find the depolarized state and purity of an initially pure qubit for $p = 0.2$.

Solution / guidance

Its Bloch vector is shortened to length 0.8. Purity is $(1 + 0.64)/2 = 0.82$; eigenvalues are 0.9 and 0.1.

8.4. Two independent expectation estimates each have variance 0.001. Find the linear extrapolation variance.

Solution / guidance

$4(0.001) + 0.001 = 0.005$, with standard deviation $\sqrt{0.005} \approx 0.0707$. Reduced bias does not guarantee smaller total mean-squared error.

8.5. Why must a tomography estimate with $(x, y, z) = (0.8, 0.8, 0.3)$ be repaired or reconsidered?

Solution / guidance

Its squared radius is $0.64 + 0.64 + 0.09 = 1.37 > 1$. One eigenvalue is negative. Sampling fluctuations, inconsistent data, or model errors must be addressed rather than interpreting it as a physical state.

CHAPTER 9

Oracle algorithms: Deutsch, Deutsch–Jozsa, and Bernstein–Vazirani

Learning goals. Define a reversible oracle, derive phase kickback, and compare three algorithms without confusing query complexity with runtime.

9.1 A black box is an explicit resource assumption

For a Boolean function $f : \{0,1\}^n \to \{0,1\}$, the standard oracle is

$$U_f|x\rangle|y\rangle = |x\rangle|y \oplus f(x)\rangle.$$

It preserves x and is reversible even when f is not. An oracle problem counts calls to this operation as units. A physical implementation must still compute f coherently, route its controls, and uncompute temporary work. A black-box separation can teach a mechanism without being a practical application.

Prepare the target in $|-\rangle = (|0\rangle - |1\rangle)/\sqrt{2}$. Since $X|-\rangle = -|-\rangle$,

$$\begin{aligned} U_f|x\rangle|-\rangle &= \frac{|x\rangle|f(x)\rangle - |x\rangle|1 \oplus f(x)\rangle}{\sqrt{2}} \\ &= (-1)^{f(x)}|x\rangle|-\rangle. \end{aligned}$$

The target returns to the same state and the phase appears on the input branch. This **phase kickback** implements $O_f|x\rangle = (-1)^{f(x)}|x\rangle$.

9.2 Deutsch's one-bit problem

There are four one-bit Boolean functions. Two are constant; two are balanced. The task is to determine $f(0) \oplus f(1)$. One classical deterministic query cannot distinguish a constant function from a balanced one agreeing at the queried input.

Prepare $|+\rangle|-\rangle$, query, and apply H to the first qubit. Ignoring the unchanged target,

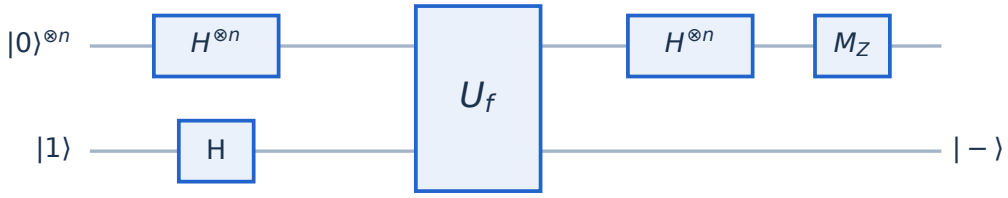


Figure 9.1: A shared circuit pattern for the three oracle algorithms. The work qubit begins in 1 and becomes minus after H; the reversible oracle imprints its function as a phase on the input register.

$$\frac{(-1)^{f(0)}|0\rangle + (-1)^{f(1)}|1\rangle}{\sqrt{2}} \xrightarrow{H} \frac{(-1)^{f(0)} + (-1)^{f(1)}}{2}|0\rangle + \frac{(-1)^{f(0)} - (-1)^{f(1)}}{2}|1\rangle.$$

If the values agree, the second amplitude cancels and result 0 is certain. If they differ, the first amplitude cancels and result 1 is certain. The algorithm reveals parity, not both function values. One query's information is reorganized through interference [11].

9.3 The Hadamard transform on bit strings

For one bit,

$$H|x\rangle = \frac{1}{\sqrt{2}} \sum_{y=0}^{1} (-1)^{xy}|y\rangle.$$

Taking tensor products yields

$$H^{\otimes n}|x\rangle = \frac{1}{\sqrt{2^n}} \sum_{y} (-1)^{x\cdot y}|y\rangle,$$

where $x \cdot y = \bigoplus_j x_j y_j$ is a binary dot product. The exponent may be computed modulo two. This is the Fourier transform of the group of bit strings under XOR. The useful cancellation identity is

$$\sum_{x} (-1)^{x\cdot z} = \begin{cases} 2^n, & z = 0^n, \\ 0, & z \neq 0^n. \end{cases}$$

If a bit of z is one, pair each x with the string obtained by flipping that bit. The two terms have opposite signs and cancel. This pairing proof is the engine behind both algorithms below.

9.4 Deutsch–Jozsa and its promise

The promise is that f is either constant or balanced over all 2^n inputs. Other functions are outside the problem. Starting in $|0^n\rangle$, apply Hadamards, the phase oracle, and Hadamards again. The amplitude of output y is

$$A(y) = \frac{1}{2^n} \sum_x (-1)^{f(x)+x\cdot y}.$$

At $y = 0^n$, this is the average of $(-1)^{f(x)}$. A constant function gives amplitude $+1$ or -1. A balanced function gives zero. Therefore observing all zeros means constant and any nonzero output means balanced, under the promise [12].

An exact deterministic classical algorithm can require $2^{n-1} + 1$ queries: after seeing the same value on half the domain, both cases remain possible. But a randomized classical algorithm can distinguish the cases with bounded error using only a number of samples logarithmic in the inverse error tolerance. For k independent uniformly sampled inputs to a balanced function, the probability all answers match is 2^{1-k}. The dramatic exact deterministic comparison is not an exponential separation from bounded-error classical computation.

9.5 Bernstein–Vazirani

Now promise $f_s(x) = s \cdot x$ for a hidden n-bit string s. The same circuit has

$$A(y) = \frac{1}{2^n} \sum_x (-1)^{s\cdot x + x\cdot y} = \frac{1}{2^n} \sum_x (-1)^{x\cdot(s\oplus y)}.$$

The cancellation identity gives $A(y) = 1$ if $y = s$, zero otherwise. One quantum oracle query recovers the whole string. Classically, querying each unit vector e_j reveals s_j, requiring n deterministic queries in the worst case [13].

An added constant offset b in $f(x) = s \cdot x \oplus b$ contributes only global phase $(-1)^b$. This circuit still returns s but does not identify b. This is a concrete example of a global phase carrying no accessible measurement information.

The BV input state after the oracle actually factors:

$$\frac{1}{\sqrt{2^n}} \sum_x (-1)^{s\cdot x} |x\rangle = \bigotimes_j \frac{|0\rangle + (-1)^{s_j}|1\rangle}{\sqrt{2}}.$$

Entanglement among these input qubits is unnecessary for this query advantage. Claims that every speedup requires entanglement in every stage would therefore be misleading.

Laboratory L13 — Oracle interference. Choose Deutsch's complete truth table, a promised Deutsch–Jozsa instance, or a BV secret. Inspect preparation, oracle phases, final amplitudes, and sampled results. The DJ panel uses a stated family of balanced parity functions; the proof applies to every promised balanced function.

9.6 Exercises

9.1. Run Deutsch algebraically for $f(0) = 1, f(1) = 0$.

Solution / guidance

The queried state is $(-|0\rangle + |1\rangle)/\sqrt{2} = -|-\rangle$. H gives $-|1\rangle$, so result 1 is certain.

9.2. For n=3 and BV secret 101, list the inputs where f is one.

Solution / guidance

$f(x) = x_0 \oplus x_2$. The inputs are 001, 011, 100, 110. Their amplitudes acquire minus signs; the final result is 101.

9.3. In Deutsch–Jozsa, suppose f is one on exactly one quarter of the domain. What is the all-zero output probability?

Solution / guidance

The amplitude is $(3/4) - (1/4) = 1/2$, giving probability $1/4$. This is outside the promise, so the circuit is not an exact classifier for this case.

9.4. How many independent classical samples make $2^{1-k} \leq 1/128$?

Solution / guidance

$1 - k \leq -7$, so k=8 suffices. This bound assumes uniformly random inputs and the exact promise.

9.5. Why does replacing $|-\rangle$ with $|+\rangle$ in the oracle target remove the useful kickback?

Solution / guidance

$X|+\rangle = |+\rangle$, so both f values leave the target and phase unchanged. The input returns to $|0^n\rangle$ after the final Hadamards.

CHAPTER 10

Simon’s algorithm and hidden XOR structure

Learning goals. Follow a two-register oracle calculation, derive the orthogonality constraint, and recover a hidden string by binary linear algebra.

10.1 A stronger oracle separation

Simon’s promise is a nonzero string $s \in \{0,1\}^n$ and a function satisfying

$$f(x) = f(x') \quad \Longleftrightarrow \quad x' = x \text{ or } x' = x \oplus s.$$

Every output therefore has exactly two preimages. An oracle computes f into a sufficiently large output register. The problem is to determine s. The function’s outputs need not have a useful numerical meaning; their equality structure matters.

With randomly labeled outputs, a classical randomized procedure generally needs on the order of $2^{n/2}$ queries to find a collision and hence recover s. The birthday scale arises because k samples contain about $k(k-1)/2$ pairs. Simon’s quantum procedure obtains useful linear constraints with polynomially many queries [14].

This is an oracle-model result. If an explicit function description exposes s directly, there may be an easier classical method. The laboratory intentionally exposes the small oracle table for education; it does not claim its own visible table is a hidden real-world problem.

10.2 Prepare, query, and condition

Start in $|0^n\rangle|0^m\rangle$, apply Hadamards to the first register, then query:

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

If the output register is measured and the value is f(z), the input collapses to

$$\frac{|z\rangle + |z \oplus s\rangle}{\sqrt{2}}.$$

We need not actually perform or retain this measurement. Tracing out the output register gives the same input measurement probabilities after later operations. The conditioned description is simply convenient for deriving them.

Apply $H^{\otimes n}$:

$$\frac{1}{\sqrt{2^{n+1}}} \sum_y \left[(-1)^{z\cdot y} + (-1)^{(z\oplus s)\cdot y}\right] |y\rangle.$$

Because $(z \oplus s) \cdot y = z \cdot y \oplus s \cdot y$, the amplitude factors as

$$\frac{(-1)^{z\cdot y}}{\sqrt{2^{n+1}}} \left[1 + (-1)^{s\cdot y}\right].$$

If $s \cdot y = 1$, the bracket is zero. If $s \cdot y = 0$, it is two. Thus y is uniformly distributed among the 2^{n-1} strings orthogonal to s over the binary field. The unknown z affects only signs, not the measured probabilities.

10.3 Binary Gaussian elimination

Each run supplies one equation $y \cdot s = 0$. Collect independent rows into a matrix Y and solve $Ys = 0$ over $\mathbb{F}_2$, where addition and subtraction are XOR. Once the rank reaches n−1, the nullspace contains exactly 0^n and the promised nonzero s.

Worked example. Let n=3 and s=101. The valid measurement strings are 000, 010, 101, 111. Suppose we obtain 010 and 101. The equations are

$$s_1 = 0, \qquad s_0 \oplus s_2 = 0.$$

The free variable $s_2 = t$ gives $(s_0, s_1, s_2) = (t, 0, t)$. Excluding the zero vector forces t=1, giving 101. The row 111 is the XOR of 010 and 101, so it adds no independent information.

After rank r has been collected in the $(n-1)$-dimensional valid subspace, the next random sample is independent with probability $1 - 2^{r-(n-1)}$. The expected sample count to reach full rank is the sum of the reciprocals of these probabilities. It is $(n-1)$ plus a bounded constant, hence O(n). The zero row and repeated rows are normal, not simulator failures.

10.4 A concrete reversible oracle

For the laboratory, define $f_s(x) = \min(x, x \oplus s)$ using the integer order of bit strings. This maps each XOR pair to the same label and has no other collisions. A reversible oracle XORs that label into its output register.

The simulator calculates the exact input marginal by grouping equal output labels and applying the Hadamard transform. It does not simply print the secret as an answer. The recovery display uses only the sampled equations; the known secret is available separately for checking the experiment.

Laboratory L14 — Simon's hidden mask. Select a nonzero mask, examine oracle pairs, sample valid equations, and inspect row reduction and candidate masks. Reduce the shot count until several candidates remain.

10.5 Failure modes and interpretation

Noise can produce y values with $s \cdot y = 1$. Naively demanding that every equation be exact may then leave no nonzero solution. Recovering a mask from noisy constraints becomes a statistical inference problem. The noiseless proof does not automatically supply a robust experimental algorithm.

The all-zero mask belongs to a different one-to-one promise case in common formulations. This laboratory restricts to the nonzero two-to-one case and rejects zero. A valid interface should explain a promise violation rather than return a confident but meaningless result.

Simon's importance is structural: coherent queries convert a hidden symmetry into measurement constraints. Shor's algorithm later uses a related idea with periodicity under ordinary modular arithmetic, but the transforms and classical reconstruction differ.

10.6 Exercises

10.1. List valid y values for n=3 and s=011.

Solution / guidance

The constraint is $y_1 \oplus y_2 = 0$, so y is 000, 011, 100, or 111, each with probability $1/4$.

10.2. For n=4, do rows 1100, 0110, and 1010 have rank three?

Solution / guidance

No. The third is the XOR of the first two. Rank is two, so the nullspace has dimension two and four candidates including zero.

10.3. Solve equations from rows 110, 011 in three bits.

Solution / guidance

$s_0 = s_1$ and $s_1 = s_2$. The nonzero solution is 111.

10.4. Why does measuring the output register not reveal s in one shot?

Solution / guidance

It reveals one output label, whose two preimages remain unknown. The useful quantum state is a coherent pair of inputs. The subsequent transform turns that pair into an equation on s.

10.5. At rank n−2, what is the probability that the next sample completes the rank?

Solution / guidance

$1 - 2^{(n-2)-(n-1)} = 1/2$. A few redundant samples near completion are expected.

CHAPTER 11

Grover search and amplitude amplification

Learning goals. Implement two reflections, derive the two-dimensional rotation, choose iteration counts, and include oracle and sampling costs.

11.1 The search problem

A predicate marks M of N candidates. We can coherently implement $O|x\rangle = -|x\rangle$ for marked x and $|x\rangle$ otherwise. For simplicity let $N = 2^n$ and prepare

$$|s\rangle = \frac{1}{\sqrt{N}} \sum_{x=0}^{N-1} |x\rangle.$$

One measurement of this state succeeds with probability M/N. Grover's algorithm increases that probability using the oracle and the diffusion operation

$$D = 2|s\rangle\langle s| - I.$$

D is a reflection: it keeps $|s\rangle$ and reverses components orthogonal to it. It is unitary because $D^2 = I$. It can be implemented using Hadamards and a phase operation on $|0^n\rangle$, up to global phase.

For amplitudes a_x, define their mean $\bar{a} = N^{-1} \sum_x a_x$. Then

$$(D|\psi\rangle)_x = 2\bar{a} - a_x.$$

The "reflection about the mean" description refers to amplitudes, which may be complex; it does not reflect probabilities.

11.2 Four-item worked example

Let N=4 and mark 10. Start with amplitudes $(1, 1, 1, 1)/2$. The oracle gives

$$\left(\tfrac{1}{2}, \tfrac{1}{2}, -\tfrac{1}{2}, \tfrac{1}{2}\right), \qquad \bar{a} = \frac{\frac{1}{2} + \frac{1}{2} - \frac{1}{2} + \frac{1}{2}}{4} = \frac{1}{4}.$$

Diffusion maps an unmarked amplitude to $2(1/4) - 1/2 = 0$, and the marked one to $2(1/4) - (-1/2) = 1$. Thus one iteration gives $|10\rangle$ exactly. This special small case illustrates interference but is not the general iteration count [15].

11.3 Rotation in a two-dimensional subspace

Define normalized uniform marked and unmarked states $|G\rangle, |B\rangle$. Then

$$|s\rangle = \sin\theta|G\rangle + \cos\theta|B\rangle, \qquad \sin^2\theta = M/N.$$

The oracle reverses the good component. Diffusion reflects about $|s\rangle$. Their product DO rotates the state in this plane by 2θ toward the marked subspace. After k iterations,

$$|\psi_k\rangle = \sin((2k+1)\theta)|G\rangle + \cos((2k+1)\theta)|B\rangle,$$
$$P_k = \sin^2((2k+1)\theta).$$

Choose a nonnegative integer k near $\pi/(4\theta) - 1/2$ and compare the neighboring integer values. When $M \ll N$, this is approximately $(\pi/4)\sqrt{N/M}$ iterations. Additional iterations can rotate past the target and reduce success.

For N=8,M=1, after one iteration

$$\sin(3\theta) = 3\sin\theta - 4\sin^3\theta = \frac{3}{\sqrt{8}} - \frac{4}{8\sqrt{8}} = \frac{5}{2\sqrt{8}},$$

so $P_1 = 25/32$. After two, $P_2 = 121/128$. The algorithm does not monotonically improve forever.

Figure 11.1: Grover success probability oscillates with iteration count.

11.4 General amplitude amplification

Suppose an algorithm A prepares $A|0\rangle = \sqrt{p}|G\rangle + \sqrt{1-p}|B\rangle$. Replace the uniform-state reflection by $A(2|0\rangle\langle 0| - I)A^\dagger$. Alternating it with a phase flip of successful outcomes amplifies the initial success amplitude. The cost includes calls to both A and $A^\dagger$ as well as the success test [16].

Unknown M complicates choosing k. Randomized iteration schedules, counting methods, or fixed-point variants can address this, with their own guarantees and costs. Applying the known-M formula to an unknown success rate is not an exact algorithm.

11.5 What the speedup does and does not establish

Grover achieves a quadratic improvement in black-box query complexity and is optimal in that setting [17]. A practical runtime includes preparation, reversible predicate computation, workspace, fault tolerance, and repetitions. Searching an ordinary database is not automatically one cheap quantum query: data access must be defined.

For an n-bit key with an efficient verification predicate, unstructured quantum search has the query scale $2^{n/2}$. That observation is not a complete cryptanalytic resource estimate. Parallelism, oracle circuits, reversible memory, and error-correction overhead affect the physical cost.

Grover also does not solve NP-complete problems in polynomial time in general. Replacing a 2^n exhaustive search by $2^{n/2}$ remains exponential in n. Structure-specific classical or quantum algorithms may perform differently; the black-box lower bound does not apply to every structured task.

Laboratory L15 — Amplitude amplification. Choose register size, marked items, iterations, shots, and seed. Inspect signed amplitudes after each iteration and the complete success curve. Deliberately over-rotate.

11.6 Exercises

11.1. For N=4,M=1, calculate success after two iterations.

Solution / guidance

$\theta = \pi/6$, so $P_2 = \sin^2(5\pi/6) = 1/4$. The second iteration undoes the perfect success of the first.

11.2. For M=N/2, does the ordinary iteration improve success?

Solution / guidance

$\theta = \pi/4$, so $\sin^2((2k+1)\pi/4) = 1/2$ for every integer k. This reflection schedule cannot improve it.

11.3. Prove D preserves the vector norm.

Solution / guidance

Let $P = |s\rangle\langle s|$, with $P^2 = P = P^\dagger$. Then $D^\dagger D = (2P - I)^2 = 4P^2 - 4P + I = I$.

11.4. A state preparation costs 100 gates, its inverse costs 100, and a success oracle costs 500. Estimate gate work for 20 amplification iterations, excluding reflection overhead.

Solution / guidance

$20(100 + 100 + 500) = 14{,}000$ gates, plus the initial preparation and omitted reflections. Counting only 20 oracle queries would conceal most details.

11.5. With independent success probability P per run, how many runs make failure at most δ?

Solution / guidance

Failure is $(1 - P)^r$. For $0 < P < 1$, choose $r \geq \log \delta / \log(1 - P)$ and round up. Verification of the proposed answer must also be included.

CHAPTER 12

Fourier transforms and phase estimation

Learning goals. Calculate a quantum Fourier transform, track bit reversal, derive the phase-estimation distribution, and account for controlled powers.

12.1 The QFT is a unitary change of amplitudes

For $N = 2^n$, define

$$F_N|x\rangle = \frac{1}{\sqrt{N}} \sum_{y=0}^{N-1} e^{2\pi i x y/N} |y\rangle.$$

The inverse uses the negative sign in the exponent. To verify unitarity, take the overlap of columns x and x':

$$\frac{1}{N} \sum_{y=0}^{N-1} e^{2\pi i (x-x')y/N} = \delta_{xx'}.$$

For x≠x', the ratio $r = e^{2\pi i(x-x')/N}$ obeys $r^N = 1$ but r≠1, so the geometric sum $(1 - r^N)/(1 - r)$ vanishes.

Worked example.

$$F_4|1\rangle = \frac{|0\rangle + i|1\rangle - |2\rangle - i|3\rangle}{2}.$$

The four probabilities are equal. The transform has encoded x in relative phases. Measuring immediately in Z does not reveal those phases or return a table of classical Fourier coefficients.

For $|\psi\rangle = \sum_x a_x |x\rangle$, the output amplitudes are $b_y = N^{-1/2} \sum_x a_x e^{2\pi i x y/N}$. A QFT circuit can transform a register with O(n²) elementary controlled-phase and Hadamard gates, but generic amplitude loading and complete output read-out can cost exponentially in n. Comparing this operation directly with an FFT that writes all N coefficients is an input-output mismatch.

12.2 Building the circuit and tracking order

Write x in binary. Factoring the phase across output bits gives tensor factors with binary-fraction phases. One standard construction applies H to a wire, controlled phase rotations from later wires, repeats down the register, and finally reverses wire order with swaps.

For two qubits in our most-significant-first convention, a forward QFT is: H on q_0; controlled phase $\pi/2$ between q_1 and q_0; H on q_1; then SWAP. The controlled phase multiplies only $|11\rangle$ by i. Omitting final swaps implements a bit-reversed output convention, not the exact F_4 defined above.

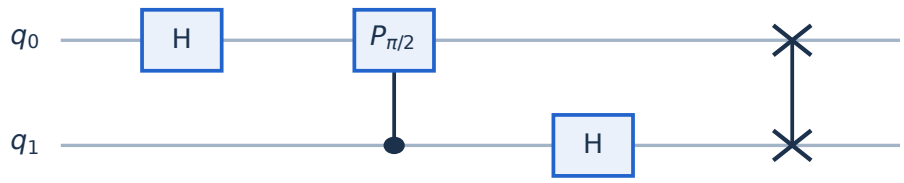


Figure 12.1: An exact two-qubit forward QFT in the book's wire order. The controlled phase contributes i only to the 11 component; the final swap fixes output order.

Small controlled rotations can sometimes be omitted for an approximate QFT with a bounded error. A valid approximation statement specifies the norm or algorithmic success criterion, not just the number of gates removed.

Laboratory L16 — Fourier transform. Prepare a basis state or an editable periodic amplitude pattern. Inspect complex output amplitudes, switch forward/inverse sign, and apply the inverse to verify recovery.

12.3 Phase estimation derives a measurable phase

Suppose $U|u\rangle = e^{2\pi i\phi}|u\rangle$, with $0 \le \phi < 1$. Use an m-qubit control register, $Q = 2^m$, prepared uniformly. Controlled powers $U^{2^{m-1-j}}$ act from control wire j under our ordering. Phase kickback gives

$$\frac{1}{\sqrt{Q}} \sum_{x=0}^{Q-1} e^{2\pi i\phi x}|x\rangle|u\rangle.$$

Apply $F_Q^\dagger$ to the controls. The amplitude of y is

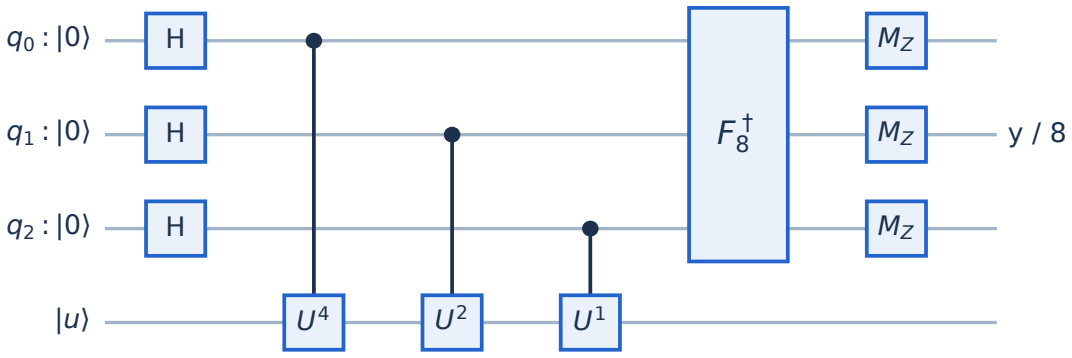


Figure 12.2: Phase estimation with three control qubits. Their powers are 4, 2, and 1 from top to bottom. The inverse QFT converts the kicked-back phase into a binary estimate.

$$A_y = \frac{1}{Q} \sum_{x=0}^{Q-1} e^{2\pi i x(\phi - y/Q)}.$$

If $\phi = k/Q$ exactly, orthogonality gives output y=k with certainty. Otherwise let $\delta = \phi - y/Q$. The geometric sum gives

$$p(y) = |A_y|^2 = \frac{\sin^2(\pi Q \delta)}{Q^2 \sin^2(\pi \delta)},$$

using the continuous limiting value one when δ is an integer. The distribution peaks near y/Q≈φ, including wraparound near zero.

Worked exact phase. Let m=3 and $\phi = 3/8 = 0.011_2$. The phase-kicked control amplitudes are $e^{2\pi i 3x/8}/\sqrt{8}$. Inverse QFT returns integer three, bit string 011. The eigentarget stays in $|u\rangle$.

For a target superposition $\sum_j c_j |u_j\rangle$, linearity creates correlated phase-register and eigenstate branches. Resolvable eigenphases appear with probabilities approximately $|c_j|^2$. Phase estimation does not generally return the expectation of U's phase; it samples eigencomponents.

Laboratory L17 — Phase estimation. Adjust φ and control precision. Inspect the phase-kicked amplitudes, inverse-QFT distribution, sampled estimates, and circular error. Compare exactly representable and nearby phases.

12.4 Precision has a physical cost

m control qubits provide grid spacing 2^{-m}. But the controlled powers collectively use U a total of $1 + 2 + \cdots + 2^{m-1} = 2^m - 1$ times if implemented by repetition. Thus increasing precision is not free because the QFT has only polynomial gate count.

Sometimes controlled powers can be implemented more efficiently using known structure, as in modular arithmetic. For Hamiltonian evolution, long controlled evolution time remains a resource. Iterative phase estimation can reduce simultaneous ancilla qubits by using repeated experiments and feed-forward, but does not remove the fundamental need for sufficient phase information [18, 19].

The eigenphase is only defined modulo one. When converting a Hamiltonian energy E using $U = e^{-iHt}$, $\phi = -Et/(2\pi)$ modulo one. The energy range and evolution time must avoid ambiguity or supply a known offset. A phase estimate without its sign and units is incomplete.

12.5 Exercises

12.1. Calculate $F_4|2\rangle$.

Solution / guidance

The phases are $e^{i\pi y} = (-1)^y$, so the amplitudes are $(1, -1, 1, -1)/2$.

12.2. What does inverse QFT do to a uniform state?

Solution / guidance

It returns $|0\rangle$, since the sum of all nontrivial roots of unity vanishes.

12.3. For m=4 and $\phi = 11/16$, give the certain bit string in our convention.

Solution / guidance

Integer 11 is 1011. The leftmost bit has weight eight.

12.4. How many repeated U applications implement all controlled powers for m=10?

Solution / guidance

$2^{10} - 1 = 1023$, excluding control overhead, preparation, QFT, and repetitions.

12.5. A target is $(|u_0\rangle + |u_1\rangle)/\sqrt{2}$ with exactly resolvable phases 1/4 and 3/4. What does phase estimation measure?

Solution / guidance

Each phase occurs with probability $1/2$, and the target is conditioned on the associated eigenstate. It does not return their mean phase 1/2.

CHAPTER 13

Period finding and Shor's factoring algorithm

Learning goals. Work modular arithmetic and continued fractions, connect periodicity to Fourier peaks, and distinguish a small demonstration from cryptographic-scale factoring.

13.1 Number theory needed for the reduction

The greatest common divisor $\gcd(a, N)$ is computed efficiently by Euclid's algorithm. Repeatedly replace (a, N) by $(N, a \bmod N)$ until the remainder is zero. Modular exponentiation also has an efficient repeated-squaring algorithm: represent the exponent in binary and multiply only the required powers.

For $\gcd(a, N) = 1$, the **order** r is the smallest positive integer with $a^r \equiv 1 \pmod{N}$. Then $a^x \bmod N$ repeats with period r.

If r is even, set $t = a^{r/2} \bmod N$. Since $t^2 \equiv 1 \pmod{N}$,

$$(t-1)(t+1) \equiv 0 \pmod{N}.$$

If $t \not\equiv \pm 1 \pmod{N}$, the gcd values $\gcd(t-1, N)$ and $\gcd(t+1, N)$ provide nontrivial factors. If r is odd or $t \equiv -1$, choose another base. A found gcd greater than one before quantum order finding is already a classical success.

Worked factorization. For N=15,a=2, powers are 1,2,4,8,1, so r=4. Then $t = 2^2 = 4$, and

$$\gcd(4-1, 15) = 3, \qquad \gcd(4+1, 15) = 5.$$

This arithmetic is the end of the algorithm, not its quantum component.

13.2 The quantum state before the Fourier transform

Choose a control range $Q = 2^m$, conventionally $N^2 \le Q < 2N^2$ for the elementary reconstruction guarantee. Compute modular exponentiation reversibly:

$$\frac{1}{\sqrt{Q}}\sum_{x=0}^{Q-1}|x\rangle|0\rangle \longmapsto \frac{1}{\sqrt{Q}}\sum_{x}|x\rangle|a^x \bmod N\rangle.$$

The operation must be a reversible circuit on all allowed basis inputs; computing into a clean work register is a shorthand for a reversible implementation with workspace. Efficient modular arithmetic is essential to Shor's polynomial complexity [20].

Condition on a work-register value. Its compatible x values form a progression $x_0, x_0 + r, \ldots, x_0 + (L-1)r$, truncated at Q. The control state is

$$\frac{1}{\sqrt{L}}\sum_{k=0}^{L-1}|x_0 + kr\rangle.$$

After inverse QFT, the amplitude at y is

$$A_y = \frac{e^{-2\pi i x_0 y/Q}}{\sqrt{LQ}}\sum_{k=0}^{L-1} e^{-2\pi i k r y/Q}.$$

The sum is large when ry/Q is near an integer. Therefore measured y tends to satisfy y/Q≈j/r. The offset contributes a phase but does not move the peaks. Summing probabilities over all possible work-register outcomes gives the unconditional control distribution.

If r divides Q exactly, peaks occur at $y = jQ/r$. For N=15,a=2,Q=256,r=4, these are 0,64,128,192, each with probability $1/4$. This unusually clean example should not be mistaken for the shape of every instance.

13.3 Continued fractions recover a candidate denominator

A rational approximation to y/Q can reveal j/r. Continued fractions successively extract an integer part and invert the remainder. Their convergents supply good rational approximations.

Suppose y/Q=171/512. Euclidean divisions begin

$$512 = 2(171) + 170, \quad 171 = 1(170) + 1, \quad 170 = 170(1).$$

Thus $171/512 = [0; 2, 1, 170]$, with convergents $0/1, 1/2, 1/3, 171/512$. The small denominator three is a candidate period divisor if the observed ratio approximated j/r with reduced denominator three.

A candidate denominator must be checked by modular exponentiation. If j and r share a factor, the reduced denominator can be smaller than r; several samples or tested multiples may be needed. The outcome y=0 supplies no useful denominator. Neither a sharp Fourier peak nor a plausible fraction eliminates the verification step.

Worked nontrivial reconstruction. For N=21,a=2, the true order is six. A peak near y/Q≈1/6 can suggest r=6; one near 1/3 may suggest only a divisor. With verified r=6, $t = 2^3 = 8$, so $\gcd(7, 21) = 7$ and $\gcd(9, 21) = 3$.

13.4 The complete classical–quantum loop

First handle even N, prime powers, and easy gcd cases using classical preprocessing. Select a coprime base, perform order finding, reconstruct and verify candidate orders, and apply the gcd reduction. Repeat when an uninformative sample or unsuitable order occurs. Efficient primality and perfect-power tests are classical; they are not quantum speedups.

The algorithm runs in time polynomial in $\log N$ with suitable fault-tolerant logical arithmetic and precision accounting. That asymptotic theorem is not a claim about today's available machines. Physical feasibility depends on error rates, logical operations, workspace, factories, decoding, and time.

> **Laboratory L18 — Small-instance order finding.** Choose N=15 or 21 and a base. The simulator computes modular exponentiation groups and the exact Fourier marginal, samples y, and shows continued-fraction candidates and gcd checks. A control register too small for the textbook guarantee is explicitly identified.

The laboratory is a classical simulation of the mathematical order-finding experiment. It is not a hardware benchmark or a full gate decomposition of cryptographic modular arithmetic. Unlike a canned "factor 15" animation, its outcome distribution is recomputed from a,N,Q, and a bad sample can fail.

13.5 Modern resource claims

A resource estimate is conditional on an architecture and error model. Gidney's 2025 analysis estimates RSA-2048 factoring with fewer than a million noisy qubits and under a week under its specified superconducting-style assumptions [52]. A March 2026 preprint by Cain and colleagues explores much smaller reconfigurable-atom counts using different coding and architecture assumptions [53]. Neither describes a completed RSA-2048 factorization on a quantum machine. Chapter 20 explains how to read these estimates, and Chapter 24 records their evidence status.

13.6 Exercises

13.1. Find the order of 4 modulo 15 and complete the factorization.

Solution / guidance

$4^1 = 4$, $4^2 \equiv 1$, so r=2. Then t=4 and gcd values are 3 and 5.

13.2. Why is a=14 unsuitable in the final reduction for N=15?

Solution / guidance

Its order is two, but $14^1 \equiv -1 \pmod{15}$. The gcd values are 1 and 15, so the reduction gives no nontrivial factor.

13.3. In the clean N=15,a=2,Q=256 experiment, what does y=128 reveal?

Solution / guidance

$y/Q = 1/2$ suggests denominator two, a divisor of the actual order four. Testing $2^2 \bmod 15 = 4$ rejects it as a full order; testing the multiple four succeeds.

13.4. What minimum power-of-two Q satisfies $Q \geq 21^2$?

Solution / guidance

$21^2 = 441$ and the next power of two is 512, requiring nine control qubits. Work and arithmetic ancillas are additional.

13.5. Why does knowing factors in advance risk invalidating a hardware demonstration of Shor's general algorithm?

Solution / guidance

A circuit can be compiled using the known period or answer, eliminating the hard modular arithmetic. Such a demonstration may test selected gates and interference, but it does not validate general scalable factoring.

CHAPTER 14

Hamiltonian simulation and modern algorithmic tools

Learning goals. Connect unitary evolution to Hamiltonians, derive a product formula, and understand the access and output assumptions behind block encoding and linear-system algorithms.

14.1 Why simulate quantum systems?

A Hamiltonian H is a Hermitian operator specifying a system's energy and dynamics. In units with $\hbar = 1$,

$$i\frac{d}{dt}|\psi(t)\rangle = H|\psi(t)\rangle, \qquad |\psi(t)\rangle = e^{-iHt}|\psi(0)\rangle$$

for time-independent H. Simulating such evolution is a natural quantum-computing task because the target itself has quantum structure. But "simulate a molecule" still needs a precise model: orbitals, truncation, initial state, desired observable, accuracy, and classical benchmark.

For n qubits, a Hamiltonian can be expanded in Pauli strings: $H = \sum_\ell h_\ell P_\ell$, with real coefficients for Hermitian H. A Pauli string is a tensor product of I,X,Y,Z. Since $P_\ell^2 = I$,

$$e^{-ih_\ell P_\ell t} = I\cos(h_\ell t) - iP_\ell \sin(h_\ell t).$$

Basis rotations and parity-computing circuits can implement these individual terms. For a Z-string, CNOTs compute parity into a target, an R_z applies the phase, and reversed CNOTs uncompute the parity.

14.2 Product formulas and their error

If A and B commute, $e^{-i(A+B)t} = e^{-iAt}e^{-iBt}$. Otherwise the equality generally fails. Expand each side for a small step δ:

$$e^{-i(A+B)\delta} = I - i(A+B)\delta - \tfrac{1}{2}(A^2 + AB + BA + B^2)\delta^2 + \cdots,$$
$$e^{-iA\delta}e^{-iB\delta} = I - i(A+B)\delta - \left(\tfrac{1}{2}A^2 + AB + \tfrac{1}{2}B^2\right)\delta^2 + \cdots.$$

The difference is proportional to $[A, B]\delta^2/2$. Repeating r steps of size t/r gives a first-order product-formula error with typical bound scaling $O(t^2\|[A, B]\|/r)$ for bounded operators. A symmetric step $e^{-iA\delta/2}e^{-iB\delta}e^{-iA\delta/2}$ cancels leading asymmetry and improves the order under appropriate assumptions [21, 22].

Worked two-term model. Let $H = X + Z$. Because XZ+ZX=0,

$$H^2 = X^2 + XZ + ZX + Z^2 = 2I.$$

Therefore

$$e^{-iHt} = \cos(\sqrt{2}t)I - i\frac{\sin(\sqrt{2}t)}{\sqrt{2}}(X + Z).$$

Acting on $|0\rangle$ gives amplitudes $\cos(\sqrt{2}t) - i\sin(\sqrt{2}t)/\sqrt{2}$ and $-i\sin(\sqrt{2}t)/\sqrt{2}$. This exact solution is a useful check on an approximate simulator.

Laboratory L19 — Product-formula experiment. Compare exact evolution under X+Z with repeated Z then X rotations. Change time and step count; inspect the state-vector error and fidelity. More steps reduce formula error but would add hardware noise in a physical implementation.

14.3 Energies through phase estimation

For an energy eigenstate $H|E\rangle = E|E\rangle$, $e^{-iHt}|E\rangle = e^{-iEt}|E\rangle$. Phase estimation can infer E after accounting for sign, time, and periodic ambiguity. Preparing a state with sufficient overlap on the desired eigenstate can be as important as simulating evolution. A precise phase estimate of the wrong eigenstate is not a successful ground-state calculation.

The full resource budget combines model error, state-preparation error, Hamiltonian-simulation error, phase-estimation error, and sampling or confidence requirements. Giving every component the entire error tolerance is invalid; the errors must be combined or budgeted with a justified bound.

14.4 Block encoding: matrices inside unitaries

An arbitrary matrix A is not usually unitary, so it cannot simply be applied as a deterministic closed-system gate. A block encoding embeds A/α into a larger unitary U:

$$(\langle 0^a| \otimes I)\, U\, (|0^a\rangle \otimes I) = A/\alpha.$$

Here $\alpha \geq \|A\|$ is a normalization and a is the ancilla count. Other blocks make the complete matrix unitary. Postselecting the ancillas on zero applies the de-

sired block with a state-dependent probability; that probability and the cost of constructing U must be counted.

For a sum $A = \sum_\ell a_\ell U_\ell$ of unitaries, one can prepare an index superposition weighted by $\sqrt{|a_\ell|/\alpha}$, apply a selected U_ℓ with its sign or phase, and unprepare the index. This illustrates how a matrix's representation determines its access cost.

Quantum signal processing and quantum singular value transformation use structured sequences of such unitaries, reflections, and phase rotations to transform eigenvalues or singular values by controlled polynomials. The required polynomial must satisfy boundedness and parity constraints, and its degree controls query cost. These methods unify important simulation, search, and linear-algebra algorithms; they do not provide arbitrary matrix functions for free [23, 24].

14.5 Linear systems and input-output restrictions

For $A|x\rangle \propto |b\rangle$, a quantum linear-system algorithm aims to prepare a state proportional to $A^{-1}|b\rangle$. In A's eigenbasis, inversion multiplies coefficient b_j by $1/\lambda_j$. Small eigenvalues cause large amplification, making the condition number κ an essential parameter [25].

The output is a quantum state, not a printed list of every component of a classical solution vector. Extracting all N entries generally removes the advertised logarithmic dependence on N. Promising use cases specify efficient access to A and b, suitable conditioning, and an output observable obtainable with limited measurements.

A classical comparison must receive comparable data access. Sampling access, sparse-entry oracles, low-rank structure, and preprocessed data can also enable powerful classical algorithms. A quantum linear-algebra primitive is a starting point for an application analysis, not a completed speedup claim.

14.6 Amplitude estimation and learning

Amplitude estimation combines amplification with phase information to estimate a success probability. Ideal coherent-query algorithms can improve the scaling with additive precision from the ordinary sampling $O(1/\epsilon^2)$ to $O(1/\epsilon)$ under appropriate access and success guarantees [16]. The improvement trades repeated shallow sampling for coherent operations, inverses, or more elaborate schedules. Noisy hardware can change that trade.

Quantum machine learning includes learning quantum data, quantum feature maps, and variational models. Loading classical data, generalization, training stability, and classical baselines are separate questions. An exponentially large Hilbert space alone proves no learning advantage. Strong results are

task-specific; broad superiority on ordinary classical datasets remains an open question in this introductory scope [26].

14.7 Exercises

14.1. Derive $e^{-itZ}|+\rangle$.

Solution / guidance

Z's eigenvalues are ±1, so the state is $(e^{-it}|0\rangle + e^{it}|1\rangle)/\sqrt{2}$. Up to global phase, its relative phase is $2t$.

14.2. Why is the product formula exact for $A = Z \otimes I$, $B = I \otimes X$?

Solution / guidance

The operators act on separate subsystems and commute. The commutator and all higher commutator corrections vanish.

14.3. A first-order formula error bound is C/r. What happens when the target error is divided by ten?

Solution / guidance

The bound requires ten times as many steps, holding C fixed. More steps also increase physical operation count; total error may not decrease if hardware noise dominates.

14.4. Why does a block encoding need $\alpha \geq \|A\|$?

Solution / guidance

A subblock of a unitary is a contraction: projecting a unit vector cannot increase its norm. Thus $\|A/\alpha\| \leq 1$.

14.5. A quantum linear-system routine prepares a 2^{20}-dimensional state. Does it output one million classical numbers?

Solution / guidance

No. It outputs a 20-qubit state encoding normalized amplitudes. Recovering a full vector requires an additional, generally expensive measurement and reconstruction task.

CHAPTER 15

Variational algorithms and their limits

Learning goals. Derive a variational energy, use a parameter-shift gradient, explore QAOA, and evaluate claims under noise and classical comparison.

15.1 A hybrid optimization loop

A parameterized circuit prepares $|\psi(\theta)\rangle$. Measurements estimate a cost such as

$$E(\theta) = \langle\psi(\theta)|H|\psi(\theta)\rangle.$$

A classical optimizer chooses new parameters, and the loop repeats. The variational principle follows by expanding the state in energy eigenstates:

$$|\psi\rangle = \sum_j c_j|E_j\rangle, \qquad E(\theta) = \sum_j |c_j|^2 E_j \geq E_{\min}.$$

This guarantee assumes the correct Hamiltonian and a physical normalized state. A finite-shot or mitigated estimate can fall below the exact ground energy due to sampling or estimator bias; that does not violate the theorem.

The variational quantum eigensolver (VQE) uses this loop to seek low-energy states [27]. Its performance depends on the ansatz, initialization, optimizer, measurement allocation, and noise. The variational principle gives an energy bound, not a guarantee that the optimizer reaches the best state.

15.2 A fully solvable example

Take $H = Z + \frac{1}{2}X$ and prepare $|\psi(\theta)\rangle = R_y(\theta)|0\rangle$. Its Bloch vector is $(\sin\theta, 0, \cos\theta)$, so

$$E(\theta) = \cos\theta + \tfrac{1}{2}\sin\theta.$$

Let $R = \sqrt{1 + 1/4} = \sqrt{5}/2$ and $\delta = \arctan(1/2)$. Then

$$E(\theta) = R\cos(\theta - \delta).$$

The minimum is $-\sqrt{5}/2$ at $\theta = \pi + \delta$ modulo 2π. Differentiating gives $E'(\theta) = -\sin\theta + \frac{1}{2}\cos\theta$. The condition $\tan\theta = 1/2$ has both maximum and minimum solutions; checking curvature or energy selects the minimum.

If θ is restricted to $[0, \pi]$ with relative phase zero, the circuit cannot reach every real direction needed for the optimum. Allowing a negative relative phase or the full 2π rotation removes that artificial restriction. Parameter ranges are part of an ansatz definition.

15.3 Parameter-shift gradients

For a parameter appearing once in a rotation $e^{-i\theta P/2}$ with $P^2 = I$, an expectation has the form $a + b\cos\theta + c\sin\theta$. Therefore

$$\frac{dE}{d\theta} = \frac{E(\theta + \pi/2) - E(\theta - \pi/2)}{2}.$$

This is exact under the stated generator and occurrence assumptions. It is not a generic two-evaluation formula for every parameterization. If a shared parameter occurs in many gates, contributions must be treated correctly.

With finite-shot evaluations, the gradient also has uncertainty. Optimizing beyond that uncertainty can waste shots or follow noise. Adaptive measurement allocation and stopping criteria belong in a serious experimental workflow.

Laboratory L20 — Variational energy. Manipulate θ and φ for H=Z+X/2, inspect the state and energy, and take gradient steps on θ. Compare with the exact ground energy. The exact classical solution is displayed because this is a learning problem.

15.4 QAOA as alternating dynamics

For a bit-string objective C(x), define the diagonal operator $C = \sum_x C(x)|x\rangle\langle x|$ and mixer $B = \sum_j X_j$. A depth-p QAOA state is

$$|\gamma, \beta\rangle = e^{-i\beta_p B} e^{-i\gamma_p C} \cdots e^{-i\beta_1 B} e^{-i\gamma_1 C} |+\rangle^{\otimes n}.$$

For MaxCut, an edge (j,k) contributes $(I - Z_j Z_k)/2$, which is one when the bits differ. The cost phase does not change probabilities by itself; the mixer converts its relative phases into a changed output distribution.

For a triangle, every nonconstant assignment cuts two edges and the two constant strings cut zero. The optimum is two. The laboratory exactly simulates

one QAOA layer and shows the entire distribution, so an improved expectation can be distinguished from certainty of the optimal string [28].

Laboratory L21 — A QAOA triangle. Change γ and β, inspect the eight probabilities, and compare expected cut value with the optimum two. The graph and objective are fixed and explicitly shown.

15.5 Trainability, noise, and honest benchmarks

A **barren plateau** is a regime where gradients become very small, often with variance decreasing rapidly with size under a specified random-circuit ensemble. It is not a theorem that every variational circuit is untrainable. Ansatz locality, depth, initialization, cost locality, and data structure matter [29].

An expressive ansatz may approximate many states but require expensive training. A restricted ansatz can be trainable yet miss the desired state. Noise can flatten a cost landscape or bias an optimum. An ideal state-vector optimizer provides information a hardware optimizer must estimate with shots.

A credible benchmark includes total quantum executions, classical optimizer effort, hyperparameter searches, state preparation, mitigation overhead, and the accuracy of the best relevant classical method. Comparing a carefully tuned quantum model with an untuned classical baseline is not evidence of a general advantage.

Annealing and adiabatic algorithms use different control models. In an ideal adiabatic argument, the system follows a changing Hamiltonian's eigenstate when evolution is sufficiently slow relative to spectral-gap and smoothness conditions. A small minimum gap can make the required time large. A device's ability to implement an Ising energy function does not establish efficient solutions to all hard optimization problems.

15.6 Exercises

15.1. Find the eigenvalues of Z+X/2 directly from its matrix.

Solution / guidance

The determinant of $\begin{pmatrix} 1-\lambda & 1/2 \\ 1/2 & -1-\lambda \end{pmatrix}$ is $\lambda^2 - 5/4$. Eigenvalues are $\pm\sqrt{5}/2$.

15.2. Verify the parameter-shift formula for E(θ)=cosθ+sinθ/2.

Solution / guidance

Substitute θ±π/2 and subtract: the cosine contribution gives $-2\sin\theta$, the sine contribution gives $\cos\theta$. Dividing by two matches the derivative.

15.3. What is the expected triangle cut value in the uniform state?

Solution / guidance

Six strings have cost two and two have zero, so the expectation is $(6 \cdot 2)/8 = 3/2$. Each edge also has probability $1/2$ of being cut.

15.4. If β=0 in one-layer QAOA, can γ change computational-basis probabilities?

Solution / guidance

No. The cost unitary is diagonal and changes only phases. Without a later noncommuting operation, the probabilities remain uniform.

15.5. A mitigated VQE energy estimate is below the exact ground energy. List two possible explanations consistent with quantum mechanics.

Solution / guidance

Finite-shot fluctuation and mitigation bias can produce such an estimate. An incorrect Hamiltonian, calibration error, or software error are additional possibilities. The variational bound applies to the exact expectation of a physical state.

CHAPTER 16

Classical simulation and the cost of representation

Learning goals. Estimate exact memory costs, identify efficient special cases, and choose a simulator matched to the circuit and required answer.

16.1 State-vector memory

A generic n-qubit pure state has 2^n complex amplitudes. At 16 bytes per complex double, raw storage is

$$M_{\text{vector}} = 16\,2^n \text{ bytes.}$$

This excludes temporary arrays, circuit metadata, and application overhead. A single-qubit gate updates pairs of amplitudes in O(2^n) work. A circuit with G such gates has naive cost O($G2^n$). A dense general $2^n \times 2^n$ matrix should not be constructed when local gate updates suffice.

Qubits	Raw state-vector memory
20	16 MiB
30	16 GiB
40	16 TiB
50	16 PiB

Each added qubit doubles memory. A generic density matrix contains 4^n complex entries, using $16\,4^n$ bytes. At n=15, that is 16 GiB. Hermiticity reduces independent real parameters but does not remove exponential scaling.

These numbers do not prove that every n-qubit circuit is difficult. They describe a dense representation. A product state needs only n small vectors, and many entangled states have compact alternative representations.

Laboratory L22 — Simulation cost. Change qubit count and numeric precision. Compare state-vector and density-matrix storage on logarithmic scales. No huge arrays are allocated by this calculator.

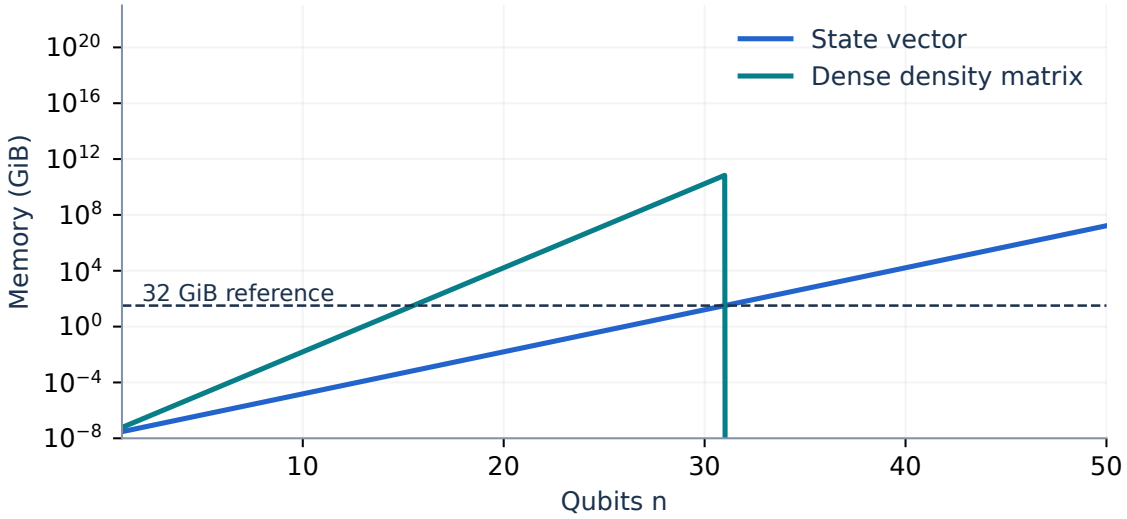


Figure 16.1: Storage for complex128 amplitudes or dense density operators. The plotted values count the array alone, before work buffers and simulator overhead.

16.2 Stabilizer simulation

The Clifford gates map Pauli strings to Pauli strings under conjugation. H exchanges X and Z, S maps X to Y, and CNOT has simple propagation rules. A stabilizer state can be stored by n commuting independent Pauli generators rather than 2^n amplitudes.

A stabilizer tableau has $O(n^2)$ binary entries, and Clifford gates and Pauli measurements admit efficient updates. This is the basis of Gottesman–Knill simulation. It can handle highly entangled states such as GHZ states; entanglement alone is therefore insufficient to imply exponential classical cost [33, 30].

Non-Clifford gates such as T generally break closure of the stabilizer representation. Specialized simulators express a state as combinations of stabilizer states or use related decompositions, with cost depending strongly on non-Clifford resources. A million Clifford operations can be easier to simulate than far fewer appropriately arranged non-Clifford gates.

Stim is an important specialized stabilizer-circuit simulator and detector-error-model tool [32]. Its speed comes from targeting a restricted but useful circuit family. It is not a generic state-vector simulator for arbitrary rotations.

16.3 Tensor networks and entanglement structure

A tensor network factors a large coefficient array into smaller tensors joined by summed indices. For a one-dimensional chain, a matrix product state has amplitudes

$$a_{x_0 \cdots x_{n-1}} = A^{[0]x_0} A^{[1]x_1} \cdots A^{[n-1]x_{n-1}}.$$

The internal matrix dimensions are bond dimensions. A bond dimension χ bounds Schmidt rank across that cut; entanglement entropy is at most $\log_2 \chi$. Low-entanglement dynamics can be represented efficiently. Generic strongly entangling dynamics may force χ to grow exponentially [31].

Truncating small Schmidt coefficients gives an approximate method. Its reliability depends on accumulated truncation error and the requested observables. A visually smooth expectation curve does not establish small global state error.

For more general tensor networks, contraction order and graph structure determine cost. A circuit's depth, geometry, and chosen output can substantially change simulation difficulty. Computing one amplitude, one local expectation, or samples from an approximate distribution are different tasks.

16.4 Trajectories, density matrices, and sampling

A density-matrix simulator exactly averages a specified noise channel but pays 4^n storage. A quantum-trajectory method samples Kraus outcomes and evolves pure states, trading memory for Monte Carlo variance. A single trajectory is a conditioned pure state, not the ensemble density operator.

For an observable, average trajectory estimates and include both trajectory randomness and measurement sampling if the workflow has both. Choosing a convenient Kraus decomposition can affect numerical efficiency without changing the channel average.

Sparse state vectors help while only a few basis amplitudes are nonzero, but Hadamard-rich circuits can quickly destroy sparsity. Distributed simulation spreads arrays across machines at the cost of communication. GPU acceleration changes constants and capacity, not the exponential asymptotic storage law.

16.5 A practical simulator selection process

Ask what output is needed. A circuit with Clifford gates and Pauli noise favors a stabilizer tool. A small arbitrary ideal circuit favors a state vector. A small open system favors a density matrix. A larger structured low-entanglement circuit may favor tensor methods. A noisy observable calculation may favor trajectories.

Then check assumptions. Does the tool support mid-circuit measurement, classical conditions, leakage, non-Pauli noise, qudits, and the intended ordering? Does its noise convention match the paper? An unsupported operation should be rejected or explicitly approximated, never silently reinterpreted.

The browser workbench intentionally limits the generic circuit to eight qubits and 40 operations for responsive teaching. Other laboratories use exact specialized formulas for larger small registers. These user-interface limits are not a statement about the maximum classical simulator capacity.

16.6 Why classical baselines keep moving

A quantum benchmark may motivate improved classical contraction, approximation, or sampling methods. A published comparison is dated evidence about a specified task, not a permanent mathematical lower bound. Claimed quantum advantage should report accuracy, hardware time, preprocessing, classical resources, and whether the baseline was later improved.

Worst-case complexity statements and empirical benchmark records are both useful, but they answer different questions. Chapter 24 keeps them separate when describing current experiments.

16.7 Exercises

16.1. How much raw complex-double memory does a 25-qubit state vector use?

Solution / guidance

$16 \cdot 2^{25} = 2^{29}$ bytes = 512 MiB. Working memory can require multiple such arrays.

16.2. What is the ratio of density-matrix to state-vector raw storage for n qubits?

Solution / guidance

$4^n / 2^n = 2^n$. The density matrix stores relations between every pair of basis states.

16.3. Can a GHZ state have high bipartite entanglement while remaining easy for a stabilizer simulator?

Solution / guidance

It is entangled but has a compact stabilizer description. Across any nontrivial bipartition its Schmidt rank is two and entropy one bit. It is also an example where multiple compact representations exist.

16.4. A matrix product state uses bond dimension 32. What is its maximum entropy across that bond?

Solution / guidance

$\log_2 32 = 5$ bits. Actual entropy can be smaller.

16.5. Why is one amplitude computed by tensor contraction not the same deliverable as a full state vector?

Solution / guidance

It gives one complex number. Repeating the calculation for all 2^n outputs can be exponentially more expensive and may not reuse all work efficiently.

CHAPTER 17

Stabilizers and quantum error correction

Learning goals. Encode an unknown state without cloning it, extract parity syndromes, understand code distance, and derive stabilizer commutation and recovery conditions.

17.1 Protect amplitudes without learning them

An encoding maps one logical qubit into a larger physical space:

$$\alpha|0\rangle + \beta|1\rangle \longmapsto \alpha|0_L\rangle + \beta|1_L\rangle.$$

It does not produce three copies $(\alpha|0\rangle + \beta|1\rangle)^{\otimes 3}$. The logical amplitudes are distributed through a code subspace. Error checks must reveal information about errors while preserving the encoded information.

The three-qubit bit-flip code uses $|0_L\rangle = |000\rangle$, $|1_L\rangle = |111\rangle$. CNOTs from the input to two zero ancillas encode it. A flip on the middle qubit produces $\alpha|010\rangle + \beta|101\rangle$. Measuring every physical bit would reveal the logical value and destroy a general superposition. Instead measure parity.

17.2 Syndrome extraction

The checks are Z_0Z_1 and Z_1Z_2. Every valid code state has eigenvalue +1 for both. A bit flip changes the signs of checks touching its location:

Error	Z_0Z_1	Z_1Z_2	Syndrome bits
I	+1	+1	00
X_0	−1	+1	10
X_1	−1	−1	11
X_2	+1	−1	01

The bit convention is 0 for eigenvalue +1 and 1 for −1. A fresh ancilla initialized to zero receives CNOTs from the two data qubits; measuring it in Z reveals their parity. The two components of the encoded superposition share the same syndrome, so the syndrome does not reveal α or β.

After syndrome 11, apply X_1: $X_1(\alpha|010\rangle + \beta|101\rangle) = \alpha|000\rangle + \beta|111\rangle$. The same correction acts on both amplitudes.

Two flips can mimic one. X_0X_1 has the same syndrome as X_2. Correcting as though X_2 occurred leaves $X_0X_1X_2$, a logical X. Thus under independent flips and perfect checks the failure probability is $3p^2 - 2p^3$. This code does not correct arbitrary phase errors.

Laboratory L23 — Repetition-code memory. Inject a specific bit-flip pattern, inspect neighboring parity checks and the chosen recovery, and compare exact odd-distance failure probabilities. Measurements and recovery are ideal in this model.

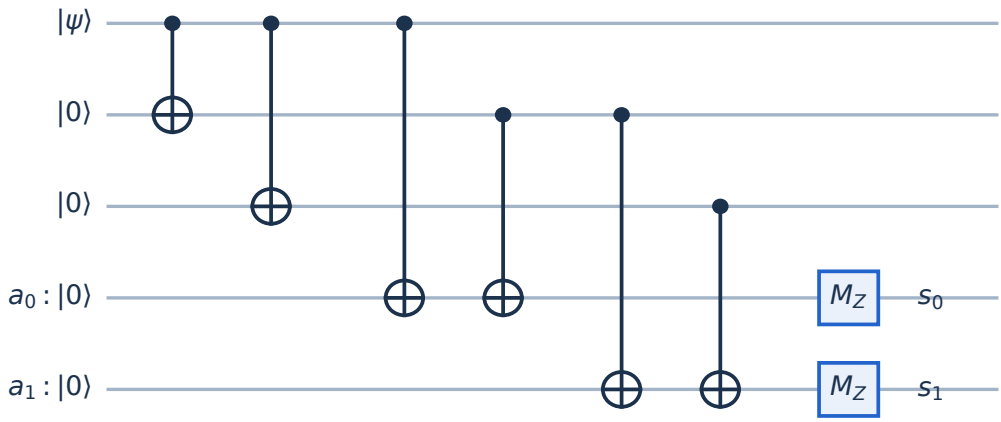


Figure 17.1: Encode an arbitrary input in the three-qubit repetition code, then extract the two Z-parity checks into separate ancillas. This ideal circuit explains the syndrome; it is not a fault-tolerant extraction schedule.

17.3 Why correcting Pauli errors is enough in principle

Any one-qubit operator can be expanded as $aI+bX+cY+dZ$. If a code corrects an appropriate error set, it also corrects coherent combinations in the span under the quantum error-correction conditions. Syndrome extraction separates error sectors without reading the logical amplitudes.

This does not mean hardware errors literally select a Pauli every time. It is a statement about a linear operator basis and recoverability. Coherent noise, correlated multi-qubit errors, leakage, and faulty syndrome circuits still require careful modeling.

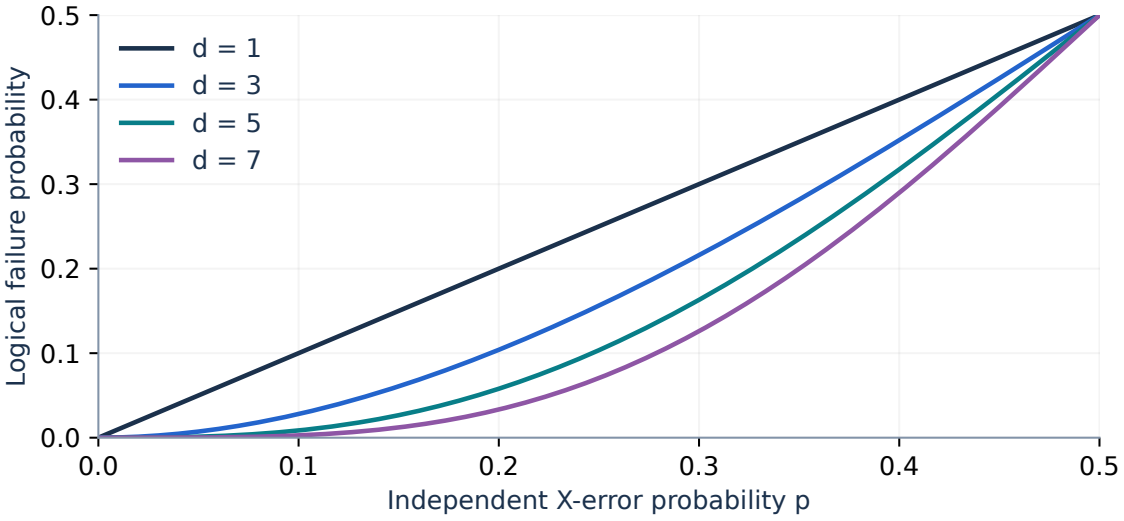


Figure 17.2: Repetition-code failure under independent X errors and perfect recovery. The improvement concerns one error type.

A phase-flip repetition code uses $|{+}{+}{+}\rangle, |{-}{-}{-}\rangle$ and X-type checks. Combining protection against bit and phase errors leads to codes such as Shor's nine-qubit code [34]. Its logical states can be written

$$|0_L\rangle = \left(\frac{|000\rangle + |111\rangle}{\sqrt{2}}\right)^{\otimes 3}, \quad |1_L\rangle = \left(\frac{|000\rangle - |111\rangle}{\sqrt{2}}\right)^{\otimes 3}.$$

The construction protects against any single physical-qubit error, not just a classical bit flip.

17.4 Stabilizer groups and logical operators

A Pauli string is a tensor product of I,X,Y,Z with a phase. Two strings commute if the number of positions where they have different nonidentity Paulis is even. For example, XX and ZZ commute: each position contributes one minus sign, so the total sign is positive.

An abelian stabilizer group S not containing −I defines the common +1 eigenspace of its generators. With r independent generators on n qubits, the code-space dimension is 2^{n-r}, encoding k=n−r logical qubits. Independence matters: a product of existing generators imposes no new constraint [33].

A logical Pauli commutes with all stabilizers but is not itself a stabilizer. For the repetition code, $\overline{X} = X_0X_1X_2$ and $\overline{Z} = Z_0$ are logical operators. Multiplying a logical operator by a stabilizer changes its physical representative but not its action within the code space.

An $[[n, k, d]]$ quantum code has distance d equal to the smallest weight of a nontrivial logical Pauli, for a stabilizer code. It can correct arbitrary errors on up to $\lfloor (d-1)/2 \rfloor$ unknown locations. Known erasures can be corrected on up to d−1 locations under the code's assumptions. The bit-flip repetition code has distance one as a full quantum code because a single Z is an undetectable logical phase operation.

17.5 The Knill–Laflamme condition

Let P project onto the code space and $\{E_a\}$ be a set of errors. A recovery correcting their span exists exactly when

$$PE_a^\dagger E_b P = c_{ab} P.$$

The constants may depend on the error labels but not on the encoded logical state. Intuitively, the environment must not learn which logical state was present while distinguishing correctable error processes [36].

For distinct logical basis states, the condition implies $\langle i_L | E_a^\dagger E_b | j_L \rangle = 0$ when i≠j and equal diagonal values for all i. Error images must preserve logical distinguishability while revealing no logical information through their error-label overlaps.

17.6 CSS codes and a next example

A CSS code uses X-only and Z-only checks specified by binary matrices H_X, H_Z. Commutation requires

$$H_X H_Z^T = 0 \pmod 2.$$

Z-type checks detect X components of errors, and X-type checks detect Z components. A Y has both components. The seven-qubit Steane code is an $[[7, 1, 3]]$ CSS code built from the classical Hamming structure [35]. Unlike the three-qubit bit-flip code, it corrects arbitrary single-qubit Pauli errors.

17.7 Exercises

17.1. Give the syndrome of X_0X_2 in the repetition code and the result of single-error recovery.

Solution / guidance

Both checks flip, giving 11. Recovery applies X_1, leaving $X_0X_1X_2$, a logical X failure.

17.2. Do XZI and ZXI commute? What about XII and ZII?

Solution / guidance

The first pair has two local anticommutations and commutes. The second has one and anticommutes.

17.3. Ten physical qubits have seven independent stabilizer generators. How many logical qubits are encoded?

Solution / guidance

k=10−7=3, so the code space has dimension eight.

17.4. What arbitrary unknown-location errors can a distance-five code correct?

Solution / guidance

Up to two physical-qubit errors. Under an erasure model with known locations, it can correct up to four erasures.

17.5. Why is measuring a parity safer than separately measuring both bits for an encoded superposition?

Solution / guidance

Parity can be identical in both logical components, preserving their coherence within the parity subspace. Separate measurements can distinguish the logical components and reveal encoded information.

CHAPTER 18

Surface codes and decoding

Learning goals. Read a surface-code patch, calculate syndromes, distinguish data and measurement errors, and interpret distance scaling with the correct noise model.

18.1 Local checks encode nonlocal information

Surface codes arrange data qubits and local X- and Z-type parity checks on a two-dimensional layout. Neighboring checks overlap on an even number of data qubits, so they commute. Error strings create changes in nearby checks; logical operators connect appropriate boundaries or wrap nontrivially around the surface [37, 38].

A rotated planar distance-d patch with one measurement ancilla per independent check has d^2 data qubits and d^2-1 syndrome ancillas, totaling $2d^2 - 1$. Extra qubits may be needed for leakage removal, routing, surgery, or spare capacity. This expression is a memory-patch count, not the physical size of an entire useful logical computer.

18.2 A complete distance-three example

Number the nine data qubits row by row, 0–2 on top, 3–5 in the middle, 6–8 on the bottom. The following eight independent commuting generators define the patch used in the laboratory:

Check	Data-qubit Pauli product
X0	X_0X_1
X1	$X_1X_2X_4X_5$
X2	$X_3X_4X_6X_7$
X3	X_7X_8
Z0	Z_3Z_6
Z1	$Z_0Z_1Z_3Z_4$
Z2	$Z_4Z_5Z_7Z_8$
Z3	Z_2Z_5

There are n=9 data qubits and r=8 independent checks, so k=1. Representatives of logical operators are

$$\overline{X} = X_0 X_3 X_6, \qquad \overline{Z} = Z_0 Z_1 Z_2.$$

They commute with every check, anticommute with each other at their one shared location, and have weight three. No weight-one or weight-two nontrivial logical Pauli exists for this code, so its distance is three.

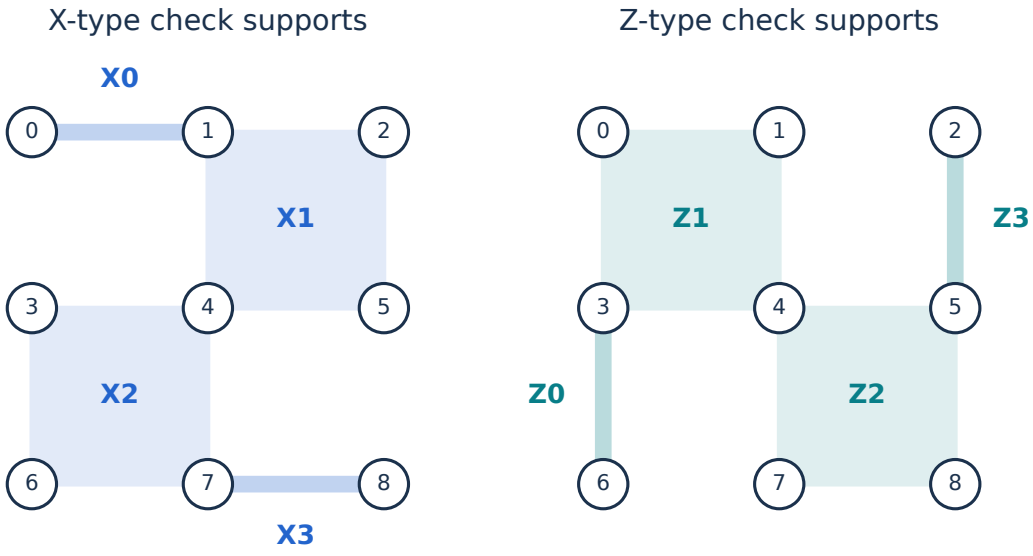


Figure 18.1: The data-qubit layout and local supports of the distance-three patch.

Worked syndrome. An X error on the center data qubit 4 anticommutes with Z1 and Z2 and commutes with the remaining checks. A Z error on 4 anticommutes with X1 and X2. A Y error has both effects. The syndrome contains information about violated parity constraints, not the logical α and β.

18.3 A syndrome does not identify a unique error

If E and E’ differ by a stabilizer, they act identically on the code space and have the same syndrome. If they differ by a nontrivial logical operator, they also have the same syndrome but act differently on the logical information.

A decoder chooses a recovery R based on the syndrome and noise model. Success requires RE to be a stabilizer (up to phase). Requiring R=E is unnecessarily strict because stabilizer-equivalent corrections are equally successful. A zero residual syndrome alone is insufficient: an undetected logical operator also has zero syndrome.

The laboratory uses a minimum-weight Pauli recovery for this fixed patch with perfect measurements and equal costs for X,Y,Z. This is a **code-capacity teach-**

ing model. It is not a full circuit-level surface-code decoder or a maximum-likelihood decoder accounting for degeneracy.

Laboratory L24 — Surface-code syndromes. Click data qubits to cycle through I,X,Y,Z errors. Inspect all eight checks, the chosen minimum-weight recovery, and the residual logical action. Try a full logical string and observe its zero syndrome.

18.4 Noisy checks require time as well as space

In a physical processor, a stabilizer measurement uses several gates and an ancilla that can itself fail. A wrong readout can mimic a data error. Repeating checks lets a decoder distinguish persistent spatial changes from isolated temporal changes.

A **detector** often compares check outcomes across consecutive rounds. Decoding graphs then have both spatial and temporal edges. A chain of data and measurement faults can connect boundaries in spacetime and cause logical failure. Minimum-weight perfect matching is one important decoding approach; correlated decoders, belief propagation, and learning-based methods address other models or structures [39].

The syndrome circuit's gate order matters. An ancilla fault can propagate to multiple data qubits as a hook error. Scheduling must preserve the intended effective distance. A stabilizer table alone does not specify a fault-tolerant extraction circuit.

18.5 Thresholds are properties of complete protocols

Below a suitable physical-noise threshold, increasing code distance can reduce logical error. A common illustrative scaling form is

$$p_L \approx A \left(\frac{p}{p_{\mathrm{th}}} \right)^{(d+1)/2}.$$

Here p summarizes a specified noise model, p_{th} depends on the code, circuit, and decoder, and A is fitted or modeled. This is not a universal exact law. Different physical error types, biased noise, correlations, leakage, and finite-size effects can change it.

If $p/p_{\mathrm{th}} = 0.1$, increasing d by two reduces the model's pL by ten. If $p/p_{\mathrm{th}} > 1$, the same expression increases instead; extrapolating a favorable below-threshold trend into that regime is invalid.

Google's Willow work reported distance scaling below threshold for surface-code memories, with a distance-seven logical error per cycle of approximately 0.143% and a suppression factor about 2.14 when increasing distance by two [47]. The reported experiment used additional hardware beyond the bare patch

and particular decoding methods. Its scope is a memory demonstration, not a completed universal large-scale computation. The 2026 correction notice and publication dates are recorded in Chapter 24.

18.6 Measuring progress toward a logical machine

Useful reports specify code distance, data and ancillary qubits, number of rounds, state bases, decoder, acceptance fraction, error bars, and failure metric. “Logical qubit” can refer to an encoded state, an error-detected state, a repeatedly corrected memory, or a subsystem supporting logical operations. These achievements should not be ranked by a single unqualified logical-qubit count.

A real-time decoder also needs adequate throughput and manageable response latency. A stream can have high throughput while each answer arrives several cycles later. Some operations tolerate delayed Pauli-frame updates; others require a timely classical decision. This motivates the architecture discussion in Chapter 19.

18.7 Exercises

18.1. Find the syndrome of X_0 in the patch.

Solution / guidance

It anticommutes only with Z1. All X checks commute with it, and the other Z supports exclude qubit 0.

18.2. Why does $\overline{X} = X_0X_3X_6$ have zero syndrome?

Solution / guidance

It overlaps Z0 on two qubits and Z1 on two, and the other Z checks on zero. All overlaps are even; X checks commute automatically.

18.3. Count bare data and syndrome qubits for d=5.

Solution / guidance

25 data plus 24 syndrome ancillas equals 49 physical qubits. Additional architecture overhead is outside this count.

18.4. For A=0.1,p/pth=0.1, compare d=3 and d=7 in the illustrative model.

Solution / guidance

$d = 3$ gives $0.1(0.1)^2 = 10^{-3}$; d=7 gives $0.1(0.1)^4 = 10^{-5}$. The two distance increments each supply a factor ten.

18.5. Does “all stabilizers +1 after recovery” prove success?

Solution / guidance

No. The residual may be a logical Pauli with zero syndrome. Compare the residual’s logical commutation or an appropriate logical measurement.

CHAPTER 19

Fault-tolerant computation and modern codes

Learning goals. Distinguish encoding from fault tolerance, propagate circuit faults, understand logical operations and magic states, and compare code families without ignoring architectural costs.

19.1 Encoding alone is not fault tolerance

A code can correct certain data errors when encoding, syndrome measurement, and recovery are perfect. A fault-tolerant protocol must tolerate faults inside those operations too. One fault should not spread into an uncorrectable data pattern faster than the protocol can identify and handle it.

CNOT propagation illustrates the issue:

$$X_c \mapsto X_c X_t, \qquad Z_t \mapsto Z_c Z_t,$$
$$Z_c \mapsto Z_c, \qquad X_t \mapsto X_t.$$

An X on a control spreads forward to its target; a Z on a target spreads back to its control under conjugation. An ancilla interacting with many data qubits can therefore correlate faults. Verified ancillas, flag qubits, careful schedules, or other constructions manage this propagation.

The threshold theorem states, under specified locality and noise assumptions, that sufficiently low physical error permits arbitrarily accurate long computation with controlled overhead. It is an existence and scaling statement, not a guarantee for arbitrary correlated noise or every engineering design [40].

19.2 Logical gates

A **transversal** gate acts separately on corresponding qubits of code blocks, limiting error spread within a block. Some codes support transversal Clifford operations; some support selected non-Clifford gates. A single finite-dimensional error-correcting code cannot have an exactly universal set of transversal logical gates under the usual assumptions [41].

Surface-code logical operations often use lattice surgery, braiding-like code deformations, or state injection. Lattice surgery measures joint logical parities by

temporarily changing check boundaries. Multiple rounds are needed to distinguish data faults from faulty measurements. Joining two patches is not equivalent to directly measuring and revealing each logical qubit.

A Pauli frame stores a classical record of inferred Pauli corrections instead of immediately applying each physical correction. Later measurement results and gates are interpreted accordingly. This can reduce physical operations, but the frame must be updated correctly, particularly when non-Clifford operations depend on it.

19.3 Magic states and a universal gate set

Clifford gates preserve stabilizer structure and are efficiently classically simulable in the stabilizer setting. A universal fault-tolerant architecture needs a non-Clifford resource. A common choice is

$$|T\rangle = T|+\rangle = \frac{|0\rangle + e^{i\pi/4}|1\rangle}{\sqrt{2}}.$$

State injection combines an appropriate ancillary state, Clifford operations, measurements, and corrections to implement a non-Clifford logical operation. The exact circuit depends on conventions and architecture.

Magic-state distillation consumes multiple noisy ancillary states to produce fewer states with lower error, under specified input-noise assumptions and acceptance criteria. For a standard idealized 15-to-1 protocol with independent small input errors ε, the leading output error is approximately $35\epsilon^3$. This scaling is a protocol model, not the cost of a complete factory including faulty logical gates [42].

A factory must also supply states quickly enough. An algorithm with low data-qubit count but high T consumption can be dominated by factories. T-count measures total consumption; T-depth measures sequential non-Clifford layers. Parallel factories exchange extra qubits for reduced waiting time.

19.4 Quantum LDPC codes

A low-density parity-check code has checks of bounded weight and a bounded number of checks touching each qubit in an asymptotic family. Quantum LDPC constructions must also satisfy commutation. The existence of asymptotically good families with constant encoding rate and distance proportional to block size is a major theoretical result [43].

This does not imply a simple two-dimensional nearest-neighbor implementation with the same overhead. Connectivity, syndrome circuits, decoder performance, logical gates, and routing remain part of the problem.

Bivariate bicycle codes provide concrete high-rate examples. Bravyi and collaborators analyzed low-overhead fault-tolerant memories, including a model preserving twelve logical qubits with 288 physical qubits under stated circuit-noise assumptions [44]. This is an analyzed protocol and connectivity proposal; it is not evidence that every 288-qubit chip supplies twelve equally capable fault-tolerant logical qubits.

19.5 Other code families

Color codes place checks on suitably colored lattices and can support attractive transversal logical operations. Their check weights, thresholds, and connectivity differ from surface codes. A favorable gate property can trade against measurement complexity or noise tolerance.

Subsystem codes introduce gauge degrees of freedom. Measuring smaller gauge operators can infer stabilizers without preserving the gauge state. The protected information is in a subsystem, not necessarily the full common eigenspace of every measured operator. Bacon–Shor codes are an instructive example.

Bosonic codes encode in an oscillator's larger Hilbert space. Cat codes use superpositions of separated oscillator states and may exploit biased noise. GKP codes use grid-like phase-space structure to diagnose small displacements. Finite-energy states, loss, imperfect controls, and ancillary hardware determine practical performance [45, 46].

Erasure-aware and loss-biased schemes use knowledge of where a qubit or photon was lost. A known location is more informative than an unknown Pauli error. But identifying loss can itself require reliable measurement and replacement, and a lost atom during an entangling operation may affect partners.

No code is "best" independently of a hardware and computation model. Compare physical costs for the same logical workload, accuracy, connectivity, and time.

19.6 A ladder of experimental claims

An encoded state is the first step. Detecting faults and discarding bad runs is stronger but has an acceptance cost. Repeated correction must preserve information through many rounds. A logical operation must be compared at matched task conditions. Universal, scalable fault-tolerant computation adds non-Clifford resources, long-duration stability, and integrated classical control.

Recent neutral-atom and trapped-ion experiments demonstrate different pieces of this ladder [49, 50]. Their significance should be assessed by the demonstrated operations, postselection, error suppression, and assumptions, rather than collapsing all progress into an undifferentiated logical-qubit count.

19.7 Exercises

19.1. An X fault occurs on an ancilla that controls CNOTs to three data qubits. What risk follows?

Solution / guidance

The X can propagate to each later target, creating a correlated multi-data error. A safe extraction design must prevent, flag, or decode that pattern.

19.2. With ε=0.01, evaluate the ideal leading 15-to-1 output estimate.

Solution / guidance

$35(0.01)^3 = 3.5 \times 10^{-5}$. The result excludes higher-order terms, logical circuit faults, rejection probability, and full factory cost.

19.3. Why can a high-rate code still have large total hardware overhead?

Solution / guidance

Encoding rate counts data qubits per logical qubit. Syndrome ancillas, routing, connectivity devices, factories, spare qubits, and operation schedules add costs.

19.4. Distinguish a logical memory benchmark from a universal-computation benchmark.

Solution / guidance

A memory preserves encoded information. Universal computation must also implement a sufficient logical gate set, manage propagation and classical feed-forward, and supply non-Clifford resources with controlled total error.

19.5. Why should accepted logical fidelity be reported with acceptance probability?

Solution / guidance

Postselection can make accepted runs accurate while most attempts are discarded. Acceptance determines throughput and success per attempt, and affects fair physical-versus-logical comparisons.

CHAPTER 20

Resource estimation: from circuits to machines

Learning goals. Allocate a failure budget, choose a code distance under assumptions, include non-Clifford throughput, and interpret dated architecture estimates.

20.1 Begin with a workload, not a qubit count

A useful resource estimate specifies the algorithm, input size, target observable or answer, numerical precision, allowed failure probability, logical circuit, compilation strategy, hardware model, error-correction protocol, and elapsed-time target. Omitting one can change the result by orders of magnitude.

Distinguish data logical qubits from temporary workspace, routing regions, and magic-state factories. Distinguish logical gate count, depth, and runtime. A "logical qubit" is not a universal conversion unit with a fixed physical cost.

20.2 A failure budget

Suppose an algorithm has L relevant fault locations, each with failure probability at most pL. The union bound gives

$$\Pr(\text{at least one failure}) \leq L p_L.$$

It does not require independence. It can be loose and assumes the locations and per-location risks are correctly defined. If independent identical events are assumed, the exact probability is $1 - (1 - p_L)^L$.

For total failure target ε, a conservative allocation is $p_L \leq \epsilon / L$. With $L = 10^8$ and ε=0.01, the required per-location error is at most 10^{-10}. A physical gate fidelity that looks excellent in isolation can therefore still be far from sufficient for a long algorithm.

Multiple error sources need a shared budget. For example, allocate ε among memory, logical Clifford operations, magic states, algorithmic approximation, and measurement inference. The allocation can be optimized; simply ignoring several components understates the risk.

20.3 Solve an illustrative distance requirement

Use the explicitly assumed model

$$p_L = A(p/p_{\text{th}})^{(d+1)/2}.$$

For p<pth, requiring $Lp_L \le \epsilon$ gives

$$\frac{d+1}{2} \ge \frac{\ln(\epsilon/(LA))}{\ln(p/p_{\text{th}})},$$

where dividing by the negative denominator reverses the intermediate inequality. Thus choose the smallest supported odd d satisfying

$$d \ge 2\frac{\ln(\epsilon/(LA))}{\ln(p/p_{\text{th}})} - 1.$$

Worked estimate. Let A=0.1,p=0.001,pth=0.01,L=10^8, and ε=0.01. Then ε/(LA)=10^{-9} and p/pth=10^{-1}, giving $(d+1)/2 \ge 9$, so d=17. A bare rotated patch uses $2(17)^2 - 1 = 577$ physical qubits.

For 100 data logical patches alone, that is 57,700 physical qubits. It omits factories, routing, spare capacity, leakage-removal hardware, and architecture-specific overhead. The numerical result is an educational model, not a prediction for a named processor.

20.4 Time and throughput

If a logical operation uses approximately d code cycles and the cycle time is τ, a depth-D sequential schedule takes roughly Ddτ. This is only a stated schedule model. Some operations have different costs, some run in parallel, and decoding or communication can add delays.

Suppose D=10^6,d=17,τ=1 microsecond. The model gives 17 seconds. If an algorithm consumes 10^8 magic states and factories jointly produce 10^5 per second, factory supply alone requires at least 1,000 seconds unless buffered or overlapped with other work. The longer bottleneck dominates an optimistic total schedule.

Classical control also matters. Syndrome streams must be decoded at sufficient throughput. A non-Clifford feed-forward decision may impose a latency constraint even when average throughput is adequate. Cooling, calibration, retries, loading, and interconnect heralding can reduce available duty cycle.

Laboratory L25 — Physical and logical resources. Adjust physical error, assumed threshold and prefactor, distance, logical patches, operation locations, cy-

cle time, and depth. Inspect failure bounds, bare patch count, and schedule time. Values above threshold are identified as outside the favorable scaling regime.

20.5 Comparing contemporary estimates

Gidney and Ekerå's 2021 study estimated RSA-2048 factoring in eight hours using about twenty million noisy qubits under a specified superconducting-style architecture [51]. Gidney's 2025 analysis estimated fewer than a million noisy qubits and less than a week with different arithmetic and error-correction arrangements, while retaining stated assumptions including a 0.1% gate error and one-microsecond code cycle [52].

The March 2026 Cain et al. preprint analyzes reconfigurable atomic architectures with high-rate codes and argues that cryptographically relevant Shor instances may be possible at much smaller physical counts [53]. Its headline count must be read alongside the targeted instance, gate times, connectivity, memory, decoding, and runtime assumptions. A small qubit estimate and a fast estimate are not necessarily the same estimate.

These studies are estimates, not demonstrated attacks. Improvements can arise from algorithms, coding, architecture, and scheduling, so extrapolating a single old qubit number as an immutable requirement is scientifically inappropriate. Equally, quoting only the smallest new headline count hides essential conditions.

20.6 Sensitivity and uncertainty

Resource inputs have uncertainty and may not remain stable as a machine scales. A pL model calibrated on small patches may encounter correlated-error floors, fabrication defects, or decoder limitations. Hardware errors differ by gate, qubit, context, and time. A scalar p is a simplification whose adequacy must be assessed.

A good estimate varies uncertain parameters and identifies what drives the result. If halving physical error reduces distance and factory overhead dramatically, improving fidelity may matter more than adding raw qubits. If the workload is communication-limited, faster local gates alone may have little effect.

Avoid false precision. A model producing 57,700 qubits does not imply engineering certainty to the last hundred. Report the model, rounded conclusions, and sensitivity range. Detailed spreadsheets or scripts are useful because they make assumptions inspectable and replaceable.

20.7 Exercises

20.1. For L=10^7 and ε=0.001, what per-location error suffices by the union bound?

Solution / guidance

$p_L \leq 10^{-3}/10^7 = 10^{-10}$.

20.2. Count bare rotated-patch qubits at d=9 and d=11.

Solution / guidance

$2(9)^2 - 1 = 161$ and $2(11)^2 - 1 = 241$. Increasing distance changes cost quadratically even though favorable logical suppression can be exponential in distance.

20.3. An estimate counts ten data logical qubits and ignores twenty factory patches. How much does including those equal-sized patches change its patch qubit count?

Solution / guidance

The count triples from ten to thirty patches. Other overheads may still remain.

20.4. Why is a per-cycle logical error not directly comparable to a per-gate physical error?

Solution / guidance

The tasks, durations, operations, and error channels differ. A fair comparison defines the same information-processing task and includes the full physical resources and timing.

20.5. If the physical error exceeds the threshold, does increasing d in the illustrative formula help?

Solution / guidance

No. The ratio p/pth exceeds one, so the modeled error grows with distance. Moreover the simple below-threshold approximation may be unreliable there.

CHAPTER 21

Physical qubits, control, and hardware benchmarks

Learning goals. Compare physical encodings, identify control and scaling bottlenecks, and read performance metrics without treating qubit count as a complete measure.

21.1 An abstract qubit needs a physical boundary

A mathematical qubit is a two-dimensional system. A physical device usually has more states, interacts with an environment, and requires preparation, gates, measurement, and reset. An implementation chooses two usable states or an encoded subspace and controls unwanted couplings.

A complete processor includes the qubits, control electronics or optics, readout, classical processors, calibration, packaging, interconnects, and thermal or vacuum infrastructure. Counting only the quantum elements understates the engineering system.

Gate-based digital computation, analog quantum simulation, annealing, and photonic sampling devices expose different operations. Results from one model cannot be ranked directly against another without a common task. A large analog array is not automatically a universal circuit computer, although analog systems can be valuable scientific instruments.

21.2 Superconducting circuits

Superconducting qubits use nonlinear electrical circuits, commonly incorporating Josephson junctions. A transmon reduces sensitivity to charge noise while retaining enough anharmonicity to address two levels. Microwave pulses implement rotations; couplers or shared circuit elements produce entangling gates; resonators assist dispersive readout [54, 55].

Key concerns include energy relaxation, dephasing, leakage into higher levels, residual couplings, calibration drift, fabrication variation, and correlated disturbances. Cryogenic packaging and wiring must scale without introducing excessive heat or noise. Fast gates permit short code cycles but impose demanding classical decoding throughput.

A reported T1 or T2 is measured under a particular protocol. It does not directly predict all gate errors in a busy processor. Simultaneous-operation performance, leakage removal, readout, reset, and stability across many rounds matter for error correction.

21.3 Trapped ions

Ions encode qubits in internal atomic levels and are confined using electromagnetic fields. Laser or microwave interactions control individual qubits. Shared motional modes or controlled transport mediate entanglement; fluorescence provides state-dependent readout [56].

Identical atomic species offer reproducible transitions and long-lived internal states. Scaling requires control over motional heating, optical addressing, transport, mode crowding, crosstalk, and measurement while protecting other qubits. A quantum charge-coupled device architecture moves ions among zones for storage, gates, and readout.

The 2026 Helios paper describes a 98-qubit trapped-ion processor with all-to-all connectivity [59]. That connectivity is an architectural capability, not a statement that all pairs can perform independent gates simultaneously at zero transport or scheduling cost. The publication should be consulted for the exact benchmark circuit, fidelities, and timing.

21.4 Neutral atoms

Optical tweezers trap individual neutral atoms. Hyperfine, nuclear-spin, or other internal levels can encode qubits; interactions involving excited Rydberg states can produce entangling gates. Arrays may be rearranged, giving connectivity options different from fixed nearest-neighbor chips [57].

Atom loss, loading, transport, laser stability, Rydberg-state lifetime, crosstalk, and mid-circuit measurement are important. An array can trap many coherent atoms before demonstrating universal gates and error-corrected operations across all of them. Those are distinct milestones.

Bluvstein and collaborators' work in a 2026 Nature issue used arrays of up to 448 atoms to investigate repeated correction, logical operations, and qubit reuse [49]. Its online publication preceded its issue date. Chapter 24 separates the demonstrated circuit scope from the larger architecture it motivates.

21.5 Photonics

Photonic qubits can use polarization, paths, time bins, or other optical degrees of freedom. Photons are natural communication carriers. Linear optics, interference, detection, ancillary photons, feed-forward, nonlinearities, and encoded resource states enable different computational architectures.

The principal challenge is often loss: source efficiency, coupling, propagation, switching, and detection all contribute. Indistinguishability and phase stability are also essential. A probabilistic entangling primitive may require multiplexing or a larger resource-state architecture to achieve useful logical throughput.

Modular photonic experiments study how components and networks can be assembled into larger systems [60]. A component crossing one modeled threshold is not the same as a complete fault-tolerant processor crossing all required thresholds simultaneously.

21.6 Semiconductor spins and other approaches

Electron or nuclear spins in semiconductor devices can use electrically controlled confinement, exchange interactions, and spin-selective readout. Their small physical footprint and relation to semiconductor manufacturing are attractive. Uniformity, charge noise, addressability, valley or orbital states, cryogenic control, and interconnects remain practical issues [58].

Defect centers combine localized spin memories with optical interfaces and are useful in networking and sensing. Bosonic encodings use oscillator states rather than a single two-level element. Topological approaches seek protection from nonlocal degrees of freedom, but demonstrating a particular parity signal is not by itself a demonstration of protected non-Abelian logical computation.

A 2025 Microsoft-associated paper reported interferometric parity measurement in hybrid devices [61]. Interpretation and diagnostic robustness have been disputed in subsequent research [62]. This book does not treat a company announcement or a candidate signature as settled evidence of a scalable topological computer.

21.7 Compare bottlenecks, not a universal winner

Platform	Typical information carrier	Important system questions
Superconducting	Engineered circuit levels	Coherence, leakage, fabrication, wiring, fast feedback
Trapped ion	Internal atomic levels	Motion, transport, addressing, parallelism
Neutral atom	Trapped-atom internal levels	Loss, rearrangement, entangling control, reuse
Photonic	Optical modes or encoded states	Loss, sources, detectors, switching, feed-forward
Semiconductor spin	Confined electronic or nuclear spin	Uniformity, noise, control integration, connectivity

Platform	Typical information carrier	Important system questions
Bosonic / hybrid	Encoded oscillator or coupled degrees	Encoding-specific errors, control, ancillas, readout

These are representative concerns, not fixed rankings. Different platforms can be suitable for different workloads, and improvements in one component may shift the dominant bottleneck elsewhere.

21.8 What a benchmark measures

Gate fidelity compares a realized operation with a target according to a defined averaging convention. It is not the probability every gate is independently perfect. **Randomized benchmarking** fits decay under random gate sequences to infer parameters under assumptions about errors and SPAM. **Process tomography** reconstructs a channel with greater measurement cost and model sensitivity. **Cross-entropy benchmarking** and related sampling metrics target particular distributions and depend on calibration and theoretical assumptions.

Quantum volume, circuit-layer metrics, algorithmic benchmarks, and logical-error rates summarize different tasks. A number without the circuit family, statistical uncertainty, acceptance policy, and runtime definition is incomplete. Error bars should describe how uncertainty was estimated; repeated calibration may reveal drift larger than simple shot uncertainty.

A practical comparison asks for the same task, target accuracy, success probability, total time, and relevant classical baseline. “Advantage” can mean a sampling benchmark, a scientific computation, or economically useful performance. The word should be qualified.

21.9 Exercises

21.1. A device traps 6,000 atoms but demonstrates entangling gates on a subset. What can its atom count establish?

Solution / guidance

It establishes a trapping-scale milestone under the reported conditions. It does not alone establish 6,000 universally controllable or logical qubits.

21.2. Why is a long idle coherence time insufficient to predict logical performance?

Solution / guidance

Logical operation includes gates, readout, reset, leakage management, crosstalk, and decoding. Their errors and correlations may dominate idle decoherence.

21.3. Does all-to-all connectivity mean constant-depth arbitrary all-pairs computation?

Solution / guidance

No. Shared resources, transport, addressing, and gate scheduling can constrain simultaneous operations.

21.4. Why can a photonic design with excellent gate components still have poor end-to-end success?

Solution / guidance

Efficiencies multiply across sources, coupling, propagation, switching, and detection. Loss and probabilistic preparation can dominate the complete protocol.

21.5. What should accompany a claim of improved logical fidelity after postselection?

Solution / guidance

The acceptance fraction, attempted and accepted shot counts, matched baseline, circuit depth, noise conditions, uncertainty, and any discarded event criteria.

CHAPTER 22

Quantum networks, BB84, and post-quantum cryptography

Learning goals. Separate entanglement distribution from classical communication, derive intercept–resend disturbance, and distinguish QKD from post-quantum cryptography.

22.1 Networks distribute a quantum resource

A quantum network connects systems using quantum channels and classical coordination. It may distribute entanglement for teleportation, remote gates, sensing, or key distribution. It does not permit faster-than-light communication. Classical messages remain necessary for heralding, synchronization, correction information, and authenticated protocol steps.

Optical attenuation limits direct transmission. Ordinary repeaters cannot amplify arbitrary unknown quantum states by cloning. Quantum repeaters instead combine entanglement generation, memories, entanglement swapping, and, depending on the architecture, purification or error correction [64, 65].

Entanglement swapping starts with Bell pairs AB and CD. A Bell-basis measurement on BC projects the distant pair AD into a Bell state determined by the measurement result. Communicating the result identifies the Pauli correction. Entanglement need not have existed directly between A and D beforehand.

A remote gate can be constructed using shared entanglement and local operations with feed-forward. Main and colleagues demonstrated distributed quantum computation across an optical link between modules in 2025 [66]. This is a concrete interconnect experiment, not an already deployed general-purpose quantum internet.

22.2 BB84 prepares incompatible alternatives

Alice chooses a random bit a and a random basis, Z or X. She sends $|0\rangle$ or $|1\rangle$ for Z, and $|+\rangle$ or $|-\rangle$ for X. Bob independently chooses Z or X and measures.

Over an authenticated classical channel they announce bases, keeping only matching-basis positions. They do not publicly announce every retained bit. With uniform independent basis choices, about half the signals survive this **sifting** step. In an ideal noiseless channel, matching bases give equal bits [67].

They sacrifice a subset to estimate errors, perform information reconciliation, account for information leaked during reconciliation, and apply privacy amplification. Authentication is essential; without it an attacker could conduct separate sessions with both parties.

22.3 Derive the intercept–resend error

Suppose Eve intercepts every signal, independently chooses Z or X, measures, and sends the resulting state to Bob. Condition on Alice and Bob choosing the same basis.

Eve chooses that basis with probability $1/2$, introducing no error. With probability $1/2$, she chooses the conjugate basis. Bob's result is then random relative to Alice's, giving error probability $1/2$. Therefore

$$Q = \frac{1}{2}(0) + \frac{1}{2}\left(\frac{1}{2}\right) = \frac{1}{4}.$$

If Eve attacks a fraction e of signals, the ideal sifted QBER is e/4. Add an independent bit flip with probability q after the measurement. Combining XOR error events gives

$$Q_{\text{total}} = q + \frac{e}{4} - 2q\frac{e}{4} = q + (1-2q)\frac{e}{4}.$$

The subtraction accounts for two flips canceling. Adding error probabilities without accounting for overlap is only a small-error approximation.

The 25% result concerns this specific attack. It is not a universal threshold guaranteeing security against all adversaries, nor does a low observed QBER prove the absence of side channels [68].

Laboratory L26 — BB84 transmission. Change intercepted fraction, teaching noise, signal count, and seed. Inspect preparation and measurement bases, sifted bits, and QBER. Export the complete simulated transcript.

22.4 From raw correlation to a secret key

An idealized asymptotic single-photon BB84 analysis with suitable symmetric assumptions yields a secret fraction per sifted bit of

$$r \geq 1 - 2h_2(Q), \qquad h_2(Q) = -Q\log_2 Q - (1-Q)\log_2(1-Q).$$

One term accounts for reconciliation and the other for privacy reduction in the ideal bound. The expression becomes zero near Q≈11%. Real finite-key analyses include parameter-estimation uncertainty, imperfect reconciliation, authen-

tication costs, loss, source assumptions, and a composable security parameter [69].

This formula is not a software recipe for production security. The browser uses it only to illustrate the difference between sifted correlation and asymptotic secret fraction. It does not generate a certified secure key.

Weak coherent laser pulses can contain multiple photons. Decoy-state methods help bound relevant contributions under a source model. Detector imperfections and implementation attacks motivate techniques such as measurement-device-independent QKD. Device-independent protocols use Bell tests under demanding assumptions; a simulated CHSH violation is not a security proof for a physical device.

22.5 QKD and PQC address different layers

Quantum key distribution uses quantum signals plus authenticated classical communication to establish key material with a security proof under physical and protocol assumptions. It requires suitable quantum hardware and channels.

Post-quantum cryptography (PQC) consists of classical algorithms designed to resist known classical and quantum attacks. It can be deployed on ordinary computing and communication systems. It does not require a quantum link. Neither “post-quantum” nor “quantum-safe” means security against every imaginable future mathematical discovery.

NIST finalized FIPS 203 (ML-KEM), FIPS 204 (ML-DSA), and FIPS 205 (SLH-DSA) in August 2024 [70]. ML-KEM is a key-encapsulation mechanism, not an encryption mode for arbitrary bulk data. ML-DSA and SLH-DSA are signature schemes. Correct protocol integration, implementations, and key management remain necessary.

Shor threatens the mathematical problems underlying widely used RSA and elliptic-curve systems at sufficient fault-tolerant scale. Grover changes generic search costs and motivates appropriate symmetric-security parameters, but it is not an immediate break of all symmetric encryption. Migration planning must consider long-lived confidential data that may be recorded now and attacked later; this is different from asserting that a cryptographically capable quantum machine already exists.

22.6 Exercises

22.1. With 2,000 ideal signals and independent uniform basis choices, what is the expected sifted count?

Solution / guidance

1,000. Actual counts fluctuate binomially; the expectation is not a guaranteed exact total.

22.2. If Eve attacks 40% of signals with intercept–resend, find ideal QBER.

Solution / guidance

$0.4/4 = 0.1$, or 10%.

22.3. Add independent teaching bit flips q=0.02 to the previous case.

Solution / guidance

$Q = 0.02 + (1 - 0.04)(0.1) = 0.116$, or 11.6%. Simple addition would give 12%, ignoring cancellations.

22.4. Why can Alice and Bob not publish all sifted bits while keeping them as a secret key?

Solution / guidance

Publishing reveals them to an adversary. Only a sacrificed sample is opened for estimation; reconciliation leakage must also be accounted for.

22.5. Match ML-KEM, ML-DSA, and BB84 to their roles.

Solution / guidance

ML-KEM: classical key encapsulation designed for post-quantum security. ML-DSA: classical digital signatures designed for post-quantum security. BB84: quantum-signal key distribution with authenticated classical processing.

CHAPTER 23

Programming reproducible quantum experiments

Learning goals. Run current local Qiskit workflows, convert wire conventions, separate sampling from exact simulation, and prepare reproducible experiments.

23.1 A program is an executable specification

A circuit diagram leaves choices that software must make explicit: register ordering, preparation, parameter units, measurement placement, classical-bit mapping, shot count, random seed, noise model, and output interpretation. These choices should travel with an experiment.

The accompanying Python examples use **Qiskit 2.5.2** and were executed during this edition's validation. This is the tested version, not a promise that it will remain the latest release. The official documentation consulted at the research cutoff uses the modern V2 primitive interfaces [71, 72, 73].

Create an isolated environment and install the pinned dependency:

```
python -m venv .venv
# Activate the environment using your operating system's command.
python -m pip install qiskit==2.5.2
python qiskit_examples.py
```

The web chapter links to the actual script and requirements file. All examples run locally without an account, hardware reservation, API key, or cloud charge.

The complete executable examples and pinned requirements accompany this chapter. Run `python qiskit_examples.py` in the environment above; the script asserts the basis, ordering, Bell-state, estimator, and channel identities before printing results.

23.2 Exact state-vector calculations

```
import numpy as np
from qiskit.quantum_info import Statevector, Pauli

psi = Statevector(np.array([2, 3j]) / np.sqrt(13))
for axis in ("X", "Y", "Z"):
```

```
    expectation = psi.expectation_value(Pauli(axis)).real
    print(axis, expectation)
```

Expected values are 0,12/13,−5/13. These are exact simulator expectations up to floating-point error, not estimates from finite measurements. State-vector simulation should omit measurement instructions when the desired object is a coherent state before readout.

A Bell circuit is equally direct:

```
from qiskit import QuantumCircuit
from qiskit.quantum_info import DensityMatrix, partial_trace

bell = QuantumCircuit(2)
bell.h(0)
bell.cx(0, 1)
state = Statevector.from_instruction(bell)
reduced = partial_trace(DensityMatrix(state), [1])
print(reduced.data)
```

The reduction is I/2. In Qiskit, the argument lists subsystems to **trace out**, not retain. The Bell example is symmetric; use asymmetric examples to test ordering.

23.3 Convert wire conventions deliberately

This book displays $|q_0 q_1 \cdots q_{n-1}\rangle$ with q0 most significant. Qiskit state-vector indices give its qubit 0 the least significant weight. To preserve the book's displayed wire labels, map book wire j to Qiskit wire n−1−j.

For three qubits, applying X to book q0 is:

```
asymmetric = QuantumCircuit(3)
asymmetric.x(2)  # book q0 maps to Qiskit wire 2
assert np.argmax(Statevector.from_instruction(asymmetric).probabilities())
 ↪  == 4
```

Index four is binary 100 in both displayed integer conventions after this mapping. A Bell state alone would not expose a reversed-order mistake.

Classical measurement registers also have display conventions. Name them, record the qubit-to-classical-bit mapping, and inspect an asymmetric known state before interpreting a histogram.

23.4 Finite shots with a V2 sampler

```
from qiskit.primitives import StatevectorSampler

measured = bell.copy()
measured.measure_all()
```

```
sampler = StatevectorSampler(seed=2026)
result = sampler.run([measured], shots=2048).result()
counts = result[0].data.meas.get_counts()
print(counts)
```

Only 00 and 11 should occur in this ideal example, with frequencies near one half. The default classical register created by measure_all is named meas, which is why the result path uses that name. A circuit with a differently named register needs the corresponding data field.

StatevectorSampler implements a local pure-state sampler and does **not** support mid-circuit measurements. A dynamic circuit needs a simulator or backend supporting its actual control flow. Removing an intermediate measurement to make code run changes the experiment and can change its answer.

23.5 Estimators and observables

V2 estimator primitives accept a circuit and observable specification. A local StatevectorEstimator can compute expectations of Pauli-sum observables:

```
from qiskit.primitives import StatevectorEstimator
from qiskit.quantum_info import SparsePauliOp

observable = SparsePauliOp.from_list([("XX", 1.0), ("ZZ", 1.0)])
value = StatevectorEstimator().run([(bell,
↪  observable)]).result()[0].data.evs
print(value)  # 2.0 for the ideal Bell state
```

A hardware estimator has sampling, compilation, and service-specific options. The local exact value is a validation reference, not evidence that a physical processor will return it without uncertainty.

For hardware work, map the logical circuit to the target's instruction set and connectivity, transform observables consistently with any layout, inspect the compiled depth and operations, then execute with explicit shots or precision. The current provider documentation governs credentials, supported modes, and costs. No cloud submission is needed for this textbook's reproducible examples.

23.6 Other tools in the ecosystem

Tool / interface	Useful role	What to verify
Qiskit	Circuit construction, quantum information, compilation, primitives	Version, register order, target instructions, V2 result shapes

Tool / interface	Useful role	What to verify
Cirq	Circuit and device modeling, state-vector and density simulation	Qubit order, supported operations and noise
PennyLane	Differentiable quantum circuits and hybrid workflows	Device, interface, gradient assumptions, shots
Stim	Large stabilizer circuits and detector models	Clifford restrictions and detector semantics
OpenQASM 3	Circuit-level language with classical control constructs	Supported subset and backend execution semantics

These descriptions follow official technical documentation [74, 75, 32, 76]. A language supporting a feature does not imply every backend implements it. Cross-platform conversion can lose timing, control-flow, or noise semantics.

23.7 A reproducibility record

Save the source circuit, dependency versions, backend or simulator identity, compilation options, layout, noise equations, shots, seeds, raw results, and the analysis code. Hardware calibrations change; a backend name alone is insufficient.

Validate small cases before scaling. Check normalization, known basis outputs, unitary inverses, analytical probabilities, and ordering conversions. Compare independent implementations where possible. A shared bug can pass a test that merely repeats the implementation's formula, so use known identities and independent libraries.

A random seed makes a software experiment repeatable within its stated implementation; it does not make a physical device's measurement outcomes deterministic. A change of library version can also change a pseudorandom stream or transpilation result while preserving correct distributions.

23.8 Exercises

23.1. Which Qiskit wire represents book q1 in a four-qubit register under the mapping?

Solution / guidance

Wire $4 - 1 - 1 = 2$.

23.2. Why should the exact Bell-state expectation of XX+ZZ be two?

Solution / guidance

The Bell state is a +1 eigenstate of both XX and ZZ, so their expectations sum to two.

23.3. Can the local StatevectorSampler example execute teleportation with intermediate measurements and feed-forward?

Solution / guidance

No. Use a compatible dynamic-circuit simulator or a branch-by-branch mathematical calculation. The browser teleportation lab computes all conditioned branches explicitly.

23.4. Why pin package versions if the formulas are timeless?

Solution / guidance

APIs, result containers, supported operations, and compilation behavior change. Pinning separates mathematical reproducibility from software-interface drift.

23.5. Name a stronger ordering test than a Bell histogram.

Solution / guidance

Prepare an asymmetric basis state such as book $|100\rangle$, map its wires explicitly, and verify its integer index and measurement-string interpretation.

CHAPTER 24

Reading the experimental frontier in 2026

Learning goals. Classify current claims, connect experiments to the theory already learned, and identify what evidence would strengthen or weaken a conclusion.

24.1 A dated scientific snapshot

Research cutoff: 5 September 2026. This chapter collects selected primary evidence relevant to an introductory course. It is not a league table or a census of every announcement. The purpose is to connect mathematical and engineering concepts to inspectable results while keeping the limits of each claim visible.

The references distinguish a journal article from its earlier preprint and online date from issue year. Where full text was inaccessible, the scope below is restricted to the accessible primary abstract or official record. The source notes identify those cases. No figure or experimental dataset in this chapter is a fabricated measurement.

24.2 Error suppression and logical operations

Experiment: a below-threshold surface-code memory. Google Quantum AI's Willow study was published online on 9 December 2024 and appeared in Nature 638 (2025). It reported distance-dependent logical suppression and a distance-seven memory beyond its physical-memory comparison. The publisher lists an author correction dated 28 April 2026; it fixes repetition-code axis and legend labels in Figure 3a, and the updated article and correction were consulted [47, 48]. Chapter 18 discusses the reported per-cycle metric. Scaling a memory establishes an important component, while universal operations and sustained large computations add further requirements.

Experiment: neutral-atom logical architecture. Bluvstein et al.'s article appeared online on 10 November 2025 and in Nature 649 (2026). Its accessible primary abstract describes repeated surface-code characterization, logical entangling operations, teleportation-based logic, and mid-circuit reuse in reconfigurable arrays [49]. The evidence supports demonstrated architectural

components under the reported circuits. It does not establish unlimited-depth operation or practical advantage for arbitrary applications.

Experiment with correction and detection. Paetznick et al., published 10 June 2026, reported logical error improvements ranging from 11-fold to 800-fold relative to several physical circuit baselines on a trapped-ion QCCD device. The work combines correction, detection, and postselection [50]. The accessible journal abstract was used here. The improvement factors must therefore be interpreted with the different baselines and acceptance conditions; they are not a universal multiplier applicable to every algorithm.

These studies demonstrate why the distinction among encoding, detection, correction, memory, and logical gates matters. A single word such as "logical" cannot carry the whole experimental specification.

24.3 Hardware scale and benchmark scope

Experiment: a 98-qubit trapped-ion processor. The Helios article was published online on 17 June 2026 [59]. Its reported all-to-all architecture is a physical platform result. This book uses the primary publication record and preprint as a platform reference, rather than treating the qubit count as evidence for a particular large algorithm's runtime.

Experiment: modular photonics. Aghaee Rad and collaborators' 2025 work investigates a modular photonic architecture and networking [60]. The important lesson is system composition: sources, optical components, detection, switching, and loss must work together. An architecture experiment is different from a complete useful fault-tolerant application.

Experiment: networked processing. Main et al.'s 2025 optical-link study connects trapped-ion modules and demonstrates distributed operations [66]. Teleportation and feed-forward become concrete engineering tools. A successful laboratory link does not by itself provide a scalable wide-area quantum network.

Experiment and classical comparison: many-body interference. Google's October 2025 Nature paper studies interference in quantum dynamics and associated classical simulation difficulty, often discussed under the "Quantum Echoes" name [63]. Such a result must be read with its observable, circuit family, accuracy, and comparison algorithm. The paper's benchmark claim does not imply that every quantum circuit is difficult to simulate or that every application now has an advantage.

24.4 Conditional resource estimates

Preprint / theory: revised RSA resources. Gidney's 21 May 2025 preprint analyzes a different space–time tradeoff for RSA-2048 factoring [52]. Its hardware assumptions include a square-grid architecture, specified gate error, code-

cycle time, and classical response time. Chapter 20 compares its scope with the peer-reviewed 2021 estimate. These are architecture analyses, not observed cryptographic attacks.

Preprint / theory: reconfigurable-atom resources. Cain et al.'s 30 March 2026 preprint explores high-rate coding and reconfigurable atomic architectures [53]. Its abstract describes cryptographically relevant instances at approximately ten-thousand-qubit scales; a 26,000-qubit scenario is associated with a few-day P-256 discrete-log estimate, whereas RSA-2048 is one to two orders of magnitude slower under its assumptions. The workloads and tradeoffs must remain attached to the numbers.

This distinction is especially important when a title's smallest qubit count is repeated without its runtime or target cryptosystem. An improved theoretical estimate can change planning priorities while still requiring substantial unbuilt engineering.

24.5 Disputed interpretations and company announcements

Microsoft-associated hybrid-device parity measurements have prompted scientific debate about what the diagnostic evidence establishes [61, 62]. This text describes the measurement work and the existence of a technical challenge to its interpretation. It does not adjudicate a disputed topological-computation claim by treating a press release as a proof.

A company roadmap states an intended future capability. It is useful evidence about that company's goals, not evidence that the capability exists or will arrive on schedule. This book deliberately gives no forecast date for a general cryptographically relevant or commercially useful fault-tolerant machine. Such a date is not established by the sources considered.

24.6 What would count as stronger evidence?

For error correction, seek replicated distance scaling, clearly defined logical operations, acceptance and discard statistics, stable long runs, realistic noise, and an integrated decoder. For an algorithm, seek a task matched to the input-output problem, total resource accounting, controlled accuracy, and comparison against strong classical methods. For hardware, seek representative simultaneous-operation results and scaling behavior, rather than only the best isolated component.

Negative results also matter. A classical method reproducing a claimed difficult observable can narrow an advantage claim. A decoder floor can invalidate a favorable large-distance extrapolation. A failed optimization can expose a state-preparation bottleneck. Scientific progress includes refining the conditions under which a claim holds.

A beginner is ready to read frontier papers when they can ask: What state was prepared? What operation was implemented? What was measured? What was conditioned on? Which errors were modeled? Which costs were counted? Which conclusion follows directly, and which is an extrapolation?

24.7 Exercises

24.1. Classify "a paper estimates an algorithm needs 20,000 qubits under specified noise" as experiment, theorem, estimate, or roadmap.

Solution / guidance

It is an estimate or theoretical analysis. It may contain mathematical theorems, but the physical requirement remains conditional on the architecture and model.

24.2. A logical demonstration reports 99.99% conditional accuracy while accepting 5% of attempts. Find accepted-correct probability per attempt.

Solution / guidance

$0.9999 \times 0.05 = 0.049995$, or 4.9995%. This does not make the experiment unimportant, but it changes its throughput interpretation.

24.3. Why retain both online and issue dates?

Solution / guidance

They distinguish when a result became available from the bibliographic year assigned to its issue. A 2026 issue can contain work publicly available in 2025.

24.4. A benchmark's classical runtime was later reduced greatly. Does the original experiment disappear?

Solution / guidance

No. Its measurements remain evidence, but the comparative advantage claim may need revision. Experimental validity and superiority over a moving baseline are separate questions.

24.5. What is missing from "our processor has the most logical qubits"?

Solution / guidance

The definition of logical, code and distance, supported operations, error and acceptance metrics, stability, physical overhead, timing, and the task being compared.

CHAPTER 25

Integrated laboratories and a path to research

Learning goals. Combine mathematical derivation, simulation, statistical inference, and evidence assessment in reproducible projects.

25.1 Project 1: identify a preparation

A source produces one of three preparations: $|+\rangle$, $|+i\rangle$, or an equal classical mixture of $|0\rangle$ and $|1\rangle$. Design a measurement plan to distinguish them statistically.

Predict X,Y,Z distributions before sampling. All three have balanced Z outcomes. In X, $|+\rangle$ is deterministic while the other two are balanced. In Y, $|+i\rangle$ is deterministic while the other two are balanced. Thus X and Y data distinguish the three preparations; Z alone cannot.

Use L02/L03 for the pure states and L11 with mixing set to one for the maximally mixed state. Choose a shot budget and explain how sampling uncertainty affects a classification. For a general unknown source, do not interpret an observed all-plus run as mathematical proof of a pure state.

Deliverable. A preparation table, exact projection calculations, simulated counts with seeds, a decision rule, and a discussion of its finite-data errors. A strong report distinguishes a model-selection exercise among three promised candidates from full tomography of an arbitrary state.

25.2 Project 2: an entanglement claim under noise

Use L07 to prepare a Bell state and inspect its local reductions. Use L08's noisy Bell model to scan visibility v and CHSH settings. Derive the optimal S value $2\sqrt{2}v$ and threshold $v > 1/\sqrt{2}$ before sampling.

Compare the Bell state with a classically correlated mixture of 00 and 11. Both have local state I/2 and perfect Z correlations, but they differ in other correlations. For the classical mixture, $\langle XX\rangle = 0$ and $\langle ZZ\rangle = 1$; for $|\Phi^+\rangle$, both are one. Local entropy alone does not certify entanglement.

Repeat at small and large shot counts near the violation boundary. Quantify uncertainty rather than reporting only whether one sampled S happened to

exceed two. The simulator assumes the specified joint distribution; it cannot test experimental loopholes.

Deliverable. Density matrices, reduced states, a correlation table, S-versus-visibility analysis, and an explanation of which statements are properties of the model and which would require a physical experiment.

25.3 Project 3: recover a hidden structure

Choose either Simon's algorithm or small-instance order finding. Record the oracle's promise separately from the information supplied to the classical reconstruction.

For Simon, generate measured rows until rank n−1 is reached. Show binary elimination explicitly. Repeat with fewer samples and explain why multiple nonzero candidates remain. For order finding, sample several Fourier outcomes and show which convergents pass the modular-order test and produce nontrivial gcd factors.

A failed sample is an expected part of these algorithms. Do not delete it from the report. Estimate the number of complete repetitions required for a target confidence, and include the oracle implementation in any runtime discussion.

Deliverable. An exact small-instance distribution, raw sampled results, reconstruction steps, verified answers, and a cost-model comparison with a relevant classical method. The visible educational oracle table should not be presented as a real hidden-input benchmark.

25.4 Project 4: choose a code under a resource budget

Begin with the repetition-code model and compare d=3,5,7 under independent X errors. Explain why its impressive suppression does not protect arbitrary quantum states. Then use the distance-three surface-code lab to identify a single error, a stabilizer-equivalent recovery, and a zero-syndrome logical string.

Use L25 to design an illustrative memory/computation budget. State A, p, p_{th}, L, ϵ, the chosen odd d, data patches, additional patches, cycle time, and logical depth. Verify the distance choice by direct substitution. Vary p by a factor of two and compare costs.

Deliverable. A code-scope comparison, syndrome calculations, a distance derivation, resource tables, and at least three omitted engineering costs. Do not label the result a prediction for a named machine without a calibrated model.

25.5 Project 5: a reproducible software comparison

Run the supplied Qiskit examples and reproduce an asymmetric three-qubit circuit in L06. Map book wire j to Qiskit wire n−1−j, and compare exact am-

plitudes up to global phase. Compare measurement probabilities rather than expecting identical seeded counts from different random-number generators.

Add a controlled gate and a basis measurement. Introduce an intentional ordering error, predict its effect, and show how an asymmetric test detects it. Then compare a single-qubit channel's density matrix with the explicit Kraus calculation in Chapter 8.

Deliverable. Source circuits, environment versions, mathematical predictions, exact numerical comparisons with tolerances, finite-shot results, and a documented failure case. The objective is to build a debugging method, not simply collect matching screenshots.

25.6 Project 6: audit a current claim

Choose one primary paper from Chapter 24. Read its abstract, methods available to you, stated metric, and limitations. Create a claim table with columns for the statement, evidence type, measured or assumed quantities, comparison baseline, and the strongest conclusion justified.

If access is limited to an abstract, state that limitation and do not infer unreported details. Check publication and correction records. Separate the experiment's result from architecture proposals or roadmaps discussed around it.

Deliverable. A short evidence-backed technical assessment. It should identify one result that is directly established, one extrapolation, and one useful next experiment or analysis. This exercise develops scientific reading, not a verdict about a company's worth.

25.7 Assessing your work

Use four criteria: correctness of the mathematical model; traceability from inputs to computed outputs; appropriate interpretation of uncertainty and assumptions; and clarity of the connection to the scientific question. A polished figure cannot compensate for a wrong basis convention. A correct formula cannot compensate for a claim about a different experiment.

For numerical work, include checks that could fail: probability normalization, channel trace preservation, known-state outputs, inverse recovery, and independent comparisons. State the tolerance and why it is sensible. Tests are strongest when they check an independent property rather than reimplementing the same expression.

For open-ended work, a negative or qualified conclusion can be excellent. A model that does not support a proposed advantage has still taught something if the analysis is transparent.

25.8 Where to go next

For information theory, study completely positive maps, entropy inequalities, channel capacities, and entanglement theory [4]. For algorithms, study query lower bounds, Hamiltonian simulation, block encodings, and quantum signal processing [23]. For error correction, study stabilizer formalism, decoding, fault-tolerant gadgets, and resource estimates [33, 37].

For hardware, choose a physical platform and learn its Hamiltonian, control methods, noise characterization, and fabrication or optical constraints. For software, contribute a small reproducible benchmark, simulator feature, or verified circuit transformation. The ability to connect assumptions, equations, code, and evidence is a useful research skill across all of these paths.

APPENDIX A

Mathematical toolkit

This appendix supplies details that are useful across chapters. It can be read when a calculation first needs them rather than memorized before beginning.

A.1 Matrix multiplication and adjoints

For matrices A of shape m×n and B of shape n×p, $(AB)_{jk} = \sum_{\ell=1}^{n} A_{j\ell} B_{\ell k}$. The inner dimensions must match. For two square matrices,

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} e & f \\ g & h \end{pmatrix} = \begin{pmatrix} ae + bg & af + bh \\ ce + dg & cf + dh \end{pmatrix}.$$

The adjoint conjugates entries and transposes: $(A^\dagger)_{jk} = A^*_{kj}$. It reverses products: $(AB)^\dagger = B^\dagger A^\dagger$. Consequently a circuit's inverse reverses gate order and replaces each gate with its adjoint.

An outer product $|u\rangle\langle v|$ is a matrix with entries $u_j v_k^*$. Its action on $|w\rangle$ is $|u\rangle(\langle v|w\rangle)$. An inner product is a scalar. Confusing the order changes the mathematical type of the result.

A.2 Eigenvectors and diagonalization

An eigenvector satisfies $A|v\rangle = \lambda|v\rangle$ with nonzero v. Eigenvalues solve $\det(A - \lambda I) = 0$. For a Hermitian two-by-two matrix

$$A = \begin{pmatrix} a & c \\ c^* & b \end{pmatrix}, \quad a, b \in \mathbb{R},$$

the characteristic equation is

$$(a - \lambda)(b - \lambda) - |c|^2 = 0,$$
$$\lambda^2 - (a + b)\lambda + ab - |c|^2 = 0.$$

Thus

$$\lambda_{\pm} = \frac{a + b \pm \sqrt{(a-b)^2 + 4|c|^2}}{2}.$$

These are real. Solve $(A - \lambda I)v = 0$ for each eigenvector and normalize. If eigenvalues are distinct, Hermiticity makes the eigenvectors orthogonal: taking matrix elements in both orders gives $(\lambda - \mu)\langle v|w\rangle = 0$.

The spectral theorem writes a Hermitian matrix as $A = \sum_j \lambda_j |v_j\rangle\langle v_j|$ with an orthonormal eigenbasis. It defines functions of A: $f(A) = \sum_j f(\lambda_j)|v_j\rangle\langle v_j|$. This explains exponentials of Hamiltonians, square roots of positive matrices, and matrix entropy.

A matrix is positive semidefinite exactly when all its eigenvalues are nonnegative. A two-by-two Hermitian density matrix with trace one is physical if its diagonal entries and determinant are nonnegative. Checking only trace and Hermiticity is insufficient.

A.3 Gram–Schmidt and projection

Given a normalized vector $|u\rangle$ and another vector $|v\rangle$, subtract its component along u:

$$|w\rangle = |v\rangle - |u\rangle\langle u|v\rangle.$$

Then $\langle u|w\rangle = \langle u|v\rangle - \langle u|u\rangle\langle u|v\rangle = 0$. If w is nonzero, normalize it. Repeating this process builds an orthonormal basis from linearly independent vectors.

For an orthonormal set, the projector onto its span is $P = \sum_j |u_j\rangle\langle u_j|$. It obeys $P^2 = P = P^\dagger$. For a complete basis, P=I. Expanding a vector using this identity yields its projection coefficients.

The Cauchy–Schwarz inequality $|\langle u|v\rangle|^2 \le \langle u|u\rangle\langle v|v\rangle$ ensures that overlap probabilities between normalized states lie in [0,1]. Equality occurs when the vectors are linearly dependent.

A.4 Singular values and Schmidt form

For any matrix M, the eigenvalues of $M^\dagger M$ are nonnegative. Their square roots are its singular values. A singular-value decomposition is $M = U\Sigma V^\dagger$, with diagonal nonnegative Σ and unitary U,V of suitable shapes.

For a bipartite pure state's coefficient matrix $M_{ab} = c_{ab}$, the singular values are square roots of the Schmidt probabilities. In the fixed product bases, the reduced operators are $\rho_A = MM^\dagger$ and $\rho_B = M^T M^* = (M^\dagger M)^T$. Indeed, $(\rho_B)_{bb'} = \sum_a M_{ab} M^*_{ab'}$. Both have the same nonzero eigenvalues. Their equality of spectra explains equal reduced entropies for a pure joint state.

The largest singular value is the operator norm ‖M‖, the maximum stretching of a unit vector. The trace norm sums singular values. The Frobenius norm is $\|M\|_F = \sqrt{\sum_{jk} |M_{jk}|^2}$. Different norms answer different error questions; state overlap, global operator error, and local observable error should not be interchanged.

A.5 Geometric sums and Fourier cancellation

For r≠1,

$$1 + r + \cdots + r^{N-1} = \frac{1 - r^N}{1 - r}.$$

Multiply the left side by 1−r and all intermediate terms cancel. For a nontrivial Nth root of unity, $r^N = 1$, so the sum is zero. For r=1, the sum is N; using the quotient without taking the limit would divide by zero.

This one identity proves QFT column orthogonality and determines finite-precision phase-estimation peaks. Binary Hadamard cancellation is a related character-orthogonality identity for XOR rather than addition modulo N.

A.6 Continued fractions by recurrence

For partial quotients $a_0, a_1, \ldots$, initialize $p_{-2} = 0, p_{-1} = 1, q_{-2} = 1, q_{-1} = 0$. Then

$$p_k = a_k p_{k-1} + p_{k-2}, \qquad q_k = a_k q_{k-1} + q_{k-2}.$$

The kth convergent is p_k/q_k. For $[0; 2, 1, 170]$: $0/1$, $1/2$, $1/3$, and $171/512$ follow directly. In order finding, a good rational approximation still needs a modular check; the recurrence does not know the physical problem's order.

A.7 Confidence and finite samples

For independent Bernoulli samples, the standard deviation of a sample proportion is $\sqrt{p(1-p)/N}$. Substituting $\hat{p}$ gives an estimated standard error, but near zero or one a naive normal interval can be misleading.

A distribution-free bound for independent bounded trials is Hoeffding's inequality:

$$\Pr(|\hat{p} - p| \geq \epsilon) \leq 2e^{-2N\epsilon^2}.$$

To make this at most δ, it suffices that $N \geq \ln(2/\delta)/(2\epsilon^2)$. This is conservative but does not require an unknown p in the bound. It still assumes independence; repeated drift or correlated experimental errors need another analysis.

For several simultaneous claims, their combined error probability must be controlled. A union bound can allocate δ across them. Reporting many favorable outcomes from a large unreported search is not equivalent to testing one prespecified prediction.

A.8 Global phase in numerical comparisons

Two vectors may represent the same pure state while having different entries. Fidelity $|\langle\psi|\phi\rangle|^2$ compares them without global-phase sensitivity. To compare entries, estimate the phase from a nonzero overlap and align one vector first.

Do not use phase-insensitive comparison when relative phases between controlled branches matter. A subroutine equal to a target up to phase may produce a different controlled operation if that phase is not tracked. The scope of the comparison must match the circuit’s use.

APPENDIX B

Glossary and laboratory directory

B.1 Terms and where to study them

Term	Meaning	Chapter
Amplitude	Complex coordinate whose coherent sums determine probabilities	2
Ancilla	Additional subsystem used as workspace, a probe, or a resource	1, 17
Basis	Orthonormal coordinate set or measurement alternatives	3
Bloch vector	Three Pauli expectations of a qubit	4, 6
BQP	Bounded-error polynomial-time quantum decision complexity	1
Channel	Completely positive trace-preserving state transformation	8
Clifford	Gate mapping Pauli strings to Pauli strings by conjugation	16, 19
Code distance	Minimum weight of a nontrivial logical error in a stabilizer code	17
Concurrence	Here, $2\|ad - bc\|$ for a pure two-qubit state	5
Density matrix	Operator describing pure or mixed-state statistics	6
Detector	Parity of measurement records expected to be fixed absent faults	18

Term	Meaning	Chapter
Eigenphase	Phase of a unitary eigenvalue, defined modulo one cycle	12
Entanglement	Joint-state nonseparability; pure-state factorization failure	5, 6
Erasure	Error with a known affected location	17, 19
Fidelity	State similarity; this book uses the squared convention	6
Gate depth	Sequential layers under an explicitly defined schedule	1, 20
Global phase	Common phase multiplying a whole state vector	2
Logical qubit	Encoded information-bearing subsystem or code-space factor	17
Magic state	Nonstabilizer ancillary resource for logical non-Clifford operations	19
Measurement instrument	Operations specifying both outcomes and conditioned states	6
Mitigation	Inference strategy estimating ideal quantities from noisy data	8
Oracle	Specified black-box operation counted as a resource	9
Partial trace	Reduction giving all statistics accessible to a subsystem	6
Pauli frame	Classical record used to interpret inferred Pauli corrections	19
Postselection	Restricting results to accepted observed events	1, 19
POVM	Positive measurement effects summing to identity	6
PQC	Classical cryptography designed to resist known quantum attacks	22
QEC	Quantum error correction	17–19

Term	Meaning	Chapter
QKD	Quantum key distribution with classical authenticated processing	22
QFT	Unitary Fourier transform of quantum amplitudes	12
QPE	Quantum phase estimation	12
QSVT	Polynomial transformation of singular values through encoded access	14
Relative phase	Phase difference between coherent components	2–4
Schmidt rank	Number of nonzero Schmidt coefficients of a pure bipartite state	6
Shot	One execution yielding measurement data	1, 23
Stabilizer	Pauli constraint fixing a code space	17
Syndrome	Check outcomes used to infer errors	17, 18
T-count	Number of non-Clifford T resources in a specified circuit model	19, 20
Threshold	Noise boundary for a specified scalable correction protocol	18
Trace distance	Operational distinguishability measure based on trace norm	6
Uncomputation	Reversing coherent work to remove intermediate information	1, 9
Unitary	Inner-product-preserving linear transformation	4

B.2 Laboratory directory

All identifiers below are present in the web edition and linked from the corresponding PDF chapter.

Lab	Experiment	Chapter
L01	Reversible logic and classical repetition	1
L02	Complex state preparation	2

Lab	Experiment	Chapter
L03	Full Z/X/Y projection workbench	3
L04	Randomized measurement practice	3
L05	Bloch sphere and single-qubit gates	4
L06	Circuit builder and state evolution	5
L07	Entanglement, density matrices, and partial trace	6
L08	Bell/CHSH correlations and finite shots	7
L09	Teleportation branches and corrections	7
L10	Superdense coding	7
L11	Exact single-qubit noise channels	8
L12	Error-mitigation bias and variance	8
L13	Deutsch, Deutsch–Jozsa, Bernstein–Vazirani	9
L14	Simon constraints and row reduction	10
L15	Grover and amplitude amplification	11
L16	Quantum Fourier transform	12
L17	Quantum phase estimation	12
L18	Small-instance Shor order finding	13
L19	Hamiltonian evolution and product formulas	14
L20	Variational energy and gradient steps	15
L21	One-layer QAOA on a triangle	15
L22	Classical representation memory cost	16
L23	Repetition-code syndromes and logical failures	17
L24	Distance-three surface-code Pauli decoding	18
L25	Illustrative logical and physical resources	20

Lab	Experiment	Chapter
L26	BB84 transmission and intercept–resend	22

B.3 Limits that travel with the models

The generic circuit laboratory supports eight qubits and forty operations. Measurement branches are resampled for each shot. The phase and Fourier laboratories use exact small-register transforms. The Shor laboratory computes a modular-exponentiation Fourier marginal for N=15 or 21 and performs classical reconstruction; it does not compile a cryptographic modular arithmetic circuit.

The noise laboratory evolves one-qubit density operators with explicitly stated Kraus models. The surface-code laboratory uses nine data qubits, perfect stabilizer information, and minimum-weight Pauli recovery. It does not simulate faulty ancilla circuits or many-round spacetime decoding. The resource calculator uses an illustrative scaling relation, and the mitigation lab uses a specified exponential model.

BB84’s classroom transcript exposes information for learning. It is not production cryptographic software and does not implement a composable finite-key proof or a secure authenticated network. These boundaries define what each tool actually computes; they are essential to interpreting its output.

References

Sources were checked against the research boundary of 5 September 2026. Publication status and access limitations remain attached to each entry.

[1] Thomas G. Wong (2022). Introduction to Classical and Quantum Computing. Rooted Grove; ISBN 979-8-9855931-0-5. **Status:** Curriculum reference; supplied PDF.

[2] Claude E. Shannon (1948). A Mathematical Theory of Communication. Bell System Technical Journal 27. **Status:** Foundational paper.

[3] Charles H. Bennett (1973). Logical Reversibility of Computation. IBM Journal of Research and Development 17. **Status:** Foundational paper.

[4] John Watrous (2018). The Theory of Quantum Information. Cambridge University Press. **Status:** Graduate reference; author-hosted materials.

[5] William K. Wootters and Wojciech H. Zurek (1982). A single quantum cannot be cloned. Nature 299. **Status:** Foundational paper.

[6] John F. Clauser, Michael A. Horne, Abner Shimony, and Richard A. Holt (1969). Proposed Experiment to Test Local Hidden-Variable Theories. Physical Review Letters 23. **Status:** Foundational paper.

[7] Bas Hensen et al. (2015). Loophole-free Bell inequality violation using electron spins separated by 1.3 kilometres. Nature 526. **Status:** Peer-reviewed experiment.

[8] Charles H. Bennett et al. (1993). Teleporting an unknown quantum state via dual classical and Einstein-Podolsky-Rosen channels. Physical Review Letters 70. **Status:** Foundational protocol.

[9] Charles H. Bennett and Stephen J. Wiesner (1992). Communication via one- and two-particle operators on Einstein-Podolsky-Rosen states. Physical Review Letters 69. **Status:** Foundational protocol.

[10] Zhenyu Cai et al. (2023). Quantum error mitigation. Reviews of Modern Physics 95, 045005. **Status:** Peer-reviewed review.

[11] David Deutsch (1985). Quantum theory, the Church-Turing principle and the universal quantum computer. Proceedings of the Royal Society A 400. **Status:** Foundational algorithm.

[12] David Deutsch and Richard Jozsa (1992). Rapid solution of problems by quantum computation. Proceedings of the Royal Society A 439. **Status:** Foundational algorithm.

[13] Ethan Bernstein and Umesh Vazirani (1997). Quantum Complexity Theory. SIAM Journal on Computing 26. **Status:** Peer-reviewed theory.

[14] Daniel R. Simon (1997). On the Power of Quantum Computation. SIAM Journal on Computing 26. **Status:** Peer-reviewed theory.

[15] Lov K. Grover (1996). A fast quantum mechanical algorithm for database search. Proceedings of STOC 1996. **Status:** Foundational algorithm.

[16] Gilles Brassard, Peter Hoyer, Michele Mosca, and Alain Tapp (2002). Quantum Amplitude Amplification and Estimation. Contemporary Mathematics 305. **Status:** Peer-reviewed theory.

[17] Christof Zalka (1999). Grover's quantum searching algorithm is optimal. Physical Review A 60. **Status:** Query lower bound.

[18] Alexei Yu. Kitaev (1995). Quantum measurements and the Abelian Stabilizer Problem. arXiv:quant-ph/9511026. **Status:** Foundational preprint.

[19] Richard Cleve, Artur Ekert, Chiara Macchiavello, and Michele Mosca (1998). Quantum algorithms revisited. Proceedings of the Royal Society A 454. **Status:** Peer-reviewed theory.

[20] Peter W. Shor (1997). Polynomial-Time Algorithms for Prime Factorization and Discrete Logarithms on a Quantum Computer. SIAM Journal on Computing 26. **Status:** Peer-reviewed algorithm; preprint version linked.

[21] Seth Lloyd (1996). Universal Quantum Simulators. Science 273. **Status:** Foundational simulation algorithm.

[22] Andrew M. Childs, Yuan Su, Minh C. Tran, Nathan Wiebe, and Shuchen Zhu (2021). Theory of Trotter Error with Commutator Scaling. Physical Review X 11, 011020. **Status:** Peer-reviewed theory.

[23] Andras Gilyen, Yuan Su, Guang Hao Low, and Nathan Wiebe (2019). Quantum singular value transformation and beyond: exponential improvements for quantum matrix arithmetics. Proceedings of STOC 2019. **Status:** Peer-reviewed theory; author preprint linked.

[24] Guang Hao Low and Isaac L. Chuang (2019). Hamiltonian Simulation by Qubitization. Quantum 3, 163. **Status:** Peer-reviewed theory.

[25] Aram W. Harrow, Avinatan Hassidim, and Seth Lloyd (2009). Quantum Algorithm for Linear Systems of Equations. Physical Review Letters 103, 150502. **Status:** Peer-reviewed theory.

[26] Hsin-Yuan Huang et al. (2021). Power of data in quantum machine learning. Nature Communications 12, 2631. **Status:** Peer-reviewed theory and numerical study.

[27] Alberto Peruzzo et al. (2014). A variational eigenvalue solver on a photonic quantum processor. Nature Communications 5, 4213. **Status:** Foundational variational experiment.

[28] Edward Farhi, Jeffrey Goldstone, and Sam Gutmann (2014). A Quantum Approximate Optimization Algorithm. arXiv:1411.4028. **Status:** Foundational preprint.

[29] Jarrod R. McClean et al. (2018). Barren plateaus in quantum neural network training landscapes. Nature Communications 9, 4812. **Status:** Peer-reviewed theory.

[30] Scott Aaronson and Daniel Gottesman (2004). Improved simulation of stabilizer circuits. Physical Review A 70, 052328. **Status:** Peer-reviewed simulation method.

[31] Guifre Vidal (2003). Efficient Classical Simulation of Slightly Entangled Quantum Computations. Physical Review Letters 91, 147902. **Status:** Peer-reviewed simulation method.

[32] Craig Gidney (2021). Stim: a fast stabilizer circuit simulator. Quantum 5, 497. **Status:** Peer-reviewed software paper.

[33] Daniel Gottesman (1997). Stabilizer Codes and Quantum Error Correction. PhD thesis, California Institute of Technology. **Status:** Foundational thesis.

[34] Peter W. Shor (1995). Scheme for reducing decoherence in quantum computer memory. Physical Review A 52, R2493. **Status:** Foundational code.

[35] Andrew M. Steane (1996). Error Correcting Codes in Quantum Theory. Physical Review Letters 77, 793. **Status:** Foundational code.

[36] Emanuel Knill and Raymond Laflamme (1997). Theory of quantum error-correcting codes. Physical Review A 55, 900. **Status:** Recovery conditions.

[37] Austin G. Fowler, Matteo Mariantoni, John M. Martinis, and Andrew N. Cleland (2012). Surface codes: Towards practical large-scale quantum computation. Physical Review A 86, 032324. **Status:** Peer-reviewed architectural review.

[38] Eric Dennis, Alexei Kitaev, Andrew Landahl, and John Preskill (2002). Topological quantum memory. Journal of Mathematical Physics 43, 4452. **Status:** Foundational theory.

[39] Oscar Higgott and Craig Gidney (2025). Sparse Blossom: correcting a million errors per core second with minimum-weight matching. Quantum 9, 1600. **Status:** Decoder method; primary author manuscript.

[40] Dorit Aharonov and Michael Ben-Or (2008). Fault-Tolerant Quantum Computation with Constant Error Rate. SIAM Journal on Computing 38. **Status:** Threshold theory; author manuscript.

[41] Bryan Eastin and Emanuel Knill (2009). Restrictions on Transversal Encoded Quantum Gate Sets. Physical Review Letters 102, 110502. **Status:** Peer-reviewed no-go theorem.

[42] Sergey Bravyi and Alexei Kitaev (2005). Universal quantum computation with ideal Clifford gates and noisy ancillas. Physical Review A 71, 022316. **Status:** Magic-state theory.

[43] Pavel Panteleev and Gleb Kalachev (2022). Asymptotically Good Quantum and Locally Testable Classical LDPC Codes. Proceedings of STOC 2022. **Status:** Peer-reviewed coding theory.

[44] Sergey Bravyi et al. (2024). High-threshold and low-overhead fault-tolerant quantum memory. Nature 627, 778–782. **Status:** Theory and circuit-noise simulation.

[45] Daniel Gottesman, Alexei Kitaev, and John Preskill (2001). Encoding a qubit in an oscillator. Physical Review A 64, 012310. **Status:** Bosonic code theory.

[46] Volodymyr V. Sivak et al. (2023). Real-time quantum error correction beyond break-even. Nature 616, 50–55. **Status:** Peer-reviewed bosonic experiment.

[47] Google Quantum AI and Collaborators (2025). Quantum error correction below the surface code threshold. Nature 638, 920–926; online 9 December 2024. **Status:** Peer-reviewed memory experiment; updated publisher page consulted.

[48] Google Quantum AI and Collaborators (2026). Author Correction: Quantum error correction below the surface code threshold. Nature 653, E5; published 28 April 2026. **Status:** Publisher correction record.

[49] Dolev Bluvstein et al. (2026). A fault-tolerant neutral-atom architecture for universal quantum computation. Nature 649, 39–46; online 10 November 2025. **Status:** Peer-reviewed experiment; primary institutional record and abstract.

[50] A. Paetznick et al. (2026). Improved quantum processor logical error rates via correction and detection. Nature 654, 349–355; published 10 June 2026. **Status:** Peer-reviewed experiment; accessible abstract, correction and postselection.

[51] Craig Gidney and Martin Ekera (2021). How to factor 2048 bit RSA integers in 8 hours using 20 million noisy qubits. Quantum 5, 433. **Status:** Peer-reviewed resource estimate.

[52] Craig Gidney (2025). How to factor 2048 bit RSA integers with less than a million noisy qubits. arXiv:2505.15917v1; submitted 21 May 2025. **Status:** Preprint resource estimate; not a hardware factorization.

[53] Madelyn Cain et al. (2026). Shor's algorithm is possible with as few as 10,000 reconfigurable atomic qubits. arXiv:2603.28627v1; submitted 30 March 2026. **Status:** Preprint architecture and resource estimate.

[54] Jens Koch et al. (2007). Charge-insensitive qubit design derived from the Cooper pair box. Physical Review A 76, 042319. **Status:** Foundational device theory.

[55] Morten Kjaergaard et al. (2020). Superconducting Qubits: Current State of Play. Annual Review of Condensed Matter Physics 11, 369–395. **Status:** Platform principles review; not a 2026 performance snapshot.

[56] Colin D. Bruzewicz, John Chiaverini, Robert McConnell, and Jeremy M. Sage (2019). Trapped-ion quantum computing: Progress and challenges. Applied Physics Reviews 6, 021314. **Status:** Platform principles review; author manuscript.

[57] Mark Saffman (2016). Quantum computing with atomic qubits and Rydberg interactions: progress and challenges. Journal of Physics B 49, 202001. **Status:** Platform principles review.

[58] Guido Burkard, Thaddeus D. Ladd, Andrew Pan, John M. Nichol, and Jason R. Petta (2023). Semiconductor spin qubits. Reviews of Modern Physics 95, 025003. **Status:** Peer-reviewed platform review.

[59] Anthony Ransford et al. (2026). A 98-qubit trapped-ion quantum computer with all-to-all connectivity. Nature; published online 17 June 2026. **Status:** Peer-reviewed experiment; primary record and arXiv:2511.05465.

[60] H. Aghaee Rad et al. (2025). Scaling and networking a modular photonic quantum computer. Nature 638, 912–919. **Status:** Peer-reviewed architecture experiment; accessible primary record.

[61] Microsoft Azure Quantum (2025). Interferometric single-shot parity measurement in InAs–Al hybrid devices. Nature 638, 651–655. **Status:** Peer-reviewed measurement study; interpretation distinguished from a topological-computer claim.

[62] Henry F. Legg (2025). Comment on Interferometric single-shot parity measurement in InAs–Al hybrid devices. arXiv:2503.08944. **Status:** Technical challenge in a preprint; not treated as final adjudication.

[63] Google Quantum AI and Collaborators (2025). Observation of constructive interference at the edge of quantum ergodicity. Nature 646, 825–830; published 22 October 2025. **Status:** Peer-reviewed experiment and classical-simulation comparison.

[64] Hans-Jurgen Briegel, Wolfgang Dur, J. Ignacio Cirac, and Peter Zoller (1998). Quantum Repeaters: The Role of Imperfect Local Operations in Quantum Communication. Physical Review Letters 81, 5932. **Status:** Foundational network theory.

[65] Stephanie Wehner, David Elkouss, and Ronald Hanson (2018). Quantum internet: A vision for the road ahead. Science 362, eaam9288. **Status:** Peer-reviewed network perspective.

[66] D. Main et al. (2025). Distributed quantum computing across an optical network link. Nature 638, 383–388; published 5 February 2025. **Status:** Peer-reviewed experiment; primary author manuscript.

[67] Charles H. Bennett and Gilles Brassard (1984 / 2014). Quantum cryptography: Public key distribution and coin tossing. 1984 conference paper; reprinted in Theoretical Computer Science 560. **Status:** Foundational protocol.

[68] Valerio Scarani et al. (2009). The security of practical quantum key distribution. Reviews of Modern Physics 81, 1301. **Status:** Peer-reviewed security review.

[69] Peter W. Shor and John Preskill (2000). Simple Proof of Security of the BB84 Quantum Key Distribution Protocol. Physical Review Letters 85, 441. **Status:** Security proof under explicit idealized assumptions.

[70] National Institute of Standards and Technology (2024). Announcing Approval of Three Federal Information Processing Standards for Post-Quantum Cryptography. FIPS 203, 204, 205; announcement 13 August 2024. **Status:** Official standards announcement; checked 5 September 2026.

[71] IBM Quantum (2026 access). Quickstart. Official documentation; examples specify qiskit~=2.5.2. **Status:** Official technical documentation; checked 5 September 2026.

[72] IBM Quantum (2026 access). StatevectorSampler. Qiskit API documentation, V2 local sampler. **Status:** Official API; no mid-circuit measurement support.

[73] IBM Quantum (2026 access). Primitives. Qiskit API documentation. **Status:** Official V2 interface documentation.

[74] Google Quantum AI (2026 access). Cirq. Official framework and simulator documentation. **Status:** Official technical documentation.

[75] PennyLane developers (2026 access). Quantum circuits. PennyLane stable documentation. **Status:** Official technical documentation.

[76] OpenQASM contributors (2026 access). OpenQASM 3 language specification. Live specification and version 3.0 documentation. **Status:** Official language specification.