Preface

This book is for a one-semester course in quantum computing and quantum information theory. It emphasizes the mathematical aspects and the historical continuity of both algorithms and information theory.

Quantum computing and quantum information theory have often been presented as radical departures from classical computation and information theory because of the "strangeness" of quantum mechanics. If one takes a broader view, the picture looks different.

Shor's factoring algorithm is one in a long line of surprising efficient algorithms. This book presents algorithms leading up to Shor's algorithm, e.g., the Schönhage-Strassen polynomial multiplication algorithm and the Miller-Rabin probabilistic algorithm to test compositeness. It then gives a rigorous and self-contained treatment of Shor's algorithm. It also gives self-contained treatments of the algorithms of Grover and Simon.

Shannon made an extraordinarily deep and versatile foundation of information theory. The foundation is so solid that quantum information theory, despite its counterintuitive aspects, is a straightforward adaptation of Shannon's theory. Both Shannon's theory and its adaptation to the quantum setting are presented in detail.

There are deep and beautiful connections between quantum information theory and representation theory. I present this material, which is generally not available in book form. It includes an exposition of the quantum marginal problem, representation-theoretic proofs of standard entropic inequalities, and even a representation-theoretic proof of strong subadditivity.

0.1. "Real world" landscape as of 2024

The primary method that banks, governments, etc., currently use to communicate securely is the RSA cryptosystem.[1] RSA relies on the assumption that it is difficult to factor a large number N into its prime factors. In 1994 P. Shor [**Sho94**] (also see [**Sho97**]) described an algorithm to factor numbers quickly on a "quantum computer", something that at the time did not really exist. This and other developments spurred interest in building one.

Shortly after Shor's algorithm, Grover [**Gro96**] developed a remarkable search algorithm that is of complexity the square root of the size of a brute-force search (essentially the best-known classical algorithm for unstructured input). Despite a large stream of excellent research in the past 30 years, see, e.g., [**STF**$^+$], [**HHL09**], and [**AJSP14**], there have not been any "game-changing" discoveries comparable to Shor's or Grover's algorithms.

In January 2016, and in more detail in October, the NSA released a document warning the world that current encryption algorithms will no longer be secure as early as 2025[2]. They were concerned that by 2025 there might be quantum computers capable of effectively running Shor's algorithm. Such a development would endanger secure communication.

In response to Shor's algorithm, mathematicians, computer scientists, and cryptographers are developing *post-quantum cryptography*, a growing body of new cryptographic paradigms that are expected to be immune to attacks using a quantum computer.

Writing now, in 2024, there are working quantum computers. You may even test one yourself by going to `https://www.ibm.com/quantum/technology`.

On the other hand, it seems likely in the medium-term future that operating quantum computers will not be able to break RSA cryptography (but see [**KSB**$^+$**20**]). The stumbling block is that all existing large quantum computers are prone to errors, and Shor's algorithm is not fault tolerant.

Is there any near-term potential advantage of quantum computers over classical computers? The answer is yes, because the modeling of complex systems *is* fault tolerant and there is the hope that quantum computers may soon be better than classical computers in tasks such as designing and testing airplane wings. See, e.g., [**KEA**$^+$**23**]. The terminology in the literature for the current state is *noisy intermediate-scale quantum computing*, NISQ. Even if the advantage does not materialize, the challenges quantum computing presents

[1] In addition to RSA, the elliptic curve Diffie-Hellman method is also used, which relies on the assumption that computing the discrete logarithm is difficult. A variant of Shor's algorithm gives an efficient evaluation of the discrete logarithm. See [**Hen21**] for a discussion.

[2] See `https://people.tamu.edu/~jml//CNSA-Suite-and-Quantum-Computing-FAQ.pdf`. For the current warning, see `https://media.defense.gov/2022/Sep/07/2003071836/-1/-1/0/CSI_CNSA_2.0_FAQ_.PDF`.

to classical computing have spurred considerable research. One such example is that of *tensor networks*, which provide a method of classically simulating quantum processes. See, e.g., [**NLD+23**] and the references therein.

Quantum information theory has brought new perspectives to several areas of research. For example, there have been recent advances in understanding the classical complexity of multiplying matrices by introducing ideas from quantum information theory [**CVZ19**, **CVZ18**].

0.2. Usage

For a one-semester graduate class, one can cover most of the book. For a one-semester undergraduate class, one can either cover Chapters 1–3 or cover up to and including §6.1 (the quantum noiseless channel theorem) while skipping the details of the algorithms of Miller-Rabin and Shor. It is also possible to give a short graduate course on either Chapters 1–3 on algorithms or Chapters 4–7 on information theory, using some material from Chapter 2.

Prerequisites. A background in undergraduate mathematics including some knowledge of elementary probability (probability distributions on finite sets), linear algebra (eigenvalues and eigenvectors of linear maps), and a semester of undergraduate algebra (basic definitions and properties of groups and rings) is needed.

Layout. All theorems, propositions, remarks, examples, etc., are numbered together within each section. For example, Definition 1.3.2 is the second numbered item in §1.3. Equations are numbered sequentially within each section. Hints and answers for exercises marked with the symbol ⊙ are given at the end of the book.

0.3. Acknowledgments

I am especially grateful to Alessandra Bernardi, who invited me to give a course on the subject at the University of Trento in the summer of 2017. I thank the Trento group for their wonderful hospitality, especially A. Bernardi, I. Carusotto, and L. Sola Conde. I thank M. Michałek for numerous suggestions after he taught a class at U. Konstanz from a draft of this book. I thank students C. Chang, R. Geng, P. Speegle, T. Tran, L. Wang, and D. Wu for extensive comments, corrections, and suggestions. I thank M. Christandl, I. Nechita, R. O'Donnell, and M. Walter for helping me learn the subject, and I thank S. Aaronson for advice on designing the course. I thank K. Dykema, M. Fanizza, I. Leigh, and especially G. Berkolaiko, P. Kuchment, J. M. Rojas, and my editor, S. Gelfand, for excellent comments on a near final version of this book.

0.4. Notation

What follows is notation that will be used throughout the book.

$i = \sqrt{-1}$.

$\omega_N = e^{\frac{2\pi i}{N}}$.

The natural logarithm is denoted ln and log denotes $\log_2$.

Binomial coefficients are $\binom{n}{p} \coloneqq \frac{n!}{p!(n-p)!}$, and multinomial coefficients are $\binom{n}{p_1,\ldots,p_d} \coloneqq \frac{n!}{p_1!\cdots p_d!}$.

For a complex number $z = x + iy$ or matrix X, $\bar{z} = x - iy$ and $\bar{X}$ denote their complex conjugates.

For functions f, g of a real variable (or integer) x:

$f(x) = O(g(x))$ if there exists a constant $C > 0$ and x_0 such that $|f(x)| \leq C|g(x)|$ for all $x \geq x_0$.

$f(x) = o(g(x))$ if $\lim_{x\to\infty} \frac{|f(x)|}{|g(x)|} = 0$.

$\mathrm{poly}(n)$ denotes a function whose growth is bounded above and below by polynomials in n.

$f(x) = \Omega(g(x))$ if there exists a constant $C > 0$ and x_0 such that $C|f(x)| \geq |g(x)|$ for all $x \geq x_0$.

Finite-dimensional complex vector spaces will generally be denoted U, V, W.

Finite-dimensional Hilbert spaces will generally be denoted $\mathcal{H}$, $\mathcal{H}_A$, $\mathcal{H}_B$, $\mathcal{H}_C$.

Subspaces of a Hilbert space will generally be denoted $\mathcal{M} \subset \mathcal{H}$.

For vector spaces V and W, V^* denotes the space of linear maps $V \to \mathbb{C}$, $\mathrm{Hom}(V, W) \cong W \otimes V^*$ denotes the space of linear maps from V to W, and $\mathrm{End}(V)$ denotes the space of linear maps from V to itself.

Elements of a Hilbert space $\mathcal{H}$ will generally be denoted $|x\rangle$ and elements of $\mathcal{H}^*$ will be denoted $\langle x|$. For $|v\rangle, |w\rangle \in \mathcal{H}$, $\langle v|w\rangle$ denotes their Hermitian inner product, which we take to be linear in the second factor and conjugate linear in the first. We use the notations $|xy\rangle \coloneqq |x\rangle \otimes |y\rangle$ and $\langle xy| \coloneqq \langle x| \otimes \langle y|$. The pairing $\mathcal{H} \otimes \mathcal{H}^* \to \mathbb{C}$ is written $|x\rangle \otimes \langle y| \mapsto \langle y|x\rangle$.

For a subspace $U \subset V$,

$$U^\perp \coloneqq \{\alpha \in V^* \mid \alpha(u) = 0 \ \forall u \in U\}$$

denotes its annihilator in the dual space.

For a subspace $\mathcal{M} \subset \mathcal{H}$,

$$\mathcal{M}^\perp \coloneqq \{|v\rangle \in \mathcal{H} \mid \langle w|v\rangle = 0 \ \forall |w\rangle \in \mathcal{M}\}$$

denotes its orthogonal complement.

$\mathcal{X} = \{a_1, \ldots, a_n\}$ denotes a finite set. We sometimes write $[n] = \{1, ..., n\}$.

When $\mathcal{X} = \{a_1, ..., a_n\}$ we let $p_{\mathcal{X}}$ denote a probability distribution on $\mathcal{X}$.

$\bar{p}_{\mathcal{X}} = (p_{\mathcal{X}}(a_1), \ldots, p_{\mathcal{X}}(a_n))$ denotes the vector representing the distribution $p_{\mathcal{X}}$.

$\Pr(M)$ is the probability that the event M occurs with respect to some understood probability distribution.

For a matrix $X \in \mathrm{Mat}_{n\times m}$, $X^{\mathbf{t}} \in \mathrm{Mat}_{m\times n}$ denotes its transpose and for a linear map $f : V \to W$, its transpose is denoted $f^{\mathbf{t}} : W^* \to V^*$, which is the induced linear map on dual spaces.

For Hilbert spaces $\mathcal{H}_A$ and $\mathcal{H}_B$ and an element $X \in \mathrm{Hom}(\mathcal{H}_A, \mathcal{H}_B)$, $X^\dagger \in \mathrm{Hom}(\mathcal{H}_B, \mathcal{H}_A)$ denotes the adjoint linear map defined by $\forall v \in \mathcal{H}_A$, $w \in \mathcal{H}_B$, $\langle Xv|w\rangle_{\mathcal{H}_B} = \langle v|X^\dagger w\rangle_{\mathcal{H}_A}$. In particular, for $X \in \mathrm{Mat}_{n\times n}(\mathbb{C})$, $X^\dagger := \bar{X}^{\mathbf{t}}$.

$\mathrm{Herm}(\mathcal{H}) \subset \mathrm{End}(\mathcal{H})$ is the subset of Hermitian endomorphisms; namely $X \in \mathrm{End}(\mathcal{H})$ such that $X^\dagger = X$.

$\mathcal{D}(\mathcal{H}) \subset \mathrm{Herm}(\mathcal{H})$ is the subset of density operators; see Definition 5.1.11.

$\mathbf{U}(n) := \{X \in \mathrm{Mat}_{n\times n}(\mathbb{C}) \mid X^\dagger = X^{-1}\}$ is the unitary group; see Definition 2.1.4.

$\mathbf{O}(n) := \{X \in \mathrm{Mat}_{n\times n}(\mathbb{R}) \mid X^{\mathbf{t}} = X^{-1}\}$ is the orthogonal group; see equation (2.1.3).

$\mathbb{C}^* = \mathbb{C}\backslash\{0\}$ is the nonzero complex numbers.

$\mathbb{F}_2$ is the field with two elements 0 and 1.

$\mathbb{Z}/M\mathbb{Z}$ denotes the ring of integers modulo M; see §A.1.

$a \bmod M$ denotes the equivalence class of a in $\mathbb{Z}/M\mathbb{Z}$ but we slightly abuse notation and use the same notation for the nonnegative integer that is the remainder when a is divided by M, i.e., the smallest nonnegative representative of the equivalence class.

For $Z \in \mathrm{End}(\mathcal{H})$ and $X \in \mathrm{End}(\mathcal{H})$, $Z \cdot X := ZXZ^\dagger$.