# Lower bounds on the complexity of quantum proofs

*Chinmay Nirkhe*

Electrical Engineering and Computer Sciences
University of California, Berkeley

Technical Report No. UCB/EECS-2022-236
http://www2.eecs.berkeley.edu/Pubs/TechRpts/2022/EECS-2022-236.html

November 23, 2022

# Lower bounds on the complexity of quantum proofs

by

Chinmay Nirkhe

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Computer Science

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Umesh Vazirani, Chair
Assistant Professor Avishay Tal
Associate Professor Nikhil Srivastava

Fall 2022

Lower bounds on the complexity of quantum proofs

# Abstract

## Lower bounds on the complexity of quantum proofs

by

Chinmay Nirkhe

Doctor of Philosophy in Computer Science

University of California, Berkeley

Professor Umesh Vazirani, Chair

The quantum PCP conjecture is one of the central open questions in quantum complexity theory. It asserts that calculating even a rough approximation to the ground energy of a local Hamiltonian is intractable even for quantum devices. The widely believed separation between the complexity classes NP and QMA necessitates that polynomial length classical proofs do not exist for calculating the ground energy. This further implies that low-energy states of local Hamiltonians cannot be described by constant-depth quantum circuits. The *No low-energy trivial states (NLTS)* conjecture by Freedman and Hastings posited the existence of such Hamiltonians.

This thesis describes a line of research culminating in a proof of the NLTS conjecture, first presented by Anshu, Breuckmann, and Nirkhe. The construction is based on quantum error correction and the thesis elaborates on how error correction, local Hamiltonians, and low-depth quantum circuits are related.

To the memory of my grandparents,
whose passion for science will inspire for generations.

# Contents

## Publications and preprints used in this thesis

The contents of this thesis are largely based on the following previous publications and preprints. The authorships are listed alphabetically by surname and reflect equal contributions by all. Some portions of this thesis were taken verbatim from these prior works and consent to do so was provided by all co-authors.

- Anurag Anshu, Nikolas Breuckmann, and Chinmay Nirkhe. *NLTS Hamiltonians from good quantum codes* [1].
- Anurag Anshu and Chinmay Nirkhe. *Circuit lower bounds for low-energy states of code Hamiltonians* [2].
- Chinmay Nirkhe, Umesh Vazirani and Henry Yuen. *Approximate low-weight check codes and circuit lower bounds for noisy ground states* [3].
- Sany Irani, Anand Natarajan, Chinmay Nirkhe, Sujit Rao, and Henry Yuen. *Quantum search-to-decision reductions and the state synthesis problem* [4].
- Anand Natarajan and Chinmay Nirkhe. *A classical oracle separation between* QMA *and* QCMA [5].

Publications [6–8] were also completed during my doctoral work but are not incorporated in this thesis.

# Acknowledgments

> "The Guide says there is an art to flying", said Ford, "or rather a knack. The knack lies in learning how to throw yourself at the ground and miss."
>
> Douglas Adams, *Life, the Universe and Everything*

I am very fortunate to have had a wonderful time in graduate school and there are many people who I need to thank for this. First and foremost, is my advisor, Umesh Vazirani. I can thank Umesh for many things but the two most important are – in order – chocolate and wisdom. I have learned much about how to be an academic, how to ask questions, and how to do research by observing Umesh. I think only now do I understand Ford's knack for flying. I just needed someone to teach me how to throw myself at the ground and miss.

For the entirety of my degree, I had an office in Room 615 Soda Hall and I've been very lucky to have Zeph Landau right next door in Room 617. For many years, I would start my day by going next door and "brain-dumping" all my new ideas onto Zeph who so generously listened to every word of mine. I cannot thank him enough for requiring explanations of every detail of every idea in its utmost simplicity. I have learned so much from Zeph and he is a bountiful source of friendship, mathematical insight, and unique perspectives.

Furthermore, I am indebted to the postdocs who have been at Berkeley who, likewise, openly offered me their time, ideas, and friendship: Rotem Arnon-Friedman, Anurag Anshu, Adam Bouland, Andrea Coladangelo, and Henry Yuen. I especially need to thank Anurag for, first and foremost, being a friend, but secondly being incredibly supportive of my ideas during the COVID-19 pandemic and writing some beautiful results with me which form the core of this thesis. In addition, Anand Natarajan has always been excited to collaborate with me on research and I'm glad that he is my friend and a partner chef. I have also been fortunate to collaborate with many other wonderful scientists: Srinivasan Arunachalam, Thom Bohdanowicz, Sergey Bravyi, Nikolas Breuckmann, Elizabeth Crosson, Bill Fefferman, Sandy Irani, Bryan O'Gorman, and Sujit Rao; research is not conducted in a vacuum.

I'd like to thank Thomas Vidick at Caltech who first instilled in me a fascination for quantum complexity theory and who is always available day or night for a conversation. Additionally, I'd like to thank Jalex and Andrea for their amazing company and support when I was just getting acquainted with academia. Furthermore, I'd like to thank Henry Yuen, Dorit Aharanov, Matt Coudron, and Anand Natarajan who hosted me on wonderful research visits. Thank you to Srinivasan Arunachalam and Kristan Temme for hosting me during the summer and fall of 2021 at IBM; I'm excited to join IBM Research with you after I finish my time at Berkeley.

Thank you Anurag Anshu, Nikolas Breuckmann, and Umesh Vaziani for many helpful discussions and comments regarding the presentation of this thesis.

I'd like to thank the members of the theory group who made Soda Hall so magical. Arun, Tarun, Rachel, Urmila, Seri, Orr, James, Manuel, Malvika, Bhaskar, Bryan, Sam (Hopkins and Gunn), Siqi, Siddhanth, Yunchao, Jonah, and Pasin, thank you. In addition, professors Prasad Raghavendra, Satish Rao, Alessandro Chiesa, Avishay Tal, John Wright, Nikhil Srivastava, and Venkat Guruswami at Berkeley have been wonderful sources of conversation and insight. Thank you, Elizabeth, for being an infinite source of conversation and friendship, especially during the COVID-19 pandemic. Thank you, Nick, for not only being the best of office mates but also the best of roommates.

Additional thanks to my friends outside the theory group who defined the last half decade. A special thanks to Yash, Elaine, Suma, and Andreea for being fantastic roommates and a true pleasure to live with. Thank you to Zach, Elaine, Brian, Chelsea, Dmac, Mara, Aritra, Erin, Marco, Coby, Eugene, Rishabh, Andrew, Linda, and Jon for reminding of life outside of research. Thank you to the friends I made running around in the hills of Berkeley: Saavan, Mike, Ryan, Doug, Josh, Kelly, Justin, Esther, Arya, Nathan, Mauricio, and Monica. Further back, my friends from childhood in Seattle, especially Langston, Alec, Andrew, and, of course, my two fantasy football teams have been wonderful company[1]. Thank you to Bobby G's Pizzeria for providing me with the liquid sustenance I needed to complete this thesis and for being a place I could always go to relax and unwind. Sheila, thank you for being such a wonderful presence and companion; I loved every moment of our time together.

A very special thanks to my family, especially my parents and Surabhi for providing me with infinite love, encouragement, and support. Despite being close and far, you were always there for me and this thesis is as much mine as it is yours.

[1] *The League of Chumpions* has been an excellent source of distraction on Sundays and for that, I have to thank my FF-nemesis Elaine, Coby, Red, Dmac, Rishabh, Brian, Bobby, Darius, and my FF-want-to-be nemesis Zach.

# Chapter 1

# Introduction

> Dr. Hoenikker used to say that any scientist who couldn't explain to an eight-year-old what he was doing was a charlatan.
>
> Kurt Vonnegut, *Cat's Cradle*[0]

## 1.1 The mathematics of proofs

The subject of mathematics begins and ends with proofs. A proof is an argument convincing a verifier beyond a reasonable doubt of the validity of a statement. Starting from Euclid of Alexandria's *Elements* to the present day, unwavering proofs, arguments written line by line starting from a core set of axioms, have been the standard of rigor for all of mathematics. In the roughly two millennia since the *Elements*, our understanding of the complexity of proofs has broadened immensely; in particular, we have begun to understand the idea of proofs through the lens of computation.

The notion of the Turing machine gave birth to the mathematical study of computation, but it wasn't until Robert Floyd, who coined the term *non-determinism* in 1967 [9], that the *computational complexity* of proofs was studied. The breakthrough results of Cook and Levin [10, 11] defined the computational class $\mathsf{NP}$, the set of languages for which a proof can be verified efficiently, and provided a complete problem of SAT, the problem of deciding the satisfiability of a boolean formula. This led to the question of the complexity of finding a proof versus deciding if a proof was correct[1] — i.e $\mathsf{P} \overset{?}{=} \mathsf{NP}$. The past half-century has even further broadened our understanding of the computational complexity of proofs. A convenient perspective is that an efficient proof is *any* string of bits that a verifier[2] can check in time $\mathrm{poly}(n)$. Proofs could be *randomized* in that the verifier can

[0] I guarantee you that this thesis will not live up to the herculean standards of Mr. Vonnegut.

[1] The nomenclature is far more variable than simply "proof and statement". Equivalent terms for proofs included solution, answer, witness or certificate while statements were sometimes referred to as questions or problems. The theory of computation gave rise to the term *language* to describe a set of questions that should be answered with yes.

[2] We think of a verifier as a computational device such as a boolean circuit or Turing machine.

afford to mistakenly reject valid proofs or mistakenly accept invalid proofs, provided this happens with a sufficiently low probability (say $< 1/3$). This defines the non-deterministic computational class MA (Merlin-Arthur) which is not believed to increase the computational power [12]. Far more variations on proofs exist and they form a beautiful theory which is the core of computational complexity theory.

### 1.1.1 Quantum proofs

The particular variant of proofs focused on in this thesis is a quantum generalization of NP and MA. In this model, a quantum proof is an entangled state consisting of $\mathrm{poly}(n)$ qubits. To check the proof, the verifier must be a *quantum device* and to enforce efficiency, it must run in polynomial time — i.e. a BQP device. The computational class of all such proof systems is called Quantum Merlin-Arthur (QMA). The quantum analog to the Cook-Levin theorem was proven by Alexei Kitaev in 1999 [13–15]; he showed that calculating the ground energy (i.e. minimum eigenvalue) of a local Hamiltonian instance was QMA-complete. Furthermore, he showed that the ground state of a local Hamiltonian was a checkable proof that the ground energy of the local Hamiltonian was small (below a fixed threshold). This proved that local Hamiltonian ground states are a general notion of a proof for all of QMA.

Since it is widely believed that $\mathsf{NP} \neq \mathsf{QMA}$, then quantum proofs *cannot* be replaced by classical proofs. In particular, ground states of local Hamiltonians cannot be classically described in any efficiently checkable manner. This makes understanding the problems for which we can and the problems for which we *provably* cannot classically describe ground (and low-energy) states such a fascinating question. This is the main question studied in this thesis.

### 1.1.2 Probabilistically checkable proofs

Parallel to the progress being made in expanding our understanding of computation from classical to quantum, a sweeping insight into the nature of classical proofs was being made[3]. Starting from the Cook-Levin theorem that constraint satisfaction problems (CSPs) are NP-complete [10, 11], the study of non-deterministic computation crescendoed with the *probabilistically checkable proofs* (PCP) theorem [17–19], arguably the *crown-jewel* of theoretical computer science.

The PCP theorem, which originated from a line of research on the complexity of interactive proofs, is commonly interpreted as a proof of hardness-of-approximation for CSPs. The hardness-of-approximation (also known as the gap-amplification) version of the PCP theorem [19] states that is NP-complete to distinguish whether a CSP of $m$ clauses is satisfiable or if no more than $m/2$ clauses can be simultaneously satisfied. In contrast, the Cook-Levin proof only implies the hardness of distinguishing whether a CSP is satisfiable or unsatisfiable (violating one clause). In other words, the PCP theorem proves that it is NP-complete to even estimate the satisfiability of a CSPs to a precision of $m/4$ (versus 1 due to Cook-Levin). This shows that, for some optimization problems, any efficient algorithm will yield solutions that are off by a constant multiplicative error (such as

[3] We recommend [16] for a history of the PCP theorem.

10%) from ideal. The PCP theorem also has a proof-checking consequence: every NP problem can be reduced to a CSP problem such that one only needs to read $O(\log(1/\epsilon))$ bits of the witness to be confident with probability $1 - \epsilon$ that the witness is legitimate. This second interpretation is the origin of the term, *probabilistically checkable proofs*. The PCP theorem completely revised our notion of proofs from a systematic sequence of local steps, each requiring meticulous verification to a robust object where a few randomly selected global checks could simultaneously verify all the local constraints.

## 1.2 Quantum probabilistically checkable proofs

With the understanding that calculating the ground energy of local Hamiltonians is QMA-complete, a reasonable conjecture was set forth of whether a quantum analog of the PCP theorem holds [20,21] for approximating the ground energy of local Hamiltonians. While variations on the statement of the quantum PCP conjecture exist, the most commonly expressed one is the hardness-of-approximation (also known as, gap-amplification) version:

**Conjecture 1.1 (Quantum PCP [20,21])** *It is* QMA*-complete to decide whether a local Hamiltonian on n qubits and* $m = \Theta(n)$ *terms has ground energy (minimum eigenvalue)* $\leq m/10$ *(*yes *instance) or* $\geq m/5$ *(*no *instance) even when promised that one of the cases holds.*

The constants of $1/10$ and $1/5$ can be replaced with any pair of small separated constants[4]; but for the rest of the introduction, we will assume this setup. A family of local Hamiltonians for which it is QMA-complete to decide the ground energy to such an approximation is called a quantum PCP local Hamiltonian (family). Despite many efforts by the community, it is unclear if we are any closer today to resolving the quantum PCP conjecture than we were when it was posited over two decades ago. Due to the wide-spreading influence of the PCP theorem on almost all branches of theoretical computer science, the quantum PCP conjecture remains the biggest open question in quantum complexity theory.

Since quantum PCPs subsume classical PCPs [17–19] which are highly *engineered* mathematical objects building on the theory of locally testable codes and expanders, quantum PCPs will also be highly engineered objects. In particular, the quantum PCP conjecture necessitates a form of robust and "exotic" entanglement[5] in the ground and low-energy space of quantum PCP local

[4] This is due to parallel repetition and adding trivial terms. All this changes is the locality of the Hamiltonian and the number of terms by constant factors.

[5] The crucial role of entanglement in the theory of quantum many-body systems is widely known with some seminal examples including topological phases of matter [22] and quantum computation with physically realistic systems [23, 24]. But entanglement also brings new challenges as the classical simulation of realistic many-body systems faces serious computational overheads. Estimating the ground energy of such systems is one of the major problems in condensed matter physics [25], quantum chemistry [26], and quantum annealing [27,28]. One of the key methods to address this problem is to construct *ansatz quantum states* that achieve as low energy as possible and are also suitable for numerical simulations. A leading ansatz, used in Variational Quantum Eigensolvers [26, 29, 30] or Quantum Adiabatic Optimization Algorithm [31], is precisely the class of quantum states that can be generated by low-depth quantum circuits.

Hamiltonians. While ground states of generic local Hamiltonian problems serve as proofs of a small ground energy, in the case of quantum PCP local Hamiltonians, the space of witnesses is more diverse and, in particular, includes all low-energy states. Since it is widely believed that quantum proofs cannot be classically described, then all quantum proofs, including all low-energy states, must be highly entangled. This is a major reason why the quantum PCP conjecture seems much harder than the classical PCP theorem. We are not sure if Hamiltonians can be constructed with the necessary exotic entanglement in the low-energy space, let alone if they can capture quantum non-deterministic computation.

To elaborate, consider a proof $|\xi\rangle$ attempting to convince a verifier that the ground energy of a local Hamiltonian is $\leq m/10$ when the verifier is *promised* that the ground energy is either $\leq m/10$ or $\geq m/5$. If the proof $|\xi\rangle$ corresponds to a state of energy $< m/5$, then it is convincing to any verifier as they can measure the energy of $|\xi\rangle$, know then that the ground energy is $< m/5$, and therefore $\leq m/10$ due to the promise. Therefore, the set of proofs for the local Hamiltonian problem includes the low-energy space as well as the ground space; in this context, the low-energy space is the subspace spanned by all states of energy $< m/5$.

In particular, the constant-temperature *thermal (Gibbs) states* of the local Hamiltonian must be exotic. The thermal state at temperature $\beta$ is the mixed state given by $e^{-\beta H}$ which at temperature $\beta = 0$ is the uniform distribution over the ground space and for small constant $\beta$, the state is low-energy. It is widely suspected that due to a decay of correlations, as $\beta$ increases, the entanglement of the thermal states of the Hamiltonian will decrease. If the entanglement was to decrease sufficiently fast so that the thermal state for small $\beta$ was always approximated by a product (i.e. classical) state, then this would put the complexity of estimating the energy of the quantum PCP local Hamiltonian in $\mathsf{NP}$. This is because the classical description of the product state would suffice as a classical proof that the thermal state was low-energy and therefore decide the local Hamiltonian problem using only a classical proof and classical computation. (This is explained in greater detail in Section 3.3). Since the problem is assumed to be $\mathsf{QMA}$-complete, if the quantum PCP conjecture is true, it follows that all thermal states of quantum PCP Hamiltonians are far from product states. This flies in face of the intuition of most condensed matter physicists who believe that for constant $\beta$, the thermal state simplifies as the universe is warmer and presumably less entangled.

We can generalize this requirement in two ways: *any* low-energy state which is the output of *a constant depth* quantum circuit can serve as a classical proof for the quantum PCP Hamiltonian problem. Such states are called *trivial* or low-depth and they naturally generalize the notion of product states. Furthermore, the circuits describing trivial states of low-energy also serve as *classical* proofs for quantum PCP Hamiltonians because of an efficient *classical* algorithm for computing the energy of any trivial state with respect to a local Hamiltonian (see Section 3.3). Therefore, assuming $\mathsf{NP} \neq \mathsf{QMA}$, the quantum PCP conjecture implies that all low-energy states of quantum PCP Hamiltonians are far from all trivial states. This is considerably stronger than the weaker requirement that the thermal states are non-trivial as the set of low-energy states may contain some *adversarial* examples[6].

[6] We use the term adversarial here to emphasize the difference between these low-depth states and thermal states. The construction of thermal states is a natural phenomenon; attaching the universe to a heat bath and slowly adding

## 1.3 No low-energy trivial states

It is reasonable to be suspicious whether there are any constructions of local Hamiltonians whose low-energy subspaces exhibit such robust entanglement[7]. In 2014, Michael Freedman and Matthew Hastings gave this suspicion a name: the *No low-energy trivial states (NLTS)* conjecture [32].

**Theorem 1.2 (NLTS (Simplified))** *There exists a family of Hamiltonians such that every low-energy state cannot be generated by a constant depth circuit.*

*See Theorem 3.10 or Theorem 4.1 for formal statements.*

The statement was conjectured in [32] and first proven in [1]. Any family of Hamiltonian which satisfies Theorem 1.2 is called an NLTS Hamiltonian (family). Proving that a Hamiltonian is NLTS is equivalent to proving a *circuit depth* lower bound for all low-energy states.

This remainder of this thesis is, in some sense, the extended story of how such circuit depth lower bounds for a family of Hamiltonians were rigorously proven (Theorem 4.1). We elaborate on the intuitions behind the lower bounds, how to mathematically turn these intuitions into fact, and the perspective on the quantum PCP conjecture in a post-NLTS world. The first family of local Hamiltonians discovered to have the NLTS property was discovered by Anurag Anshu, Nikolas Breuckmann, and Chinmay Nirkhe [1]. The proof showed that a family of local Hamiltonians corresponding to *quantum error correcting codes* of good rate and good distance satisfied the NLTS property.

The proof that there exist local Hamiltonians satisfying the NLTS property is a positive sign for the quantum PCP conjecture. If the NLTS statement had been answered in the negative, it would have proven that every local Hamiltonian's ground energy could be approximated to accuracy constant $\epsilon > 0$ by an $\mathsf{NP}$ algorithm. This would effectively[8] disprove the quantum PCP conjecture.

Before the proof of the NLTS theorem, there was a sequence of results reaffirming the physicists' intuition that there must be trivial low-energy states [33–35]. In particular, Brandão and Harrow [34] provided a construction of product states that were low-energy states of Hamiltonians whose interaction graphs were too expanding; the proof is a consequence of monogamy of entanglement and studying correlation strength with quantum De Finetti theorems. When combined with a folklore proof that the interaction graph of NLTS Hamiltonians must be somewhat expanding, the result of [34] was ominous because it insisted that NLTS Hamiltonians, if they existed at all, sat in a *goldilocks regime* in terms of interaction graph expansion.

energy results in a thermal state. However, the adversarial state may be far from the thermal states. An example of an adversarial state is the rotation of the ground state by a small angle.

[7] For intuition, one can think of the minimum depth of a circuit generating a state is a proxy for the entanglement of the state. This is not exactly true as the minimum depth of a circuit generating a classical probability distribution can also be large. What circuit depth does capture is non-trivial correlations between many qubits.

[8] We say effectively because it would only prove that the quantum PCP conjecture is false in a world where $\mathsf{NP} \neq \mathsf{QMA}$. However, if $\mathsf{NP} = \mathsf{QMA}$, then a quantum PCP conjecture is trivial since the classical PCP conjecture would suffice.

On the optimistic side of the NLTS coin was persistent progress on the construction of quantum error correcting codes. In Chapters 2 and 3, we explain, in-depth, the connection between error correction and circuit depth lower bounds which is the foundation for the proofs in this thesis. A systemic progress in the construction of quantum *low-density parity check codes* [36–43] is largely to thank for the NLTS theorem. These highly engineered mathematical objects are the true engine; if anything, our previous works [1–3] are simply proofs that the properties of quantum error correction are *sufficient* for NLTS Hamiltonians. Through [1–3] along with [44, 45], we developed a deeper understanding of how to make robust folklore circuit depth lower bounds for error correction. This "robustification" of lower bounds, along with surprising insights into the influence of the rate, paved the path to the NLTS theorem.

## 1.4 Outline

In Chapters 2 and 3 we develop the intuition and prerequisites for proving the NLTS property. In particular, Section 3.3 formalizes the definition of the quantum PCP conjecture and its relationship to the NLTS theorem and standard quantum complexity classes. Chapter 4 is the proof the NLTS theorem [1]. Chapter 5 describes the implications of the NLTS theorem on the quantum PCP conjecture. Appendix A describes a sequence of intermediate results on lower bounds for the *description complexity* of quantum states that were proved during my graduate work but do not directly influence the construction of NLTS Hamiltonians. Specifically, Appendix A.1 describes non-trivial circuit depth lower bounds on the low-energy space of all codes [2] (regardless of whether they have good rate, good distance, or even are CSS or stabilizer).

## 1.5 Notation

We assume the reader is familiar with basic Dirac notation and elementary quantum computation notation; for a review, we suggest [46]. For a review on basic complexity theory, we suggest [47]. The set of integers $\{1, 2, \ldots n\}$ is abbreviated as $[n]$. Given a composite system of $n$ qubits, we will often omit the register symbol from the states (being clear from context). For a set $A \subseteq [n]$, $-A$ will denote the set complement $[n] \setminus A$ and $\mathrm{tr}_A$ will denote the partial trace operation on qubits in $A$ and $\mathrm{tr}_{-A} \overset{\text{def}}{=} \mathrm{tr}_{[m]\setminus A}$. Therefore, $\mathrm{tr}_{-\{i\}}(\cdot)$ gives the reduced marginal on the $i$th qubit. The uniformly distributed quantum state on a Hilbert space $\mathcal{H}$ will be represented by $\nu_{\mathcal{H}} \overset{\text{def}}{=} \mathbb{I}_{\mathcal{H}}/|\mathcal{H}|$.

**Quantum states** A quantum state is a positive semi-definite matrix with unit trace, acting on a finite-dimensional complex vector space (a Hilbert space) $\mathcal{H}$. In this thesis, we will only concern ourselves with Hilbert spaces coming from a collection of qubits, i.e. $\mathcal{H} = (\mathbb{C}^2)^{\otimes n}$. A pure quantum state is a quantum state with rank 1 (i.e. it can be expressed as $|\psi\rangle\langle\psi|$ for some unit vector $|\psi\rangle$). In which case we will refer to the state as $|\psi\rangle$ when interested in the unit vector representation and $\psi$ when interested in the positive semi-definite matrix representation. Given two Hilbert spaces $\mathcal{H}_A, \mathcal{H}_B$, their tensor product is denoted by $\mathcal{H}_A \otimes \mathcal{H}_B$. For a quantum state $\rho_{AB}$ acting on $\mathcal{H}_A \otimes \mathcal{H}_B$, the reduced state on $\mathcal{H}_A$ is denoted by $\rho_A \overset{\text{def}}{=} \text{tr}_B(\rho_{AB})$, where $\text{tr}_B$ is the partial trace operation on the Hilbert space $\mathcal{H}_B$. The partial trace operation is a type of quantum channel. More generally, a quantum channel $\mathcal{E}$ maps quantum states acting on some Hilbert space $\mathcal{H}_A$ to another Hilbert space $\mathcal{H}_B$.

Every quantum state $\rho$ acting on a $D$-dimensional Hilbert space has a collection of eigenvalues $\{\lambda_i\}_{i=1}^{D}$, where $\sum_i \lambda_i = 1$ and $\lambda_i \geq 0$. The von Neumann entropy of $\rho$, denoted $\mathrm{S}(\rho)$, is defined as $\sum_i \lambda_i \log \frac{1}{\lambda_i}$. The notation log signifies a base 2 logarithm while ln signifies the natural logarithm.

# Chapter 2

# Quantum error correction

> When you stir your rice pudding, Septimus, the spoonful of jam spreads itself round making red trails like the picture of a meteor in my astronomical atlas. But if you stir backwards, the jam will not come together again. Indeed, the pudding does not notice and continues to turn pink just as before. Do you think this is odd?
>
> Tom Stoppard, *Arcadia*

## 2.1 The features of error correction

Just as one cannot stir backwards the red trails of jam out of the rice pudding, one cannot undo the thermalization or decoherence of a generic quantum state due to the environment. This is the fundamental issue with quantum information; its inherent fragility and therefore its tendency to decohere, thermalize, corrupt and eventually lose all information of its starting state.

When classical information succumbs to noise, a natural cure is to use repetition. Attempt to talk across a noisy room, and you will find yourself shouting the same message over and over again until your point is clear. The trouble is that quantum information cannot be cloned and therefore no simple repetition-like procedure exists for thwarting the fragility of quantum information.

Enter the notion of quantum error correction, first introduced by Shor [48]. Shor's seminal idea was to consider a 1-qubit subspace amongst 9-qubits. The preliminary goal was simple — a la classical error correction, where the only error is bit-flips, the most rudimentary of quantum error correcting codes must be able to correct bit-flips and phase-flips. A phase-flip is the bit-flip but in the Hadamard basis. Shor's 9-qubit code [48] was capable of correcting only a single bit-flip or a single phase-flip on any qubit. But by linearity, since the Pauli matrices $X$, $Z$ and $XZ = iY$ span the space of 1-qubit operators, the 9-qubit code can correct any 1-qubit error.

The notion of protected encoding can be generalized from 1 to $k$ qubits within a generalized $n$ qubit space [49]. If the space is protected against *any* $d$ qubit Pauli error, then we call the code a $[[n, k, d]]$ quantum error correcting code. There are many ideas baked into this simple definition that are worth recognizing. But for the NLTS theorem; we are going to only highlight three.

**Depolarizing noise channel** First, Shor's intention for correcting single-qubit errors was due to an assumption that a reasonable noise model for quantum information is the *depolarizing noise channel*: For an $n$ qubit state $\rho$, on each qubit independently apply the following process $\mathcal{M}$: with probability $\epsilon$, apply a noisy channel $\mathcal{N}$, and with probability $1-\epsilon$ apply the identity channel $\mathcal{I}$. The resulting state is

$$\begin{aligned}\mathcal{M}(\rho) &= ((1-\epsilon)\mathcal{I} + \epsilon\mathcal{N})^{\otimes n}(\rho) && (2.1a)\\ &= \sum_{S\subseteq[n]} (1-\epsilon)^{n-|S|}\epsilon^{|S|}\mathcal{N}^S(\rho) && (2.1b)\\ &\approx \sum_{S:|S|\leq 2\epsilon n} (1-\epsilon)^{n-|S|}\epsilon^{|S|}\mathcal{N}^S(\rho). && (2.1c)\end{aligned}$$

When $\epsilon = O(d/n)$, the number of qubits corrupted by the depolarizing channel is, with high probability, $\leq d$. In Shor's case, $n = 9$ and $d = 1$, so under a constant error threshold, the 9-qubit code was correcting errors. The depolarizing channel is very similar to thermalization, so any construction of NLTS Hamiltonians must, at a minimum, argue that ground states after passing through a $\epsilon$-depolarizing channel remain non-trivial. But an NLTS Hamiltonian construction must have no low-energy trivial states including any adversarial constructions. A priori, it isn't obvious that error correction provides this level of robustness. This is because an adversarial construction of a low-energy state could involve a *global* manipulation of a ground state. The intention of error correction was only meant to handle errors that were localized to $< d$ qubits; what we find, and a fundamental insight into why error correcting codes generate NLTS Hamiltonians is that error correction, for high-rate or high-distance codes, allows for *some* control over global manipulations of ground states.

**Erasure errors** While often stated that a $[[n,k,d]]$ error correcting code can correct any error supported on $\leq d$ qubits, a useful characterization of error correction is that if $\leq d$ qubits are traced out — i.e. removed from the system — then the information can still be recovered from the other $n-d$ qubits. This is because the *completely depolarizing channel* (also known as the Pauli one-time pad)'s action on a state can be expressed as

$$\mathcal{E}_S(\rho) = \rho_{-S} \otimes \nu_S \tag{2.2}$$

where for subset $S \subset [n]$, the channel $\mathcal{E}_S$ acts on the qubits of $S$ and is defined as

$$\mathcal{E}_S(\cdot) = \frac{1}{4^{|S|}} \sum_{a,b\in\{0,1\}^S} \left(X^a Z^b\right)(\cdot)\left(Z^b X^a\right). \tag{2.3}$$

For any subset $S$ such that $|S| \leq d$, we can interpret the action of $\mathcal{E}_S$ as a linear combination of correctable Pauli errors. Therefore, it is correctable by linearity. However, the action of $\mathcal{E}_S$ is equivalent to tracing out the qubits in $S$ and replacing them with the maximally mixed state. This is equivalent to an erasure error.

The ability for quantum error correction to correct erasure errors along with the no-cloning theorem gives us a fundamental property of the quantum error correcting codes that differentiates it from classical error correction. To see this consider a code-word $\rho$ and an erasure error applied on qubits denoted by $S$ for $|S| \leq d$. Since the erasure error is correctable, there exists a recovery channel $\mathcal{R}_S$ such that $\mathcal{R}_S(\rho_{-S}) = \rho$. But what happens to the system $\rho_S$ that was traced out? Since all the information of $\rho$ can be recovered from $\rho_{-S}$, it follows from the no-cloning theorem that no information about $\rho$ was contained in $\rho_S$. To see this, notice that $S(\rho_{-S}) = S(\rho)$ since we can convert between the two states using the channels $\mathcal{R}_S$ and $\mathrm{tr}_S$. Therefore, the mutual information between registers $S$ and $-S$ is $S(\rho_{-S})$, or equivalently these states are uncorrelated. Thus, $\rho_{-S}$ is uncorrelated from $\rho$. As this holds for every code-state $\rho$, the following fact is easy to prove.

**Fact 2.1 (Local Indistinguishability)** *Let $C$ be a $[[n, k, d]]$ error correcting code and $S$ a subset of the qubits such that $|S| < d$. Then the reduced density matrix $\rho_S$ of any code-state $\rho$ on the set $S$ is an invariant of the code. Equivalently, for all code-states $\rho, \varsigma$, it holds that $\rho_S = \varsigma_S$.*

We also give a derivation of this fact from the Knill-Laflamme conditions [49] in the following section.

Let us quickly note that this is not a property of classical error correction; in fact, classical error correction is built on repetition. Key examples are the repetition code and the Hadamard code. In Chapter 5, we discuss the difficulties this induces in using quantum codes in a proof of the quantum PCP conjecture.

**A Hamiltonian defining the codespace** The 9-qubit code by Shor belongs to the class of Calderbank-Shor-Steane (CSS) codes [50, 51] which is a subclass of all *stabilizer* codes. A convenient property of all stabilizer codes is that the codespace can be easily specified as the unique +1 eigenspace of a set of commuting Pauli matrices (Definition 2.2), which in turn also provides a way of testing membership in the codespace[1]. The operators are called *stabilizers*; a more precise definition is given in the following section. For every set of stabilizers, $C_1, \ldots, C_m$, there exists a convenient local Hamiltonian $\mathbf{H}$ whose ground space is the +1 eigenspace of each stabilizer:

$$\mathbf{H} = \sum_{i=1}^{m} h_i \overset{\mathsf{def}}{=} \sum_{i=1}^{m} \frac{\mathbb{I} - C_i}{2}. \tag{2.4}$$

Since each stabilizer is a Pauli matrix, its eigenvalues are ±1 and therefore each $h_i$ is a projector onto the +1 eigenspace of $C_i$. Since the projectors commute, the eigenvalues of $\mathbf{H}$ are easy to calculate as $0, 1, 2, \ldots, m$. The ground space of the Hamiltonian is equivalent to the code-space. The family of NLTS Hamiltonians we consider in Theorem 4.1 will be the Hamiltonians $\mathbf{H}$ corresponding to a family of CSS codes.

[1] For example, the 1-qubit subspace among 7-qubits defined by the Steane code is stabilized by $IIIZZZZ$,$IZZIIZZ$, $ZIZIZIZ$, $IIIXXXX$,$IXXIIXX$,$XIXIXIX$ where we use a short-form of ignoring $\otimes$ symbols. To test membership, one can measure each of these stabilizers. The logical operators of the code become $\overline{X} = X^{\otimes 7}$ and $\overline{Z} = Z^{\otimes 7}$.

## 2.2 Definitions

**Quantum error correcting code** We will refer to a code $C$ as a $[[n,k,d]]$ code where $n$ is the number of physical qubits (i.e. the states are elements of $(\mathbb{C}^2)^{\otimes n}$), $k$ is the dimension of the code-space, and $d$ is the distance of the code. In the context of a system with more than $n$ qubits, the qubits corresponding to the physical code will be referred to as the code register — i.e. for state $\rho$, the reduced density matrix of $\rho$ on the code register is referred to as $\rho_{\text{code}}$. We say that a state $\rho$ (on $n' \geq n$ qubits) is a code-state if $\rho_{\text{code}}$ is a mixed state supported on the vectors of $C$. We can define distance precisely using the Knill-Laflamme conditions [49].

Let $\{|\overline{x}\rangle\} \subseteq C$ be an orthonormal basis for $C$ parameterized by $x \in \{0,1\}^k$. The Knill-Laflamme conditions state that the code can correct an error $E$ iff

$$\langle \overline{x}| E |\overline{y}\rangle = \begin{cases} 0 & x \neq y \\ \eta_E & x = y \end{cases} \tag{2.5}$$

where $\eta_E$ is a constant dependent on $E$. This is equivalent to

$$\Pi_C E \Pi_C = \eta_E \Pi_C \tag{2.6}$$

where $\Pi_C$ is the projector onto the code-space. We say that the code $C$ has distance $d$ if it can correct all Pauli-errors of weight $< d$. By linearity, it is equivalent to correcting all errors of weight $< d$. Furthermore, given a set $S$ of fewer than $d$ qubits, the reduced density matrix $\rho_S$ of any code-state $\rho$ on the set $S$ is an invariant of the code (Fact 2.1). It can be derived as a direct consequence of the Knill-Laflamme conditions.

**Proof of Fact 2.1:** Let $E$ be any operator whose support is entirely contained in $S$. Then for any code-state $\rho$,

$$\begin{aligned} \operatorname{tr}(E\rho) &= \operatorname{tr}(E\Pi_C\rho\Pi_C) && (2.7a) \\ &= \operatorname{tr}(\Pi_C E \Pi_C \rho) && (2.7b) \\ &= \operatorname{tr}(\eta_E \Pi_C \rho) && (2.7c) \\ &= \eta_E && (2.7d) \end{aligned}$$

where eq. (2.7a) is the Knill-Laflamme condition (eq. (2.6)), eq. (2.7b) is due to cyclicality of trace, eq. (2.7c) is an application of eq. (2.6), and eq. (2.7d) is because $\rho$ has trace 1. Since this equality holds for any operator $E$ and $\eta_E$ is a constant independent of $\rho$ such that $\eta_E = \operatorname{tr}(E\rho) = \operatorname{tr}(E\rho_S)$, then $\rho_S$ is an invariant of the code-state $\rho$. □

Given a code $C$ and a state $\sigma$ on $n$ qubits, we define the trace-distance between $\sigma$ and $C$ as $\inf_{\rho \in C} \|\rho - \sigma\|_1$. If the code-space $C$ can be defined as the ground space of a commuting set of projectors (see Definition 3.6) we will call it a commuting code. A special subcase is *stabilizer* codes when it can be expressed as the simultaneous eigenspace of a subgroup of Pauli operators.

### 2.2.1 Stabilizer codes

**Definition 2.2 (Pauli group)** *The Pauli group on n qubits, denoted by $\mathcal{P}_n$ is the group generated by the n-fold tensor product of the Pauli matrices*

$$\mathbb{I}_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad Y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \textit{ and } \quad Z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{2.8}$$

**Definition 2.3 (Stabilizer Code)** *Let $\{C_i\}_{i\in[m]}$ be a collection of commuting Pauli operators from $\mathcal{P}_n$ and $\mathcal{S}$ be the group generated by $\{C_i\}$ with multiplication. The stabilizer error correcting code $C$ is defined as the simultaneous $+1$ eigenspace of each element of $\mathcal{S}$:*

$$C = \left\{ |\psi\rangle \in (\mathbb{C}^2)^{\otimes n} : C_i |\psi\rangle = |\psi\rangle \ \forall i \in [m] \right\}. \tag{2.9}$$

*More generally, for every $s \in \{0,1\}^m$, define the space $D_s$ as*

$$D_s = \left\{ |\psi\rangle \in (\mathbb{C}^2)^{\otimes n} : C_i |\psi\rangle = (-1)^{s_i} |\psi\rangle \ \forall i \in [m] \right\}. \tag{2.10}$$

*In this language, $C = D_{0^m}$. The logical operators $\mathcal{L}$ are the collection of Pauli operators that commute with every element of $\mathcal{S}$ but are not generated by $\mathcal{S}$:*

$$\mathcal{L} = \{ P \in \mathcal{P}_n : PC_i = C_i P \ \forall i \in [m] \} \setminus \mathcal{S}. \tag{2.11}$$

*We say that the code is $\ell$-local if every $C_i$ is trivial on all but $\ell$ components of the tensor product and that each qubit of the code is non-trivial in at most $\ell$ of the checks $\{C_i\}$.*

Given a stabilizer code defined by $\{C_i\}_{i\in[m]}$, the associated local Hamiltonian is defined by

$$\mathbf{H} = \sum_{i\in[m]} h_i \overset{\text{def}}{=} \sum_{i\in[m]} \frac{\mathbb{I} - C_i}{2}. \tag{2.12}$$

This Hamiltonian is therefore commuting and it is a $\ell$-local low-density parity check Hamiltonian where $\ell$ is the locality of the code. Furthermore, the eigenspaces of $H$ are precisely the spaces $\{D_s\}$ with corresponding eigenvalues of $|s|$, the Hamming weight of $s$. If the rate of the stabilizer code is $k$, we can identify a subset of $2k$ logical operators denoted as

$$\overline{X_1}, \overline{Z_1}, \ldots, \overline{X_k}, \overline{Z_k} \tag{2.13}$$

such that all operators square to identity and pairwise commute except $\overline{X_i}$ and $\overline{Z_i}$ which anti-commute for all $i \in [k]$.

**Fact 2.4** *Let $\rho$ be a state such that $\rho_{\text{code}} \in D_s$ for a string $s$. For a logical pauli $P \in \mathcal{L}$, define $\rho' = P\rho P$. It holds that for any region $T \subset [n']$ of size less than d, $\rho_T = \rho'_T$. In general, let $\rho, \rho'$ be states such that $\rho_{\text{ancilla}} = \rho'_{\text{ancilla}}$ and $\rho_{\text{code}}, \rho'_{\text{code}} \in D_s$ for a string s. It holds that for any region $T \subset [n']$ of size less than d, $\rho_T = \rho'_T$.*

**Proof:**

$$\operatorname{tr}(P\rho) = \operatorname{tr}\left(P\left((-1)^{s_i} C_i \rho\right)\right) = (-1)^{s_i+1} \operatorname{tr}(C_i P \rho) = (-1)^{s_i+1} \operatorname{tr}(\rho C_i P) = -\operatorname{tr}(P\rho). \qquad (2.14)$$

where we used the cyclicality of trace twice. □

### 2.2.2 CSS codes [50, 51]

To formalize this property, recall a CSS code with parameters $[[n, k, d]]$. The code is constructed by taking two classical codes $C_{\mathsf{X}}$ and $C_{\mathsf{Z}}$ such that $C_{\mathsf{Z}} \supset C_{\mathsf{X}}^{\perp}$. The code $C_{\mathsf{Z}}$ is the kernel of a row- and column-sparse matrix $H_{\mathsf{Z}} \in \mathbb{F}_2^{m_{\mathsf{Z}} \times n}$; the same for $C_{\mathsf{X}}$ and $H_{\mathsf{X}} \in \mathbb{F}_2^{m_{\mathsf{X}} \times n}$. The rank of $H_{\mathsf{Z}}$ will be denoted as $r_{\mathsf{Z}}$ and likewise $r_{\mathsf{X}}$ is the rank of $H_{\mathsf{X}}$. Therefore, $n = k + r_{\mathsf{X}} + r_{\mathsf{Z}}$. If the code is constant-rate and linear-distance, then $k, d, r_{\mathsf{X}}, r_{\mathsf{Z}} = \Omega(n)$. For the codes considered in this work, we also have $m_{\mathsf{Z}}, m_{\mathsf{X}} = \Omega(n)$.

**Definition 2.5 (Code distance metric)** *In addition to the standard Hadamard metric on the boolean hypercube, it will be helpful to define the following two "code distance metrics": $|\cdot|_{C_{\mathsf{X}}^{\perp}}$ and $|\cdot|_{C_{\mathsf{Z}}^{\perp}}$ where for any subset $S \subset \{0,1\}^n$, let the distance measure $|\cdot|_S$ as $|y|_S = \min_{s \in S} |y + s|$ where $|\cdot|$ denoted Hamming weight.*

The $X$- and $Z$- distances of a code can be expressed as

$$d_{\mathsf{X}} \overset{\text{def}}{=} \min \left\{ |y|_{C_{\mathsf{Z}}^{\perp}} \ : \ y \in C_{\mathsf{X}} = \ker H_{\mathsf{X}} | \right\} \qquad (2.15a)$$

$$d_{\mathsf{Z}} \overset{\text{def}}{=} \min \left\{ |y|_{C_{\mathsf{X}}^{\perp}} \ : \ y \in C_{\mathsf{Z}} = \ker H_{\mathsf{Z}} | \right\} \qquad (2.15b)$$

$$d = \min\{d_{\mathsf{X}}, d_{\mathsf{Z}}\}. \qquad (2.15c)$$

### 2.2.3 Quantum LDPC and locally testable codes

**Definition 2.6 (QLDPC)** *Consider a code $C$ on $n$ qubits which is the simultaneous $+1$ eigenspace of a collection of projectors $\Pi_1, \ldots, \Pi_m$ on $n$ qubits. It is a* low-density parity check *(LDPC) code if each $\Pi_i$ acts non-trivially on at most $\ell$ physical qubits and each physical qubit is acted on non-trivially by at most $\ell$ projectors for $\ell = O(1)$.*

The construction of quantum LDPC codes has been a problem of great interest; in particular, the construction of good quantum LDPC codes which are codes with linear-rate and linear-distance scaling parameters [36–43].

The other object of interest is quantum locally testable codes. Although we do not use quantum locally testable codes in the proof of the NLTS theorem, it is worth noting their definition as [52] previously showed that locally testable codes of linear-distance would prove the NLTS theorem. We discuss their implications on the quantum PCP conjecture in Chapter 5.

**Definition 2.7** *For any state $|\psi\rangle$ and subspace $C$, define the error-distance between $|\psi\rangle$ and $C$ as the minimum weight error $E$ such that $E|\psi\rangle \in C$. The weight of $E$ is the number of physical qubits that the operator $E$ acts non-trivially on.*

**Definition 2.8 (Testable codes [53])** *A code $C$ is $\alpha$-testable if for any state $|\psi\rangle$, if $D$ = the error-distance between $|\psi\rangle$ and $C$, then*

$$\sum_{i=1}^{m} \langle\psi|\Pi_i|\psi\rangle \geq \alpha \cdot \frac{Dm}{n}. \tag{2.16}$$

*Any LDPC $\alpha$-testable code with $\alpha = \Omega(1)$ is a locally testable code.*

Currently, no constructions for constant $\alpha$ are known to exist.

## 2.3 Tanner codes

For a regular graph $G = (V, E)$ with degree $d$ and $|V| = n$ vertices and a classical linear code $C \subset \{0, 1\}^d$, we can construct a classical LDPC error correcting code on $m = |E|$ bits. Given a family of $d$-regular graphs for scaling $n$, this gives a construction of classical codes. The code will be referred to as $T = T(C, G)$.

The physical bits of $T$ correspond to the edges of the graph and for each vertex $v \in V$, the corresponding local check term verifies that the bits on the edges adjacent to $v$ are a member of $C$ (the order of edges adjacent to $v$ is implicitly defined). Therefore, there are $n$ local check terms and $m = \Theta(n)$ physical bits.

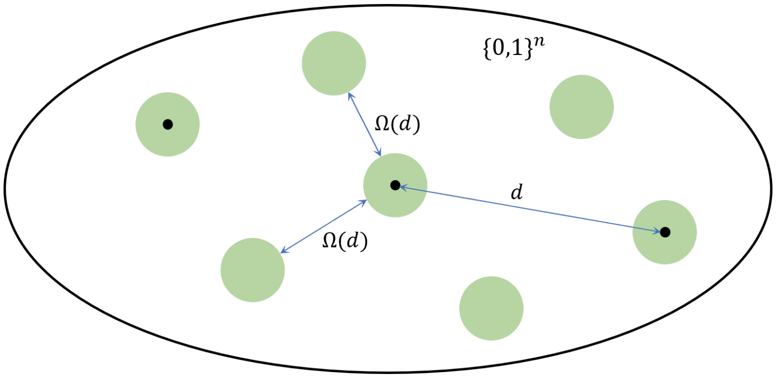

Figure 2.1: A cartoon of the low-energy space of a Tanner code. The black dots are codewords; the low-energy space of a Tanner code is a collection of clusters with the distance between clusters scaling as $\Omega(d)$. The upper bound on the diameter of each cluster is $O(\delta n)$. Most clusters do not have a codeword within them; this is the difference between a generic Tanner code and a locally testable code.

Define $\Delta_0$ as the distance of $C$, and $\lambda$ is the spectral expansion of $G$ — i.e. $\lambda = \max\{|\lambda_2, \lambda_n|\}$ where $\lambda_1 \geq \ldots \geq \lambda_n$ are the eigenvalues of the adjacency matrix of $G$. It is well known (see [45, Lemma 4]) that the distance of $T$ is lower bounded by $\frac{\Delta_0^2}{2d} \cdot m$ and the rate is $m - n$.

When the expansion of the graph $G$ is good, we notice the following property.

**Property 2.9** *If $y \in \{0,1\}^m$ is a word satisfying most checks, $|Hy| \leq \delta n$ for small $\delta$ where $H$ is the linear check matrix of $T$, then there exist constants $c_1, c_2$ such that*

$$\textit{either } |y| \leq c_1 \delta n \qquad \textit{or} \qquad |y| \geq c_2 n. \tag{2.17}$$

We call this property the "clustering of approximate code-words" since the low-energy subspace of the code forms clusters that are all far from each other (see Figure 2.1). Property 2.9 was used to prove the combinatorial NLTS theorem [45, Theorem 4.3]. The quantum version of this property, Property 4.2, is the additional ingredient needed to prove the NLTS theorem.

While this property holds for Tanner codes, the following lemma shows that more general classical codes with small-set expanding interactions graphs satisfy Property 2.9.

**Definition 2.10** *Let $G$ be a $d$-left-regular bipartite graph between vertex sets $L$ and $R$. A subset $A \subset L$ is said to be $\gamma$-expanding if $|\Gamma(A)| \geq (1-\gamma)d|A|$ where $\Gamma(A) \subset R$ is the set of neighbors of $A$. We say that $G$ is $(\gamma, \alpha)$-small set expanding if every set $A$ of size $\leq \alpha|L|$ is $\gamma$-expanding. (See Figure 2.2).*

**Lemma 2.11** *For a classical error correcting code with check matrix $H \in \mathbb{F}_2^{m\times n}$ , draw the interaction graph $G$ between the set of vertices, $V = [n]$, and the set of checks, $C = [m]$, with an*

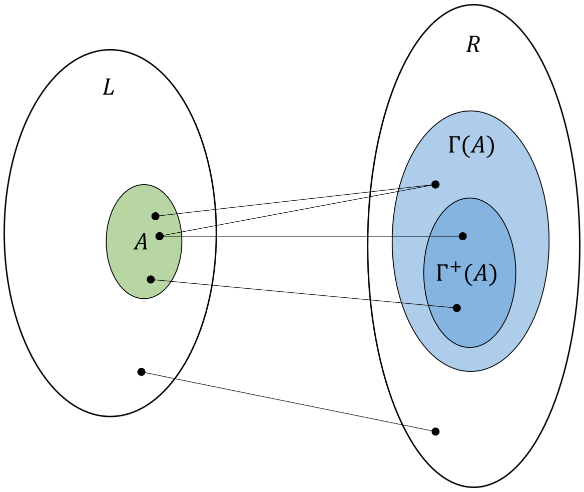

Figure 2.2: A cartoon of small-set expansion (Definition 2.10).

*edge $v \sim c$ if $v$ participates in the check $c$. If $G$ is $(\gamma, \alpha)$-small set expanding for $\gamma < \frac{1}{2}$, then the code satisfies Property 2.9.*

**Proof:** Consider any $y \in \{0,1\}^n$. If $|y| < \alpha n$, then $y$ is the indicator vector for a small subset $A \subset V$, and $|\Gamma(A)| \geq (1-\gamma)d|A|$. Let $\Gamma^+(A)$ be the subset of $\Gamma(A)$ with a unique neighbor in $A$. Since the number of edges between $A$ and $\Gamma(A)$ is $d|A|$, then

$$d|A| \geq \left|\Gamma^+(A)\right| + 2 \cdot \left(|\Gamma(A)| - \left|\Gamma^+(A)\right|\right) \tag{2.18a}$$

$$= -\left|\Gamma^+(A)\right| + 2(1-\gamma)d|A| \tag{2.18b}$$

Therefore, $|\Gamma^+(A)| \geq (1-2\gamma)d|A|$. Since every check in $\Gamma(A)$ is adjacent to a unique vertex in $A$, $\Gamma^+(A)$ is a subset of the checks that will be violated by $y$. Set $c_2 \overset{\text{def}}{=} \alpha$ and $c_1 \overset{\text{def}}{=} \frac{m}{(1-2\gamma)dn}$. If $|y| < \alpha n$, then

$$\delta m \geq |Hy| \geq \left|\Gamma^+(A)\right| \geq (1-2\gamma)d|A| = (1-2\gamma)d|y|. \tag{2.19}$$

This shows that, in fact, $|y| < c_1 \delta n$. □

### 2.3.1 Quantum Tanner codes

Quantum Tanner codes, defined by Leverrier and Zémor [43] are one generalization of Tanner codes to the quantum setting. In [43], Leverrier and Zémor showed that the construction generates linear-rate and linear-distance quantum codes. We show in Section 4.2 that quantum Tanner codes also cluster approximate code-words (Property 4.2). This is necessary to show these codes are NLTS.

**Definition of quantum Tanner codes** For a group $G$, consider a right Cayley graph $\mathrm{Cay}^r(G, A)$ and a left Cayley graph $\mathrm{Cay}^\ell(G, B)$ for two generating sets $A, B \subset G$, which are assumed to be symmetric, i.e. $A = A^{-1}$ and $B = B^{-1}$ and of the same cardinality $\Delta = |A| = |B|$. Further, we define the double-covers of $\mathrm{Cay}^r(G, A)$ and $\mathrm{Cay}^\ell(G, B)$ that we will denote $\mathrm{Cay}_2^r(G, A)$ and $\mathrm{Cay}_2^\ell(G, B)$.[2] The vertex sets of $\mathrm{Cay}_2^r(G, A)$ and $\mathrm{Cay}_2^\ell(G, B)$ are $\{\pm\} \times G$ and $G \times \{\pm\}$, respectively. The edges of $\mathrm{Cay}_2^r(G, A)$ are labeled by $A \times G$ and are of the form $(g, +) \sim (ag, -)$. Similarly, the edges of $\mathrm{Cay}_2^\ell(G, B)$ are labeled by $G \times B$ and are of the form $(+, g) \sim (-, gb)$.

Quantum Tanner codes are defined on the balanced product of the two Cayley graphs $X' = \mathrm{Cay}_2^r(G, A) \times_G \mathrm{Cay}_2^\ell(G, B)$, see [40, Section IV-B]. It is given by the Cartesian product $\mathrm{Cay}_2^r(G, A) \times \mathrm{Cay}_2^\ell(G, B)$ with the (canonical) anti-diagonal action of $G$ factored out. To understand the set of vertices $V'$ of $X'$, we first note that the vertices of the Cartesian product are labeled by $\{\pm\} \times G \times G \times \{\pm\}$. The group $G$ acts via right-multiplication on the left copy of $G$ and via inverse left-multiplication on the right copy of $G$. Factoring out this action identifies the vertices $(\pm, a, b, \pm)$ with $(\pm, ag, g^{-1}b, \pm)$

[2] The reason for defining the double-covers is convenience; the covering allows us to label each edge directly by specifying a vertex (group element) and a generator, which is not immediately possible in the original Cayley graphs.

for all $g \in G$. This means that two vertices $(\pm, a, b, \pm)$ and $(\pm, c, d, \pm)$ are identified if and only if $ab = cd$ and the outer signs agree. By passing from these equivalence classes to $ab \in G$, we obtain a unique labeling of the vertices $V'$ of $X'$ by $\{\pm\} \times G \times \{\pm\}$. Thus, $V'$ can be partitioned into the *even-parity vertices* $V'_0$, which are all vertices of the form $(+, g, +)$ and $(-, g, -)$, and the *odd-parity vertices* $V'_1$, which are all vertices of the form $(+, g, -)$ and $(-, g, +)$. The complex $X'$ is called the "quadripartite version" in [43].

Note that besides the natural action of $G$, there is an addition action of $\mathbb{Z}_2 = \langle \sigma \rangle$ on $\mathrm{Cay}^r_2(G, A)$ and $\mathrm{Cay}^\ell_2(G, B)$, which operates on the labels $\{\pm\}$ via $\sigma(+) = -$ and $\sigma(-) = +$. Hence, there is an operation of the group $G \times \mathbb{Z}_2$. We can thus analogously define the alternative balanced product complex $X = \mathrm{Cay}^r_2(G, A) \times_{(G\times\mathbb{Z}_2)} \mathrm{Cay}^\ell_2(G, B)$. The complex $X$ is called the "bipartite version" in [43]. Here, we will consider the complex $X$ instead of $X'$. Using the same arguments as previously for $X'$, we see that the vertices $V$ of $X$ can be labeled by $G \times \{\pm\}$ which fall into the sets $V_0$, which are all vertices of the form $(g, +)$, and $V_1$, which are all vertices of the form $(g, -)$.

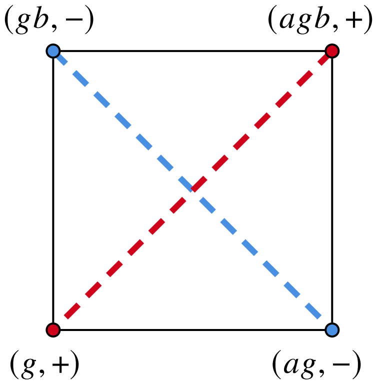

Figure 2.3: A face of the balanced product complex $X = \mathrm{Cay}^r_2(G, A) \times_{(G\times\mathbb{Z}_2)} \mathrm{Cay}^\ell_2(G, B)$. Each face is incident to two vertices in $V_0$ (red) and two vertices in $V_1$ (blue). This fact is used in [43] to define two graphs $\mathcal{G}^\square_0$ and $\mathcal{G}^\square_1$ whose edges connect the vertices in $V_0$ (red dashed line) and $V_1$ (blue dashed line), respectively. Importantly, the edge-sets of $\mathcal{G}^\square_0$ and $\mathcal{G}^\square_1$ are both in one-to-one correspondence with the faces of $X$ (and thus with each other).

The quantum Tanner code is now defined as follows. From the balanced product complex $X$ we define two graphs $\mathcal{G}^\square_0$ and $\mathcal{G}^\square_1$. The vertices of $\mathcal{G}^\square_0$ are the vertices in $V_0$. Note that there are exactly two vertices belonging to $V_0$ per face in $X$, see Figure 2.3. Hence, we connect two vertices by an edge in $\mathcal{G}^\square_0$ if and only if they belong to the same face, or equivalently, all edges in $\mathcal{G}^\square_0$ are of the form $(g, +) \sim (agb, +)$. Similarly, we can define the graph $\mathcal{G}^\square_1$ using the fact that there are exactly two vertices in $V_1$ per face. Note that both $\mathcal{G}^\square_0$ and $\mathcal{G}^\square_1$ are regular graphs of degree $\Delta^2$, as edges surrounding a vertex are labeled by $A \times B$. Further, $\mathcal{G}^\square_0$ and $\mathcal{G}^\square_1$ are expanders: Let $\lambda(\mathcal{G}) = \max\{|\lambda_2(\mathcal{G})|, |\lambda_n(\mathcal{G})|\}$, where $\lambda_2(\mathcal{G}), \lambda_n(\mathcal{G})$ are the second largest and the smallest eigenvalues of the adjacency matrix of the graph $\mathcal{G}$.

**Lemma 2.12 ( [43, Lemma 4])** *If* $\mathrm{Cay}^r(G,A)$ *and* $\mathrm{Cay}^\ell(G,B)$ *are Ramanujan graphs, then*

$$\lambda(\mathcal{G}_0^\square), \lambda(\mathcal{G}_1^\square) \leq 4\Delta. \tag{2.20}$$

Taking two suitable local codes $C_A, C_B \subset \mathbb{F}_2^\Delta$, we define $C_0 = C_A \otimes C_B$ and $C_1 = C_A^\perp \otimes C_B^\perp$. Finally, we define Tanner codes $C_{\mathsf{z}} = C(\mathcal{G}_0^\square, C_0^\perp)$ and $C_{\mathsf{x}} = C(\mathcal{G}_1^\square, C_1^\perp)$ [54,55]. It can be shown [43] that $C_{\mathsf{z}} \supset C_{\mathsf{x}}^\perp$, so that we obtain a well-defined CSS code.

For these codes to have linear-rate and linear-distance, the graphs and local codes need to fulfill certain conditions: The Cayley graphs are required to be Ramanujan expanders [56, 57]. Further, the local codes are required to be *robust* and *resistant to puncturing*. More precisely, we call $C_1^\perp = (C_A^\perp \otimes C_B^\perp)^\perp = C_A \otimes \mathbb{F}_2^B + \mathbb{F}_2^A \otimes C_B$ *$w$-robust* if any codeword $|x|$ of Hamming weight bounded as $|x| \leq w$ has its support included in $|x|/d_A$ columns and $|x|/d_B$ rows, where $d_A$ and $d_B$ are the minimum distances of $C_A$ and $C_B$, respectively. Further, $C_1^\perp$ has $w$-robustness with *resistance to puncturing $p$* if for any $A' \subset A$, $B' \subset B$ with $|A'|, |B'| \geq \Delta - w'$ with $w' \leq p$ the code $C_1^\perp$ remains $w$-robust when punctured outside of $A' \times B'$.

# Chapter 3

# Trivial states, Hamiltonians, and indistinguishability

"There is still one of which you never speak."
Marco Polo bowed his head.
"Venice," the Khan said.
Marco smiled. "What else do you believe I have been talking to you about?"
The emperor did not turn a hair. "And yet I have never heard you mention that name."
And Polo said: "Every time I describe a city I am saying something about Venice."

Italo Calvino, *Invisible Cities*

## 3.1 Quantum circuits

Quantum circuits are a model for quantum computation generalizing classical boolean (reversible) circuits. A quantum circuit is a unitary $U$ parametrized by a depth $t$ such that $U = U_t \dots U_1$ where each $U_j$ is a unitary acting on $(\mathbb{C}^2)^{\otimes n}$ and $U_j = \bigotimes_i U_{ji}$, a tensor product of disjoint two-qubit unitaries $U_{ji}$. This describes a quantum circuit with gates of fan-in and fan-out 2 and all-to-all connectivity; the choice of fan-in and fan-out of 2 is equivalent to any other constant since we will be caring about asymptotic behavior. We often care about the output of circuits run from the initial state $|0^{n'}\rangle$.

**Definition 3.1 (Circuit Complexity)** *Let $\rho$ be a mixed quantum state of $n$ qubits. Then the circuit complexity*[1] *of $\rho$,* $\mathrm{depth}(\rho)$*, is defined as the minimum depth over all $n'$-qubit quantum circuits $U$*

[1] We note that while our definition for circuit complexity of $\rho$ is given as the minimum depth of any circuit exactly generating a state $\rho$, we could have equivalently defined the circuit complexity of $\rho$ as the minimum depth of any circuit generating a state $\rho'$ within a small ball $B_\delta(\rho)$ of $\rho$ for some $\delta > 0$. This would not have changed our results except for constant factors. This is because our results will be concerned with lower bounding the circuit complexity of all states of energy $\leq \epsilon$. If $\rho$ is a state of energy $\leq \epsilon$, then every state $\rho' \in B_\delta(\rho)$ has energy $\leq \epsilon + \delta$. Therefore, by

*such that* $U|0^{\otimes n'}\rangle \in (\mathbb{C}^2)^{\otimes n'}$ *is a purification of* $\rho$. *Equivalently,*

$$\mathrm{depth}(\rho) \overset{\mathsf{def}}{=} \min\left\{\mathrm{depth}(U) : \mathrm{tr}_{[n']\setminus[n]}\left(U|0^{\otimes n'}\rangle\langle 0^{\otimes n'}|U^\dagger\right) = \rho\right\}. \tag{3.1}$$

*A family of states* $\{\rho_n\}$ *for growing n is called* trivial *if there exists a constant c such that* $\mathrm{depth}(\rho_n) \le c$ *for all n.*

**Lightcones** For an operator $A$ (think of $A$ as a projector onto a single qubit) and a circuit $U$, let the lightcone of $A$, with respect to $U$, be the set of qubits on which $UAU^\dagger$ acts non-trivially. For a qubit $i$, let the lightcone of qubit $i$ be the union of lightcones over all operators $A$ supported on qubit $i$.

**Fact 3.2** *For a circuit U of depth t, the size of the lightcone of qubit i is* $\le 2^t$.

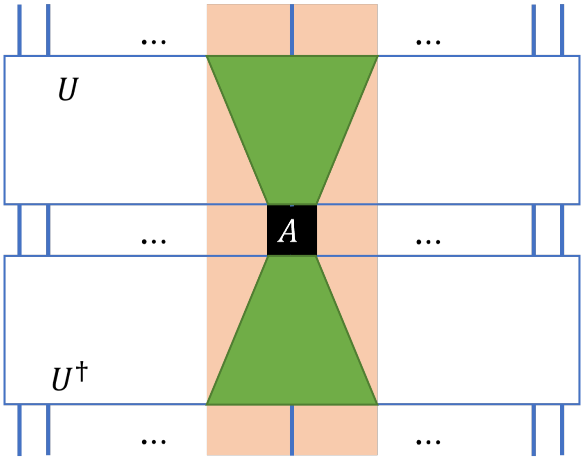

Figure 3.1: Every gate outside the green regions cancel with their conjugate across the picture. Only the gates in the green regions (the lightcone) do not trivially cancel with another gate. The pink region denotes the set of qubits in the lightcone. Notice that $UAU^\dagger$ equals the product of the unitaries in the pink region. If we denote $U_R$ as the unitaries $U$ restricted to only the qubits of the region $R$, then $UAU^\dagger = U_{L_A}AU_{L_A}^\dagger$ where $L_A$ is the lightcone of $A$.

**Proof:** One can show, intuitively, that $U_i \ldots U_1 A U_1^\dagger \ldots U_i$ acts non-trivially on at most $2^i$ qubits since each $U_i$ is the tensor product of 2 qubit unitaries. □

If the circuit was geometrically constrained to a lattice of a fixed constant dimension $D$, then the simple upper bound would be $O((tD)^D)$. All our proofs can easily be translated into lower bounds for geometric circuits on a lattice using this substitution.

redefining $\epsilon \leftarrow \epsilon - \delta$, we can switch to the alternate definition of circuit complexity. We use the listed definition in our proofs as it vastly simplifies legibility. However for variations of the NLTS theorem, such as the combinatorial NLTS theorem [45], we need to consider a robust definition since the same argument no longer holds.

**Fact 3.3** *Consider a quantum state $\psi$ acting on $(\mathbb{C}^2)^{\otimes m}$. For any $i \in [n']$, let $L_i$ denote the support of the lightcone of $i$ with respect to $U$. It holds that*

$$\mathrm{tr}_{-\{i\}}(U\psi U^\dagger) = \mathrm{tr}_{-\{i\}}\left(U(\psi_{L_i} \otimes \nu_{-L_i})U^\dagger\right). \tag{3.2}$$

*In other words, the reduced density matrix on qubit $i$ only depends on the reduced density matrix on the lightcone, $\psi_{L_i}$*

**Proof:** where the uniformly distributed quantum state on a Hilbert space $\mathcal{H}$ is represented by $\nu_{\mathcal{H}} \overset{\text{def}}{=} \mathbb{I}_{\mathcal{H}}/|\mathcal{H}|$. For any operator $O$ supported on qubit $i$, consider

$$\mathrm{tr}_{\{i\}}(O\,\mathrm{tr}_{-\{i\}}(U\psi U^\dagger)) = \mathrm{tr}\Big(U^\dagger O U \psi\Big) = \mathrm{tr}\Big(U^\dagger O U \psi_{L_i} \otimes \nu_{-L_i}\Big) \tag{3.3a}$$

$$= \mathrm{tr}_{\{i\}}(O\,\mathrm{tr}_{-\{i\}}(U(\psi_{L_i} \otimes \nu_{-L_i})U^\dagger)). \tag{3.3b}$$

The second equality uses $U^\dagger O U = U_{L_i}^\dagger O U_{L_i}$ where $U_{L_i}$ is the circuit restricted to the region $L_i$ (see Figure 3.1 for proof). This proves the fact. □

**Fact 3.4** *Consider a quantum state $|\phi\rangle = U|0^{\otimes m}\rangle$. Let $R \subset [n']$ and define $|\phi'\rangle = U_R|0^{\otimes n'}\rangle$. We have* $\mathrm{tr}_{-R}(\phi) = \mathrm{tr}_{-R}(\phi')$.

**Proof:** The proof is very similar to that of Fact 3.3. For any operator $O$ supported on $R$, consider

$$\mathrm{tr}(O\,\mathrm{tr}_{-R}(\phi)) = \mathrm{tr}\Big(U^\dagger O U |0^{n'}\rangle\langle 0^{n'}|\Big) = \mathrm{tr}\Big(U_R^\dagger O U_R |0^{n'}\rangle\langle 0^{n'}|\Big) = \mathrm{tr}(O\,\mathrm{tr}_{-R}(\phi')). \tag{3.4}$$

This completes the proof. □

## 3.2 Local Hamiltonians

### 3.2.1 Constraint satisfaction problems

The fundamental object in the study of *classical* non-deterministic computation, NP, is the $k$-local constraint satisfaction problem (CSP). A $k$-variable clause over an alphabet $\Sigma$ (in most cases $|\Sigma| = 2$ — i.e. boolean alphabet) is a function $C : \Sigma^k \to [0,1]$ assigning a score to any assignment to $k$-variables over $\Sigma$.

**Definition 3.5 (Constraint Satisfication Problem)** *A $k$-CSP $C$ is a formula on $n$ variables over $\Sigma$, composed of $m$ $k$-variable clauses $C_i$ on subsets of the $n$ variables. The value of the $k$-CSP $C$ for any assignment $x \in \Sigma^n$ is*

$$C(x) \overset{\text{def}}{=} \sum_{i=1}^{m} C_i(x) \tag{3.5}$$

*where $C_i(x)$ equals $C_i(x|_{S_i})$ where $S_i$ is the subset of $k$-variables that clause $C_i$ acts on. A CSP is satisfiable if $\exists x \in \Sigma^m$ s.t. $C(x) = 0$.*

The Cook-Levin theorem [10, 11] is equivalent to a proof that deciding whether or not a CSP instance is satisfiable is NP-complete and the PCP theorem [17–19] is equivalent to a proof that deciding whether a CSP is satisfiable or $\min_{x \in \{0,1\}^n} C(x) \geq m/2$ is NP-complete.

### 3.2.2 A quantum analog: local Hamiltonians

CSPs capture the local-to-global phenomenon of classical non-deterministic computation. Perhaps surprisingly, there is an analog for quantum objects; a quantum local-to-global phenomenon that captures the complexity of quantum non-deterministic computation. Even more surprising, the objects generated have been central to the study of condensed matter physics. Condensed matter physics concerns itself with the properties of $n$ interacting particles whose interactions are governed by quantum mechanics. Calculating the minimum energy of a condensed matter system — i.e. the ground energy — is a central problem in that field. The operator which describes the energy is called a *Hamiltonian* and in physical systems of interest, the operator is the sum of many smaller Hamiltonian terms each governing the interaction of a few particles. A classical CSP on $n$ variables corresponds to a local Hamiltonian $\mathbf{H} = \sum_{i=1}^m h_i$ acting on $n$ qubits[2]. The analog of a solution to the CSP is an $n$ qubit quantum state solution to the local Hamiltonian, and the number of violated constraints corresponds to the energy (eigenvalue) of that quantum state. The NP-hardness of deciding if a CSP $C$ is satisfiable ($\exists x$ s.t. $C(x) = 0$) or is unsatisfiable ($\forall x,\ C(x) \geq 1/m$) corresponds to the QMA-hardness of deciding whether a local Hamiltonian $\mathbf{H}$ has minimum eigenvalue at most $a$ or at least $b$ for given $a, b$ such that $b - a = 1/\mathrm{poly}(n)$.

**Definition 3.6 (Local Hamiltonian Problem)** *A $k$-local Hamiltonian $\mathbf{H}$ is an operator acting on $n$ qudits (of constant dimension $d$) composed on $m$ $k$-local Hamiltonian terms $h_i$. Each term $h_i$ is a linear operator $\mathcal{L}((\mathbb{C}^d)^{\otimes k})$ such that $h_i^\dagger = h_i$ (Hermitian) with $\|h_i\| \leq 1$. Each local term $h_i$ is accompanied with a subset $S_i \subset [n]$ of $|S_i| = k$ denoting the terms that $h_i$ acts on. The local Hamiltonian $\mathbf{H}$ equals*

$$\mathbf{H} = \sum_{i=1}^{m} h_i \otimes \mathbb{I}_{[n]\setminus S_i} \tag{3.6}$$

*which we write as $\mathbf{H} = \sum_i h_i$ for brevity. The minimum energy, also known as the ground energy, of the Hamiltonian $\mathbf{H}$ is the minimum eigenvalue of the Hermitian matrix $\mathbf{H}$:*

$$\lambda_{\min}(\mathbf{H}) \overset{\text{def}}{=} \min_{|\psi\rangle \in (\mathbb{C}^d)^{\otimes k}} \langle\psi|\mathbf{H}|\psi\rangle = \min_{\rho} \mathrm{tr}(\mathbf{H}\rho) \tag{3.7a}$$

$$= \min_{|\psi\rangle \in (\mathbb{C}^d)^{\otimes k}} \sum_{i=1}^{m} \langle\psi|h_i|\psi\rangle = \min_{\rho} \sum_{i=1}^{m} \mathrm{tr}(h_i\rho). \tag{3.7b}$$

[2] For normalization, we assume that the terms of a local Hamiltonian have spectral norm at most 1.

*A Hamiltonian is said to be frustration-free if* $\lambda_{\min}(\mathbf{H}) = 0$*. The gap between the smallest and second-smallest eigenvalues of* $\mathbf{H}$ *is called the* spectral gap *of the Hamiltonian.*

The energy of a local Hamiltonian term or more generally interpreting $\langle h|\psi|h\rangle$, can be interpreted physically. If we write $h$ in a basis in which it is diagonal $h = \sum_j \lambda_j \, |j\rangle\langle j|$, then $\langle\psi|h|\psi\rangle = \sum_j \lambda_j |\langle j|\psi\rangle|^2$. Note this implicitly defines a probability distribution $p$ with $p(j) = |\langle j|\psi\rangle|^2$. Therefore $\langle\psi|h|\psi\rangle = \mathbf{E}_{j\sim p}\,\lambda_j$, or equivalently this is an expectation over the eigenvalues of $h$ over the probability distribution naturally induced by $|\psi\rangle$.

We call the eigenvalues of $\mathbf{H}$ the *energy levels* of the system. The energy of a state $|\psi\rangle$ with respect to $\mathbf{H}$ is $\langle\psi|\mathbf{H}|\psi\rangle$. A state is called a *ground state* if $\langle\psi|\mathbf{H}|\psi\rangle = \lambda_{\min}(\mathbf{H})$; analogously we define mixed ground states. The ground space is the linear span of all ground states and all mixed states supported on the space. Low-energy states[3] are all states of energy near the ground energy; in this thesis, it will refer to all states of energy $\leq \epsilon m$).

Computing the ground energy is a central problem in condensed matter physics. It was shown to be QMA-complete.

**Definition 3.7** (QMA) *A quantum circuit U acting on n qubits and consisting of T gates is a* QMA-*verifier circuit iff there exists* $w \leq n$ *qubits that are designated the* witness *register and the rest of the qubits form the* ancilla *register, and it satisfies the promise that either there exists a state* $|\xi\rangle \in (\mathbb{C}^2)^{\otimes w}$ *such that*

$$\mathbf{Pr}(U \textit{ accepts } |\xi\rangle \otimes |0^{n-w}\rangle) \geq 2/3 \tag{3.8}$$

*or for all states* $|\xi\rangle$*,*

$$\mathbf{Pr}(U \textit{ accepts } |\xi\rangle \otimes |0^{n-w}\rangle) \leq 1/3. \tag{3.9}$$

*By accept, we mean the event that measuring the first qubit of the state* $U|\xi\rangle \otimes |0^{n-w}\rangle$ *in the standard basis yields the* $|1\rangle$ *state. The constants* $2/3$ *and* $1/3$ *are arbitrary; they only need to be separated by a universal constant. The two cases are denoted* yes *and* no *instances, respectively. Deciding if a quantum circuit U is a* yes *or* no *instance is the canonical* QMA*-complete problem.*

**Theorem 3.8** (QMA**-completeness [13–15])** *There exists a quantum polynomial-time reduction from any* QMA*-verifier circuit U of n qubits and T gates to a local Hamiltonian* $\mathbf{H}$ *acting on* $n+T$ *qubits such that* $\mathbf{H}$ *has minimum energy* $\leq a$ *iff U is a* yes *instance and* $\mathbf{H}$ *has minimum energy* $\geq b$ *iff U is a* no *instance. The difference* $b - a \leq 1/\mathrm{poly}(n,T)$ *and is called the* promise gap *of the local Hamiltonian problem.*

Often, the QMA-completeness of the local Hamiltonian problem is expressed in terms of being hard to estimate the ground energy of a $n$-qubit local Hamiltonian problem to precision $1/\mathrm{poly}(n)$. While correct, it can mislead one into thinking that progress towards the quantum PCP conjecture

[3] Some papers express the Hamiltonian $\mathbf{H}$ as the expectation over terms instead of the sum. In which case it is the set of states of energy $\leq \epsilon$.

can be made by improving the polynomial $\mathrm{poly}(n)$ in precision. However, as stated, the precision can be improved to *any* polynomial by simply considering parallel copies of the original Hamiltonian problem. The original proof of Kitaev had a precision of $1/O(T^3)$ and this was improved to $1/O(T^2)$ in [58, 59]. Any improvement past $1/O(T^2)$ would be novel. Therefore, it is better[4] to define the precision in terms of the parameters of the original problem being reduced from and not the intermediate as to not cause this confusion.

## 3.3 Statements of the quantum PCP conjecture and NLTS theorem

**Conjecture 3.9 (Quantum PCP (formal statement) [20, 21])** *For some universal constants, $\ell, c$, there exists a quantum polynomial-time reduction from any* QMA*-verifier circuit $U$ of $n$ qubits and $T$ gates to a $\ell$-local Hamiltonian $\mathbf{H}$ acting on $n_1 = \mathrm{poly}(n, T)$ qubits and $m_1 = \Theta(n)$ terms such that $\mathbf{H}$ has minimum energy $\leq a$ iff $U$ is a* yes *instance and $\mathbf{H}$ has minimum energy $\geq b$ iff $U$ is a* no *instance where the difference $b - a \geq cm_1$.*

It is not too difficult to see that under quantum polynomial time reductions, this version of the quantum PCP conjecture is equivalent to a *proof-checking* version in which only a constant number of qubits of the proof are measured [21].

**Theorem 3.10 (NLTS (formal statement) [32])** *There exists a fixed constant $\epsilon > 0$ and an explicit family of $\ell$-local Hamiltonians $\mathbf{H}$ for an infinite set of integer values $n$, where $\mathbf{H}$ acts on $n$ particles, consists of $m = \Theta(n)$ local terms, such that for any family of states $\psi$ satisfying*

$$\mathrm{tr}(\mathbf{H}\psi) \leq \epsilon m + \lambda_{\min}(\mathbf{H}), \tag{3.10}$$

*the minimum depth of any quantum circuit generating $\psi$, $\mathrm{depth}(\psi)$, grows faster than any constant*[5]*.*

The following lemma proves that the NLTS theorem was a necessary consequence of the quantum PCP conjecture.

**Lemma 3.11** *If the NLTS theorem was false and the quantum PCP conjecture is true, then* NP = QMA.

[4] However, a feature of defining the quantum PCP conjecture in terms of a constant-fraction promise gap is that it automatically avoids this misconception.

[5] This definition is the one originally expressed by Freedman and Hastings in [32]. However, a consequence of the quantum PCP conjecture and NP ≠ QMA would be a circuit complexity lower bound of $\omega(\log\log n)$. For this reason, we will be more interested in circuit lower bounds of $\omega(\log\log n)$. Furthermore, if QCMA ≠ QMA, then the necessary consequence of the quantum PCP conjecture is a circuit lower bound of $\omega(\mathrm{poly}(n))$. Our techniques make no obvious progress towards this strengthened conjecture as we study stabilizer codes whose circuit complexity is $O(\log n)$. Some progress towards super-polynomial NLETS was made by Nirkhe, Vazirani, and Yuen [3].

**Proof:** If the NLTS theorem was false, then for every Hamiltonian $\mathbf{H}$ (of $n$ qubits and $m$ terms) and every $\epsilon > 0$, there exists a circuit $U$ of depth $O_\epsilon(1)$ such that $|\psi\rangle = U|0^{n'}\rangle$ such that $\mathrm{tr}(H\psi) \leq \epsilon m + \lambda_{\min}(\mathbf{H})$. Since the quantum PCP conjecture is true, there exists a family of Hamiltonians $\mathbf{H}$ and some values of $a$ and $b$ with promise gap $\geq cm$, such that it is QMA-hard to decide if the energy is $\leq a$ or $\geq b$.

Assume $\mathbf{H}$ is yes instance and for

$$\epsilon < \frac{b - \lambda_{\min}(\mathbf{H})}{m} \tag{3.11}$$

consider the witness which is the classical description of the circuit $U$. Then $\mathrm{tr}(\mathbf{H}\psi) < b$, a convincing proof that $\lambda_{\min}(\mathbf{H}) \leq a$ (due to the promise). Given a classical description of $U$, it is easy to verify that $\mathrm{tr}(H\psi) < b$, since

$$\mathrm{tr}(\mathbf{H}\psi) = \sum_{i=1}^{m} \langle 0^{n'}|U^\dagger h_i U|0^{n'}\rangle \tag{3.12}$$

and $U^\dagger h_i U$ acts non-trivially on at most $\ell \cdot 2^{O_\epsilon(1)} = O_\epsilon(1)$ qubits due to Fact 3.2. Therefore, we can calculate classically $\langle 0^{n'}|U^\dagger h_i U|0^{n'}\rangle$ by multiplying out the gates in the lightcone of $h_i$ and calculating the upper-left corner entry.

If $\mathbf{H}$ is a no instance, then $\lambda_{\min}(\mathbf{H}) \geq b$ and the minimum eigenvalue when restricted to low-depth states is at least $\lambda_{\min}(\mathbf{H})$ so a false proof will not be accepted. Therefore, the problem is in NP. However, since the problem was QMA-hard, then NP = QMA. □

## 3.4 Trivial states and local Hamiltonians

There is a fundamental connection between trivial states and local Hamiltonians that is incredibly useful in proving circuit depth lower bounds. In particular, it is useful for tying the theory of low-depth circuits to the theory of approximate ground state projectors [60].

To begin, we consider the 0-depth state, $|0^{n'}\rangle$. We notice that it is the ground state of the 1-local Hamiltonian:

$$\mathbf{H}_0 \stackrel{\text{def}}{=} \sum_{i=1}^{n'} |1\rangle\langle 1|_i \tag{3.13}$$

where $|1\rangle\langle 1|_i$ is the projector onto 1 of the $i$-th qubit. The Hamiltonian $\mathbf{H}_0$ has some simple properties worth recognizing: it is frustration-free, has a unique ground state of $|0^{n'}\rangle$, is commuting and has a spectral gap of 1. The eigenvectors are all the basis vectors $|x\rangle$ for $x \in \{0,1\}^{n'}$ and their respective eigenvalues are $|x|$.

For any circuit $U$ on $n'$ qubits of depth $t$, we can consider the Hamiltonian

$$\mathbf{H}_U \stackrel{\text{def}}{=} U\mathbf{H}_0U^\dagger = \sum_{i=1}^{n'} U|1\rangle\langle 1|_i U^\dagger. \tag{3.14}$$

Since $U$ is a unitary, the analysis of $\mathbf{H}_0$ directly translates to $\mathbf{H}_U$: it is frustration-free, has a unique ground state of $U\,|0^{n'}\rangle$, is commuting, and has a spectral gap of 1. The eigenvectors are all the vectors $U\,|x\rangle$ for $x \in \{0,1\}^{n'}$ and their respective eigenvalues are $|x|$. Furthermore, by Fact 3.2, the Hamiltonian $\mathbf{H}_U$ is $2^t$-local. This underlines the fundamental relationship between trivial states and local Hamiltonians.

**Fact 3.12** *Every trivial state $|\psi\rangle \in (\mathbb{C}^2)^{\otimes n'}$ is the unique ground state of a $O(1)$-local frustration-free commuting Hamiltonian with spectral gap* 1.

**Approximate ground state projectors** At this time it is helpful to recall the notion of approximate ground state projectors (AGSPs) which can help us improve the relationship between trivial-states and local Hamiltonians.

**Definition 3.13 (Approximate ground state projector)** *For a frustration-free Hamiltonian* $\mathbf{H}$ *with a unique ground state $|\Omega\rangle$. An operator $K$ is an $\delta$-approximate ground state projector (AGSPs) if*

$$\| |\Omega\rangle\langle\Omega| - K \| \leq \delta. \tag{3.15}$$

We can see that $(\mathbb{I} - \mathbf{H}_U/n')$ is a $(1 - 1/n')$-AGSP for $\mathbf{H}_U$ since 1 is the spectral gap of the Hamiltonian $\mathbf{H}_U$. It follows that polynomials $(\mathbb{I} - \mathbf{H}_U/n')^f$ are better AGSPs at the cost of locality: They are roughly $(1 - f/n')$-AGSPs but are $f \cdot 2^t$-local. We could ask if there is a better trade-off between the locality of the AGSP and $\delta$. Finding an optimal polynomial of $\mathbf{H}_U$ which maximizes the trade-off of the AGSP was studied and answered by [61, Theorem 3.1] (similar results exist in [62, 63]):

**Lemma 3.14 (Optimal polynomial approximation to the AND function [61])** *There exists a polynomial $P$ of degree $f \in (\sqrt{n'}, n')$ such that*

$$P(0) = 1, \qquad |P(i)| \leq \exp\left(-\frac{f^2}{100n'}\right) \textit{for } i = 1, 2, \ldots, n'. \tag{3.16}$$

A construction of $P$ can be built using Chebyshev polynomials; see [61, Theorem 3.1] for construction. By construction, $P(\mathbf{H}_U)$ will be a $\exp\left(-\frac{f^2}{100n'}\right)$-AGSP with a locality of $f \cdot 2^t$. Therefore, states of depth $O(\log n)$ are the ground states of very good AGSPs of locality $o(n)$. This comes in handy in proving circuit depth lower bounds.

## 3.5 Local indistinguishability and circuit depth lower bounds

Previously, we noted (Fact 2.1) that for every $[[n, k, d]]$ quantum error correcting codes, any code state $\rho$ and any subset $S \subset [n], |S| \leq d$, that $\rho_S$ was an invariant over the code space. This is an example of a more general property called local indistinguishability[6] which can be used to prove circuit depth lower bounds.

[6] It is called global entanglement in [21].

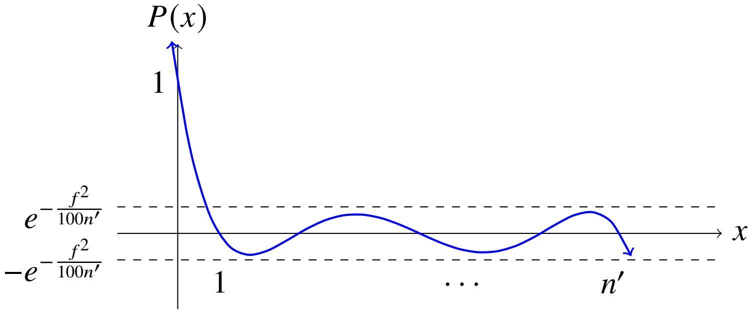

Figure 3.2: A cartoon of an optimal polynomial approximation to the AND function.

**Definition 3.15 (Local indistinguishability)** *Two states $\rho$ and $\varphi$ on $n$ qubits are said to be $d$-locally indistinguishable if for every subset $S \subset [n], |S| \leq d$, the density matrices $\rho_S$ and $\varphi_S$ are equal.*

Aside from error correcting codes, the best example of locally indistinguishable states are the $|\text{🐱}_+\rangle$ and $|\text{🐱}_-\rangle$ states:

$$|\text{🐱}_\pm\rangle \overset{\text{def}}{=} \frac{|0^n\rangle \pm |1^n\rangle}{\sqrt{2}}. \tag{3.17}$$

These states are some of the simplest examples of global multipartite entanglement. The ± phase of the state can only be recognized by viewing all $n$ qubits. View fewer than $n$ qubits and the reduced density matrix of either state is

$$\left(\text{🐱}_\pm\right)_{S\subsetneq[n]} = \frac{|0\rangle\langle 0|^{|S|} + |1\rangle\langle 1|^{|S|}}{2} \tag{3.18}$$

Consider a circuit constructing $|\text{🐱}_+\rangle$; the complexity of generating the state comes from the long-range correlations between all $n$ qubits. A common misconception is that it comes from the states $|\text{🐱}_+\rangle$ and $|\text{🐱}_-\rangle$ being far apart; that is not the case since a $Z$-gate applied to any qubit converts one to the other. Nevertheless, having both states allows for a simple circuit depth lower bound proof.

**Lemma 3.16 (Simple local indistinguishability lower bounds)** *If any two different $n$-qubit pure states $|\rho\rangle$ and $|\varphi\rangle$ are $d$-locally indistinguishable, then neither state can be generated by a circuit of depth $\Omega(\log d)$.*

**Proof:** Assume there exists a circuit $U$ of depth $t$ acting on $n$ qubits such that $|\rho\rangle = U|0^n\rangle$. Assume $2^t < d$. Consider the Hamiltonian $\mathbf{H}_U$ from earlier; its *unique* ground state will be $|\rho\rangle$ and it is $2^t$-local. Let $h_i = U|1\rangle\langle 1|_i U^\dagger$ be the local terms of $\mathbf{H}_U$; therefore $h_i$ acts non-trivially only on lightcone $L_i$. Then,

$$\langle\varphi|\mathbf{H}_U|\varphi\rangle = \sum_{i=1}^{n} \langle\varphi|h_i|\varphi\rangle \overset{(\star)}{=} \sum_{i=1}^{n} \operatorname{tr}\left(h_i\varphi_{L_i}\right) \overset{(\dagger)}{=} \sum_{i=1}^{n} \operatorname{tr}\left(h_i\rho_{L_i}\right) \overset{(\star)}{=} \langle\rho|\mathbf{H}_U|\rho\rangle = 0 \tag{3.19}$$

where both equality equations ($\star$) follow from the fact that the energy of the term $h_i$ only depends on the reduced density matrix state on the lightcone $L_i$ and the equality equation ($\dagger$) follows because the size of the lightcones $L_i$ are at most $2^t < d$ and therefore $\varphi_{L_i} = \rho_{L_i}$ by local indistinguishability. Therefore, $|\varphi\rangle$ is a ground state of $\mathbf{H}_U$ implies that it equals $|\rho\rangle$, a contradiction. So, $2^t \geq d$, proving the lower bound. □

Eq. 3.19 is useful enough that it's worth stating as its own fact.

**Fact 3.17** *If $|\rho\rangle$ and $|\phi\rangle$ are $d$-locally indistinguishable and $\mathbf{H}$ is a $d$-local Hamiltonian, then $\langle\rho|\mathbf{H}|\rho\rangle = \langle\phi|\mathbf{H}|\phi\rangle$. In other words, $\mathbf{H}$ cannot distinguish $\rho$ and $\phi$.*

**Robust lower bounds** The previous lemma easily applies to the state $|\text{🐱}_\pm\rangle$ for proving circuit depth lower bounds of $\Omega(\log n)$ since they are $(n-1)$-locally indistinguishable. It also easily gives a $\Omega(\log d)$ lower bound for every pure code state of a $[[n, k, d]]$ code where $k \geq 1$. This is by noticing that two orthogonal code states are $d$-locally-indistinguishable (Fact 2.1). But it has flaws that need addressing. First, it is only bounding pure states on $n$ qubits whereas we will, in general, be interested in lower bounding mixed states on $n$ qubits. Second, the bounds prove depth lower bounds for circuits *exactly* generating the state $|\rho\rangle$. This is not a robust definition[7], in general, we will be interested in proving lower bounds for all states within a small $\delta$-radius of the state $\rho$. This is a prerequisite for NLTS. Here we present a collection of robust lower bounding techniques which will be necessary for proving the NLTS theorem [1].

## 3.6 Lower bounds for well-spread distributions

The lower bound for $|\text{🐱}_+\rangle$ immediately implies a $\Omega(\log n)$ lower bound for the probability distribution with half its mass on $0^n$ and the other half on $1^n$. To make the lower bound robust, we would like to show that every distribution $D$ on $\{0,1\}^n$ with $D(0^n), D(1^n) \geq \mu$ for some constant $\mu$ has a non-trivial quantum circuit lower bound. We can even further generalize to sets $S_1$ and $S_2$ with $D(S_1), D(S_2) \geq \mu$ and a large Hamming distance between $S_1$ and $S_2$. We call such distributions "well-spread" and show a generalized lower bound for them by modifying our previous lower bound argument and using $P(\mathbf{H}_U)$ instead of $\mathbf{H}_U$ where $P(\cdot)$ is the polynomial from Lemma 3.14. Versions of the following lower bound versions for well-spread distributions can be found in [45, Theorem 4.6], [52, Corollary 43], [2, Lemma 13].

**Lemma 3.18 (Fact 4 of [1])** *Let $D$ be a probability distribution on $\{0,1\}^n$ generated by measuring the output of a quantum circuit in the standard basis. If two sets $S_1, S_2 \subset \{0,1\}^n$ satisfy $D(S_1), D(S_2) \geq \mu$, then the depth of the circuit is at least*

$$\frac{1}{3}\log\left(\frac{\operatorname{dist}(S_1, S_2)^2}{400n \cdot \ln\frac{1}{\mu}}\right). \tag{3.20}$$

[7] Robust lower bounds for the classical variant of this question, the uniform distribution over a classical code, are answered in full by Lovett and Viola [64].

The only ingredient not previously discussed in this proof is the simple observation that for an $\ell$-local operator term $h$ and basis vectors $|x\rangle, |y\rangle$ for $x, y \in \{0,1\}^n$ with Hamming distance $|x \oplus y| > \ell$, $\langle x| h |y\rangle = 0$. Therefore, for sets $S_1$ and $S_2$ any local operator $\mathbf{H}$ of locality $< \mathrm{dist}(S_1, S_2)$,

$$\Pi_{S_1}\mathbf{H}\Pi_{S_2} = 0 \tag{3.21}$$

where $\Pi_{S_1}, \Pi_{S_2}$ are the projections on the strings in sets $S_1, S_2$, respectively.

**Proof of Proof of Lemma 3.18:** Let $\rho$ be a mixed state such that measurement in the standard basis results in the distribution $D$ and assume let $|\rho\rangle$ be the output of a depth $t$ circuit $U$ such that $|\rho\rangle$ is a purification of $\rho$ — i.e. $\mathrm{tr}_{[n']\setminus[n]}(|\rho\rangle\!\langle\rho|) = \rho$. We first recognize that, without loss of generality, $n' \leq 2^t \cdot n$. This is a lightcone argument (Fact 3.2) as any qubits outside $\bigcup_{i\in[n]} L_i$, the union of the lightcones of the first $n$ qubits, do not influence the state on the first $n$ qubits. Therefore, they can be removed without changing the state on the first $n$ qubits.

Now, consider the AGSP $P(\mathbf{H}_U)$ previously constructed (Lemma 3.14) using $P$ of degree $f$:

$$\|\, |\rho\rangle\!\langle\rho| - P(\mathbf{H}_U)\| \leq \exp\left(-\frac{f^2}{100\cdot n'}\right) \leq \exp\left(-\frac{f^2}{100\cdot 2^t n}\right) \tag{3.22}$$

Furthermore, $P(\mathbf{H}_U)$ is a $f \cdot 2^t$ local operator. Setting $u \overset{\text{def}}{=} \mathrm{dist}(S_1, S_2)$ and choosing $f \overset{\text{def}}{=} \frac{u}{2^{t+1}}$, we obtain

$$\|\, |\rho\rangle\!\langle\rho| - P(\mathbf{H}_U)\| =\leq \exp\left(-\frac{u^2}{400\cdot 2^{3t} n}\right). \tag{3.23}$$

From eq. (3.21), we have $\Pi_{S_1}P(\mathbf{H}_U)\Pi_{S_2} = 0$ which implies

$$\begin{aligned}
\|\Pi_{S_1}\, |\rho\rangle\!\langle\rho|\, \Pi_{S_2}\| &= \|\Pi_{S_1}\, |\rho\rangle\!\langle\rho|\, \Pi_{S_2} - \Pi_{S_1}P(\mathbf{H}_U)\Pi_{S_2}\| && (3.24\text{a})\\
&\leq \|\, |\rho\rangle\!\langle\rho| - P(\mathbf{H}_U)\| && (3.24\text{b})\\
&\leq \exp\left(-\frac{u^2}{400\cdot 2^{3t}\cdot n}\right). && (3.24\text{c})
\end{aligned}$$

However,

$$\|\Pi_{S_1}\, |\rho\rangle\!\langle\rho|\, \Pi_{S_2}\| = \sqrt{\langle\rho|\, \Pi_{S_1}\, |\rho\rangle \cdot \langle\rho|\, \Pi_{S_2}\, |\rho\rangle} = \sqrt{p(S_1)p(S_2)} \geq \mu. \tag{3.25}$$

Thus, $2^{3t} \geq \frac{u^2}{400\cdot\ln\frac{1}{\mu}\cdot n}$, which rearranges into the fact statement. □

**Applications of Lemma 3.18** Note that Lemma 3.18 is a generalization of the lower bounds of |🐱⟩ states and, in particular, allows for mixed states. A downside of the lemma is that it only provides a non-trivial lower bound when $\mathrm{dist}(S_1, S_2) \geq \omega(\sqrt{n})$. In some cases, this is fine, since we can consider a restriction of the distribution to fewer bits to reduce $n$ while still preserving the distance. However, in some situations, this may not be optimal. This is why, in some sense, this is a robustness of the local indistinguishability lower bounds for certain states but it is different than the aforementioned local indistinguishability arguments.

For example, it is unclear how to robustly use Lemma 3.18 to lower bound the circuit depth of code states of all quantum codes; it is, however, easy enough to see how to use Lemma 3.18 to lower bound the circuit depth of the code states of CSS codes (and any state within small $\ell_1$-distance of the code space) for $d \geq \omega(\sqrt{n})$. In [52], Eldar and Harrow notice that given a $[[n, k, d]]$ stabilizer code, we can find two logical operators $\overline{X}$ and $\overline{Z}$ such that these operators are anti-commuting Pauli matrices (see Definition 2.3). Moreover, we can find Pauli matrices such that $\overline{X}$ is a $X$-type Pauli and $\overline{Z}$ is a $Z$-type Pauli. Because two anti-commuting operators that square to 1 form a qubit[8], they satisfy an uncertainty principle:

**Fact 3.19** *For any state $\rho$ and an operator $A$, let*

$$\mathbf{Var}\ A = \mathrm{tr}\left(A^2\rho\right) - \mathrm{tr}(A\rho)^2. \tag{3.26}$$

*For any anti-commuting operators $A, B$ that square to 1,* $\mathbf{Var}\ A + \mathbf{Var}\ B \geq 1$ *and* $\mathbf{Var}\ A, \mathbf{Var}\ B \geq 0$.

When applied to $\overline{X}$ and $\overline{Z}$ we get that at least one $\mathbf{Var}\ \overline{X}$ or $\mathbf{Var}\ \overline{Z}$ is $\geq \frac{1}{2}$. Without loss of generality, assume it to be $\overline{Z}$. Then,

$$\frac{1}{2} \leq \mathbf{Var}\ \overline{Z} = 1 - \mathrm{tr}\left(\overline{Z}\rho\right)^2 \implies \frac{1}{\sqrt{2}} \geq \mathrm{tr}\left(\overline{Z}\rho\right). \tag{3.27}$$

Therefore, a $\overline{Z}$-measurement of $\rho$ would yield either option with a probability $\geq \frac{1}{2} - \frac{1}{\sqrt{2}}$. Since $\overline{Z}$ is a $Z$-type measurement, $\overline{Z} = Z^b$ for $b \in \{0,1\}^n$. It is a *sub-measurement* of measuring all qubits in the $Z$-basis (standard basis). Recall from the description of CSS codes, for a code state measured all qubits in the $Z$-basis, the resulting distribution will be supported on $\ker H_{\mathsf{Z}}$. We can divide $\ker H_{\mathsf{Z}}$ into two components:

$$S_1 \overset{\mathsf{def}}{=} \ker H_{\mathsf{Z}} \cap \{y : y^\top b = 0\}, \qquad S_2 \overset{\mathsf{def}}{=} \ker H_{\mathsf{Z}} \cap \{y : y^\top b = 1\}. \tag{3.28}$$

$S_1$ corresponds to measuring $|Z\rangle$ and getting the outcome +1 and $S_2$ the outcome −1. Since $\overline{Z}$ is a logical operator, the distance between $S_1$ and $S_2$ is at least $d$. This is because the $|\cdot|_{C_{\mathsf{X}}^\perp}$ distance between any two points in $\ker H_{\mathsf{Z}}$ is at least $d$ and $|\cdot|_{C_{\mathsf{X}}^\perp} \leq |\cdot|$. Therefore, we have a well-spread distribution for $\mu = \frac{1}{2} - \frac{1}{\sqrt{2}}$ and $\mathrm{dist}(S_1, S_2) \geq d$. This gives a non-trivial lower bound for any circuit of super-quadratic distance; the bound can easily be seen to be robust to small $\ell_1$-perturbations (as seen in [52, Proposition 44]).

[8] They define a two-dimensional subspace; see [65] for an introduction to this intuition.

**A local indistinguishablity perspective** Although not necessary for the remainder of the thesis, it may be illustrative to the reader to see a proof sketch of Lemma 3.18 through the lens of local indistinguishability; this is inspired by [52]. Recall, we are interested in lower bounding the circuit depth of any state $|\rho\rangle$ on $n'$ qubits such that measuring $|\rho\rangle$ in the standard basis yields the distribution $D$. To apply Lemma 3.16, we would need to create a state $|\varphi\rangle$ with the same reduced density matrices. Instead, we create a state with approximately the same reduced density matrices.
**Proof** *(Sketch)* **:** Let $u = \mathrm{dist}(S_1, S_2)$ and let $B_i$ be the Hamming ball of radius $i \cdot \ell$ around $S_1$ for $i = 1, \ldots, u/\ell$ (see Figure 3.3). Consider the sequence of states

$$|\varphi_i\rangle = \left( \sum_{x \in \{0,1\}^n} (-1)^{x \in B_i} |x\rangle\langle x| \right) |\rho\rangle \,. \tag{3.29}$$

In other words, $|\varphi_i\rangle$ flips the sign of the basis strings inside the ball $B_i$. Notice that $|\langle\rho|\varphi_i\rangle|^2 \leq 1 - O(\mu)$ since $S_1 \subset B_i$. Consider now the AGSP $P(\mathbf{H}_U)$ previously constructed (Lemma 3.14) where $U$ is a depth $t$ circuit generating $|\rho\rangle$ and $P$ is of degree $f \sim \sqrt{n}$. Then, $\mathrm{tr}(P(\mathbf{H}_U)\varphi_i) \geq \Omega(\mu)$ by approximate orthogonality.

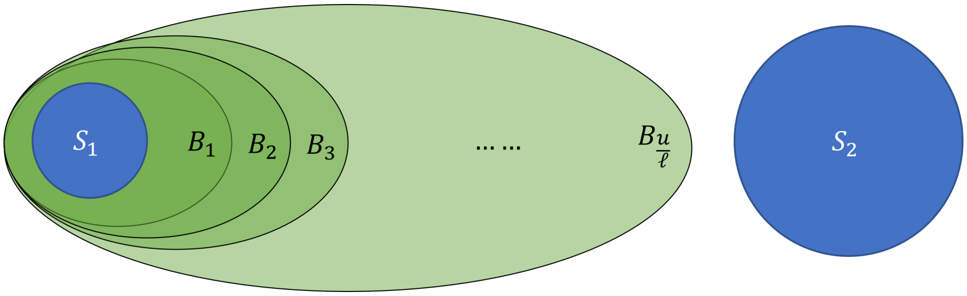

Figure 3.3: A cartoon of the local indistinguishability lower bound for well-spread distributions.

On the other hand, let $\ell \sim 2^t\sqrt{n}$ be the locality of $P(\mathbf{H}_U)$. Then, it suffices to note that

$$|\langle\rho|P(\mathbf{H}_U)|\rho\rangle - \langle\varphi_i|P(\mathbf{H}_U)|\varphi_i\rangle| \leq O(D(B_{i+1} \setminus B_{i-1})). \tag{3.30}$$

This is because $\langle x|\, P(\mathbf{H}_U)\, |y\rangle = 0$ unless $|x \oplus y| \leq \ell$ and the only locations at which the phases differ for the states $\rho$ and $\varphi_i$ are near the perimeter of ball $B_i$. This is precisely the difference in balls $B_{i+1}$ and $B_{i-1}$, and by a Cauchy-Schwartz inequality, the difference is upper-bounded by the probability mass in the region. By a counting argument, there must exist some $i^\star$ such that $O(D(B_{i^\star+1} \setminus B_{i^\star-1})) \leq O(\ell/u)$. Therefore, $\mu \leq O(\ell/u) = O(2^t\sqrt{n}/u)$. Therefore $|\rho\rangle$ and $|\varphi_{i^\star}\rangle$ are approximately locally indistinguishable. Rewriting the equation, gives us a bound of $t \geq \Omega\left(\log\left(\frac{u\mu}{\sqrt{n}}\right)\right)$ which is similar to that of Lemma 3.18. □

## 3.7 More lower bounds from code definitions

The following section is not necessary for understanding the proof of the NLTS theorem; however, it addresses generalizing circuit depth lower bounds to more quantum error correcting codes. This includes codes that are not simultaneously linear-rate and linear-distance or codes that are not necessarily CSS (or even stabilizer). In [2], we studied this problem in depth and extended some of the folklore local indistinguishability arguments in interesting manners for these results. These lower bounds were insufficient for proving NLTS as we were, at the time, unable to incorporate the utility of simultaneously being linear-rate and linear-distance. However, we proved simple arguments for codes that were *either* high-rate or high-distance codes. The results of the two sections can be expressed succinctly by the following theorem.

**Theorem 3.20 (Theorem 1 of [2])** *Let $C$ be a $[[n,k,d]]$ stabilizer code of constant locality $\ell = O(1)$ and let $\mathbf{H} = \sum_i h_i$ be the corresponding Hamiltonian (see eq. (2.12)) for each code check $C_i$. For any $\epsilon > 0$ and any state $\psi$ on n-qubits with energy $\leq \epsilon n$, the circuit depth of $\psi$ is at least*

$$\operatorname{depth}(\psi) \geq \Omega\left(\min\left\{\log d, \quad \log \frac{k+d}{n\sqrt{\epsilon \log \frac{1}{\epsilon}}}\right\}\right). \tag{3.31}$$

The proof of Theorem 3.20 is provided in Appendix A.1.

### 3.7.1 High-distance codes

Furthermore, we can prove non-trivial lower bounds as to the fidelity of low-depth states and any code state. We prove lower bounds in both the high-rate and high-distance regimes. The high-rate lower bounds are presented in the appendix as Lemma A.3 and the high-distance lower bounds are the following lemma.

**Lemma 3.21 (Lemma 14 of [2])** *Given $|\psi\rangle = U\left|0^{\otimes n'}\right\rangle$ with $U$ of depth $t$ and let $C$ be a code with distance $d$. Let $\Pi$ be the projector onto the codespace and $F = \sqrt{\langle\psi|\left(\Pi \otimes \mathbb{I}_{\text{ancilla}}\right)|\psi\rangle}$ be the fidelity of $|\psi\rangle$ with the codespace. If $2^t \leq \frac{d}{2}$ then*

$$F^2 \leq 2\exp\left(-\frac{d^2}{400 \cdot 2^{2t} n'}\right). \tag{3.32}$$

**Proof:** From Lemma 3.14, there is a polynomial $P$ of degree $\frac{d}{2^{t+1}}$ such that

$$\|P(\mathbf{H}_U) - |\psi\rangle\langle\psi|\,\| \leq \exp\left(-\frac{d^2}{400 \cdot 2^t n'}\right) \tag{3.33}$$

Note that each multinomial term in $P(\mathbf{H}_U)$ is supported on $\leq 2^t \cdot \frac{d}{2^{t+1}} \leq \frac{d}{2}$ terms. Let $|\phi\rangle = \Pi\, |\psi\rangle\, / f$ be the code state having largest overlap with $|\psi\rangle$. Consider[9] any vector $|\phi'\rangle$ with $\phi'_{\text{code}} \in C$ that is orthogonal to $|\phi\rangle$ (and hence orthogonal to $|\psi\rangle$). Since $P(\mathbf{H}_U)$ is a $\leq \frac{d}{2}$-local Hamiltonian, then Fact 3.17, ensures that $P(\mathbf{H}_U)$ cannot distinguish $|\phi\rangle$ and $|\phi'\rangle$. Therefore,

$$\begin{aligned}
F^2 &= |\langle\phi|\psi\rangle|^2 && (3.35a)\\
&\leq \langle\phi|P(\mathbf{H}_U)|\phi\rangle + \|P(\mathbf{H}_U) - |\psi\rangle\!\langle\psi|\| && (3.35b)\\
&\leq \langle\phi'|P(\mathbf{H}_U)|\phi'\rangle + \|P(\mathbf{H}_U) - |\psi\rangle\!\langle\psi|\| && (3.35c)\\
&\leq \Big(|\langle\phi'|\psi\rangle|^2 + \|P(\mathbf{H}_U) - |\psi\rangle\!\langle\psi|\|\Big) + \|P(\mathbf{H}_U) - |\psi\rangle\!\langle\psi|\| && (3.35d)\\
&\leq 2\|P(\mathbf{H}_U) - |\psi\rangle\!\langle\psi|\| && (3.35e)
\end{aligned}$$

where eq. (3.35b) and eq. (3.35d) follow from triangle inequalities, eq. (3.35c) is the indistinguishability by local Hamiltonians, and eq. (3.35e) is because $\langle\phi'|\psi\rangle = 0$. With eq. (3.33), the proof is complete.

□

We can extend this proof to prove circuit depth lower bounds for low-energy states of generic codes due to their distance.

**Lemma 3.22** *Let $C$ be a $[[n, k, d]]$ stabilizer code of locality $\ell$ defined by checks $\{C_i\}_{i\in[m]}$. Let $\mathbf{H}$ be the corresponding Hamiltonian. Suppose there is a state $|\phi\rangle$ on $n'$ qubits with $\operatorname{tr}(\mathbf{H}\psi) \leq \epsilon m$ and circuit complexity $t \overset{\text{def}}{=} \operatorname{depth}(\phi) < \log(d) - 1$. Then,*

$$t \geq \frac{1}{3}\log\left(-\frac{d^2}{1200\sqrt{\epsilon}\cdot \ell n^2}\right). \tag{3.36}$$

A proof is provided in Appendix A.1.

### 3.7.2 High-rate codes

In [2], we observed, for the first time, that another parameter plays a key role in circuit lower bounds: the rate of the code. Inspired by [66], we used novel entropic arguments to prove that states of low circuit complexity are significantly far in $\ell_1$−distance from high rate code spaces. Formally, we showed that all states of circuit complexity $\leq \log d$ are at a $\ell_1$-distance of $\geq \Omega\big(k^2/n^2\big)$ from the code space.

[9] One way to construct $|\phi'\rangle$ is to expand $|\phi\rangle = \sum_x |\bar{x}\rangle\,|\phi_x\rangle$ (with $\{|\bar{x}\rangle\}_x$ a basis for $C$ and $|\phi_x\rangle$ unnormalized) and then define $|\phi'\rangle \propto \sum_x \alpha_x\, |\bar{x}\rangle\, |\phi_x\rangle$. The complex numbers $\{\alpha_x\}_x$ are chosen such that

$$\sum_x \alpha_x\,\langle\phi_x|\phi_x\rangle = 0 \implies \langle\phi|\phi'\rangle = 0. \tag{3.34}$$

**Lemma 3.23 ( [2])** *Let $C$ be a $[[n, k, d]]$ code and $\psi$ a state on $m$ qubits. Let $\psi_{\text{code}}$ be the reduced state on the $n$ code qubits. If the trace-distance between $\psi_{\text{code}}$ and $C$ is $0 < \delta < 1/2$ and the code is of rate at least $k > 2\delta \log(1/\delta) m$, then the circuit complexity* $\operatorname{depth}(\psi) > \log d$.

This observation alone does not suffice to address a central challenge of bounding circuit depth: the space of low-energy states is much larger than the code space or even its small neighborhood. A general strategy in earlier works [44, 52] was to build a low-depth decoding circuit to bring each low-energy state closer to the code space. But this required assuming that the code was locally testable; such codes are not known to exist in the desired parameter regime. We instead appeal to the observation that every eigenspace of a stabilizer code Hamiltonian possesses the local indistinguishability property (Fact 2.1). Instead of attempting to construct a decoding circuit, we measure the syndrome using a constant-depth circuit (which uses the LDPC nature of the code Hamiltonian). This allows us to decohere the low energy state into a mixture of orthogonal states that live within each of the eigenspaces. A key realization is that measurement of the syndrome for low-energy states is a gentle measurement in that it does not perturb the state locally. This is used to show that a state of low energy satisfies an approximate version of local indistinguishability. This, coupled with the argument for codes of high rate, completed a proof for high-rate codes.

**Theorem 3.24 (Theorem 20 of [2])** *Let $C$ be a $[[n, k, d]]$ stabilizer code of locality $\ell$ defined by checks $\{C_i\}_{i\in[m]}$. Let $\mathbf{H}$ be the corresponding Hamiltonian. Suppose there is a state $|\phi\rangle$ on $m$ qubits with $\operatorname{tr}(H\phi) \le \epsilon m$ and circuit complexity $t \overset{\text{def}}{=} \operatorname{depth}(\phi) < \log(d) - 2\ell^3$. Then, for a constant $c_\ell$ depending only on $\ell$ and not the size of the code,*

$$2^{2t} > \frac{k}{c_\ell n \cdot \epsilon \log \frac{1}{\epsilon}}. \tag{3.37}$$

A proof of Lemma 3.23 and a proof of Theorem 3.24 are given in Appendix A.1.

### 3.7.3 Spatially local Hamiltonians

A key property of an NLTS Hamiltonian is that it cannot live on a *Euclidean* lattice of dimension $D$ for a fixed constant $D$ [21]. This is because of a "cutting" argument: Let $\mathbf{H}$ be a local Hamiltonian in $D$ dimensions and $\Psi$ a ground state of $\mathbf{H}$. For a fixed constant $\epsilon$, partition the lattice into $D$ dimensional rectangular chunks so that the side length of each rectangular chunk is $O((D\epsilon)^{-1/D})$. Let $\rho_i$ be the reduced state of $\Psi$ on a chunk $i$, and $\rho = \bigotimes_i \rho_i$ be a state over all the qubits. It's not hard to check that $\rho$ violates at most a $\epsilon$-fraction of the terms of $\mathbf{H}$ (only the boundary terms of the rectangular division) and yet has circuit complexity at most $\exp(((D\epsilon)^{-1/D})^D) = O(\exp(1/D\epsilon)) = O(1)$; so it is not NLTS.

This circuit complexity upper bound can be further improved for the specific case of stabilizer Hamiltonians on a lattice, due to the result of Aaronson and Gottesman [67]. Since the circuit complexity of each chunk is at most logarithmic in its size $O(1/\epsilon^{1/D})$, the aforementioned quantum state $\rho$ can be prepared by a circuit of depth $O(\min(\log n, \log(1/\epsilon)))$. Note that this holds for any

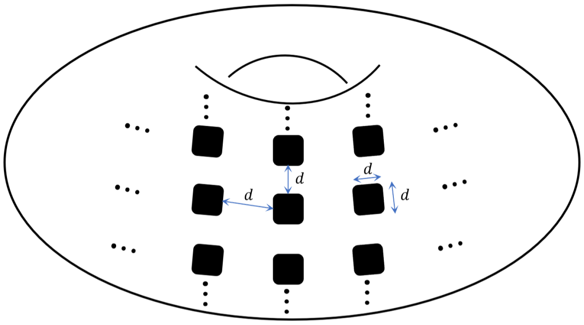

Figure 3.4: The punctured toric code is a quantum surface code on a $\sqrt{n} \times \sqrt{n}$ torus with punctures of size $d \times d$ and distance between punctures of $d$. These codes saturate the rate-distance tradeoff bounds in two dimensions [66] with distance = $O(d)$ and rate = $\Omega(n/d^2)$.

$0 < \epsilon < 1$, not just a constant. Theorem 3.20 shows that the 2D punctured toric code (Figure 3.4) Hamiltonians on $n$ qubits with distance $d$ (which is a free parameter) requires a circuit of depth $\Omega(\log d)$ for an approximation to ground energy better than $O(n/d^3)$. Therefore, our lower bound in the case of nearly linear rate and polynomial distance codes (such as the punctured toric code) matches the upper bound – up to constant factors – closing the question on the circuit complexity of the approximate ground states of these codes.

# Chapter 4

# The NLTS Theorem

> Each week I plot your equations dot for dot, $x$'s against $y$'s in all manner of algebraical relation, and every week they draw themselves as commonplace geometry, as if the world of forms were nothing but arcs and angles. God's truth, Septimus, if there is an equation for a curve like a bell, there must be an equation for one like a bluebell, and if a bluebell, why not a rose? Do we believe nature is written in numbers?
>
> Tom Stoppard, *Arcadia*[0]

In this chapter, we give an argument that likely nature is *not* written in numbers (even approximately) and that describing ground-states (solutions) or low-energy states (approximate solutions) of local Hamiltonians likely requires quantum mechanics to describe.

**Theorem 4.1 (No low-energy trivial states [1])** *There exists a fixed constant $\epsilon > 0$ and an explicit family of $O(1)$-local frustration-free commuting Hamiltonians $\{\mathbf{H}^{(n)}\}_{n=1}^{\infty}$ where $\mathbf{H}^{(n)} = \sum_{i=1}^{m} h_i^{(n)}$ acts on $n$ particles and consists of $m = \Theta(n)$ local terms such that for any family of states $\{\psi_n\}$ satisfying $\mathrm{tr}\left(\mathbf{H}^{(n)}\psi\right) < \epsilon n$, the circuit complexity of the state $\psi_n$ is at least $\Omega(\log n)$.*

The local Hamiltonians for which we can show such robust circuit depth lower bounds correspond to linear-rate and linear-distance quantum LDPC CSS error correcting codes with an additional property related to the clustering of approximate codewords of the underlying classical codes. We show that the property holds for the *quantum Tanner code* construction of Leverrier and Zémor [43] (Section 4.2). We suspect that the property is true for other constructions of linear-rate and linear-distance QLDPC codes [40, 42, 68], however we do not prove this outright. While we show that the property is sufficient for NLTS, it is an interesting open question if the property is inherently satisfied by all linear-rate and linear-distance constructions.

[0] Tom Stoppard's quote is about *chaos theory* which was a subject he was quite fascinated by as a writer. His plays often reflect his curiosity about mathematical notions such as the butterfly effect and I wonder if he has had any introduction to quantum mechanics and what literary conclusions he has drawn from this subject.

For any subset $S \subset \{0,1\}^n$, recall the distance measure $|\cdot|_S$ as $|y|_S = \min_{s \in S} |y+s|$ where $|\cdot|$ denoted Hamming weight. For a $[[n,k,d]]$ CSS code defined by classical codes $(C_{\mathsf{X}}, C_{\mathsf{Z}})$, we define $G_{\mathsf{Z}}^{\delta}$ as the set of vectors which violate at most a $\delta$-fraction of checks from $C_{\mathsf{Z}}$, i.e. $G_{\mathsf{Z}}^{\delta} = \{y : |H_{\mathsf{Z}} y| \le \delta m_{\mathsf{Z}}\}$. We similarly define $G_{\mathsf{X}}^{\delta}$.

**Property 4.2 (Clustering of approximate codewords)** *We say that a $[[n,k,d]]$ CSS code defined by classical codes $(C_{\mathsf{X}}, C_{\mathsf{Z}})$* clusters approximate codewords *if there exist constants $c_1, c_2, \delta_0$ such that for sufficiently small $0 \le \delta < \delta_0$ and every vector $y \in \{0,1\}^n$,*

1. *If $y \in G_{\mathsf{Z}}^{\delta}$, then either $|y|_{C_{\mathsf{X}}^{\perp}} \le c_1 \delta n$ or else $|y|_{C_{\mathsf{X}}^{\perp}} \ge c_2 n$.*
2. *If $y \in G_{\mathsf{X}}^{\delta}$, then either $|y|_{C_{\mathsf{Z}}^{\perp}} \le c_1 \delta n$ or else $|y|_{C_{\mathsf{Z}}^{\perp}} \ge c_2 n$.*

This property can be viewed as a quantum — or, in the language of [68, 69], boundary and co-boundary — clustering of approximate codewords. Recall the discussion that the classical clustering of approximate codewords for classical codes occurs for codes with small-set expanding interaction graphs. In particular, this includes classical Tanner codes.

**Local Hamiltonian definition** The aforementioned quantum codes lead to a natural commuting frustration-free local Hamiltonian (as defined in eq. (2.12)): For every row $w_{\mathsf{Z}}$ of $H_{\mathsf{Z}}$ – i.e. a stabilizer term $Z^{w_{\mathsf{Z}}}$ of the code, we associate a Hamiltonian term $\frac{1}{2}(\mathbb{I} - Z^{w_{\mathsf{Z}}})$. We define $\mathbf{H}_{\mathsf{Z}}$ as the sum of all such terms for $H_{\mathsf{Z}}$. $\mathbf{H}_{\mathsf{X}}$ is defined analogously and the full Hamiltonian is $\mathbf{H} = \mathbf{H}_{\mathsf{X}} + \mathbf{H}_{\mathsf{Z}}$. The number of local terms is $m_{\mathsf{X}} + m_{\mathsf{Z}} = \Theta(n)$ and $\mathbf{H}$ has zero ground energy.

## 4.1 A short proof

The proof, that the local Hamiltonian corresponding to a linear-rate and linear-distance CSS code satisfying Property 4.2 is NLTS, is divided into a few steps. We first show that the classical distributions generated by measuring any low-energy state in the standard or Hadamard bases are approximately supported on a particular structured subset of vectors. Then, we show that the subsets cluster into a collection of disjoint components which are far in Hamming distance from each other. Finally, we show that the distribution in one of the two bases cannot be too concentrated on any particular cluster. This shows that the distribution is *well-spread* (like Lemma 3.18) which can be used to prove a circuit depth lower bound.

**The supports of the underlying classical distributions** Consider a state $\psi$ on $n$ qubits such that $\mathrm{tr}(\mathbf{H}\psi) \le \epsilon n$. Let $D_{\mathsf{X}}$ and $D_{\mathsf{Z}}$ be the distributions generated by measuring the $\psi$ in the (Hadamard) $X-$ and (standard) $Z-$ bases, respectively. We find that $D_{\mathsf{Z}}$ is largely supported on $G_{\mathsf{Z}}^{O(\epsilon)}$. Formally, this is because, by construction,

$$\epsilon n \ge \mathrm{tr}(\mathbf{H}\psi) \ge \mathrm{tr}(\mathbf{H}_{\mathsf{Z}}\psi) = \mathop{\mathbf{E}}_{y \sim D_{\mathsf{Z}}} |H_{\mathsf{Z}} y|. \tag{4.1}$$

Here, the last equality holds since for a Pauli operator $Z^a$, $\langle y| \frac{\mathbb{I}-Z^a}{2} |y\rangle = \frac{1-(-1)^{a.y}}{2} = a.y$. Let $q \overset{\text{def}}{=} D_{\mathsf{Z}}(G_{\mathsf{Z}}^{\epsilon_1})$ be the probability mass assigned by $D_{\mathsf{Z}}$ to $G_{\mathsf{Z}}^{\epsilon_1}$. Then,

$$\mathop{\mathbf{E}}_{y \sim D_{\mathsf{Z}}} |H_{\mathsf{Z}} y| \geq 0 \cdot q + (1-q) \cdot \epsilon_1 m_{\mathsf{Z}} = (1-q)\epsilon_1 m_{\mathsf{Z}}. \tag{4.2}$$

Therefore, $D_{\mathsf{Z}}(G_{\mathsf{Z}}^{\epsilon_1}) \geq 1 - \epsilon n/(\epsilon_1 m_{\mathsf{Z}})$. A similar argument shows that $D_{\mathsf{X}}(G_{\mathsf{X}}^{\epsilon_1}) \geq 1 - \epsilon n/(\epsilon_1 m_{\mathsf{X}})$. With the choice $\epsilon_1 = \frac{200n}{\min\{m_{\mathsf{X}}, m_{\mathsf{Z}}\}} \cdot \epsilon$, we find

$$D_{\mathsf{Z}}(G_{\mathsf{Z}}^{\epsilon_1}), D_{\mathsf{X}}(G_{\mathsf{X}}^{\epsilon_1}) \geq \frac{199}{200} \tag{4.3}$$

for both the bases.

**The supports are well clustered** Given that $D_{\mathsf{Z}}$ is well supported on $G_{\mathsf{Z}}^{\epsilon_1}$, it is helpful to understand the structure of $G_{\mathsf{Z}}^{\epsilon_1}$. For $x, y \in G_{\mathsf{Z}}^{\epsilon_1}$, notice that $x \oplus y \in G_{\mathsf{Z}}^{2\epsilon_1}$ since $x \oplus y$ satisfies every check that both $x$ and $y$ satisfy. By Property 4.2 (and assuming $2\epsilon_1 \leq \delta_0$), then either

$$|x \oplus y|_{C_{\mathsf{X}}^{\perp}} \leq 2c_1\epsilon_1 n \quad \text{or else} \quad |x \oplus y|_{C_{\mathsf{X}}^{\perp}} \geq c_2 n. \tag{4.4}$$

Define a relation '$\sim$' such that for $x, y \in G_{\mathsf{Z}}^{\epsilon_1}$, $x \sim y$ iff $|x \oplus y|_{C_{\mathsf{X}}^{\perp}} \leq 2c_1\epsilon_1 n$. To prove that the relation is transitive and therefore an equivalence relation, notice that if $x \sim y$ and $y \sim z$, then

$$|x \oplus z|_{C_{\mathsf{X}}^{\perp}} \leq |x \oplus y|_{C_{\mathsf{X}}^{\perp}} + |y \oplus z|_{C_{\mathsf{X}}^{\perp}} \leq 4c_1\epsilon_1 n. \tag{4.5}$$

However, $x \oplus z \in G_{\mathsf{Z}}^{2\epsilon_1}$ and for sufficiently small $\epsilon_1$ such that $4c_1\epsilon_1 < c_2$, Property 4.2 implies that $|x \oplus z|_{C_{\mathsf{X}}^{\perp}} \leq 2c_1\epsilon_1 n$. Thus, $x \sim z$ and hence $\sim$ forms an equivalence relation. We can now divide the set $G_{\mathsf{Z}}^{\epsilon_1}$ into clusters $B_{\mathsf{Z}}^1, B_{\mathsf{Z}}^2, \ldots$, according to the equivalence relation $\sim$. Furthermore, the distance between any two clusters is $\geq c_2 n$, since for $x$ in one cluster and $x'$ in another cluster, we have $|x \oplus x'| \geq |x \oplus x'|_{C_{\mathsf{X}}^{\perp}} \geq c_2 n$. Therefore, the picture for both bases looks like Figure 2.1 with norm $|\cdot|_{C_{\mathsf{X}}^{\perp}}$ as the clusters contain most of the support. Lastly, the same argument holds for $G_{\mathsf{X}}^{\epsilon_1}$.

**The distributions are not concentrated on any one cluster** To apply known circuit depth lower bounding techniques to $D_{\mathsf{Z}}$, it suffices to show that $D_{\mathsf{Z}}$ is not concentrated on any one cluster $B_{\mathsf{Z}}^i$. However, it is not immediate how to show this property for $D_{\mathsf{Z}}$. Instead, what we can show is that is impossible for both $D_{\mathsf{Z}}$ to be concentrated on any one cluster $B_{\mathsf{Z}}^i$ and $D_{\mathsf{X}}$ to be concentrated on any one cluster $B_{\mathsf{X}}^j$.

**Lemma 4.3** *For $\epsilon_1$ such that $2c_1\epsilon_1 \leq \left(\frac{k-1}{4n}\right)^2$, either $\forall\ i$, $D_{\mathsf{Z}}(B_{\mathsf{Z}}^i) < 99/100$ or else $\forall\ j$, $D_{\mathsf{X}}(B_{\mathsf{X}}^j) <$* 99/100.

**Proof:** Assume there exists some $i$ such that $D_{\mathsf{Z}}(B_{\mathsf{Z}}^i) \geq 99/100$. We will employ the following fact that captures the well-known uncertainty of measurements in the standard and Hadamard bases; the proof of the fact is given immediately after this proof.

**Fact 4.4** *Given a state $\psi$ and corresponding measurement distributions $D_{\mathsf{X}}$ and $D_{\mathsf{Z}}$, for all subsets $S, T \subset \{0,1\}^n$, $D_{\mathsf{X}}(T) \le 2\sqrt{1 - D_{\mathsf{Z}}(S)} + \sqrt{|S| \cdot |T|/2^n}$.*

For any $j$, we employ this fact with $S = B^i_{\mathsf{Z}}$ and $T = B^j_{\mathsf{X}}$. To bound $|B^i_{\mathsf{Z}}|$, fix any string $z \in B^i_{\mathsf{Z}}$. Any other string $z' \in B^i_{\mathsf{Z}}$ has the property that its Hamming distance from $z \oplus w$ (for some $w \in C^{\perp}_{\mathsf{X}}$) is at most $2c_1\epsilon_1 n$. Since $\left|C^{\perp}_{\mathsf{X}}\right| = 2^{\dim C^{\perp}_{\mathsf{X}}} = 2^{n - \dim C_{\mathsf{X}}} = 2^{r_{\mathsf{X}}}$, the size of the cluster $B^i_{\mathsf{Z}}$ is at most

$$2^{r_{\mathsf{X}}} \cdot \binom{n}{2c_1\epsilon_1 n} \le 2^{r_{\mathsf{X}}} \cdot 2^{2\sqrt{2c_1\epsilon_1}n}. \tag{4.6}$$

A similar bound can be calculated of $\left|B^j_{\mathsf{X}}\right| \le 2^{r_{\mathsf{Z}}} \cdot 2^{2\sqrt{2c_1\epsilon_1}n}$. Then applying Fact 4.4 with the bound on $\epsilon_1$ as stated in the Lemma,

$$\forall\, j, \quad D_{\mathsf{X}}\left(B^j_{\mathsf{X}}\right) \le \frac{1}{5} + \sqrt{2^{r_{\mathsf{X}} + r_{\mathsf{Z}} - n} \cdot 2^{4\sqrt{2c_1\epsilon_1}n}} = \frac{1}{5} + 2^{\frac{-k}{2} + 2\sqrt{2c_1\epsilon_1}n} < \frac{99}{100}. \tag{4.7}$$

□

**A lower bound using the *well-spread* nature of the distribution** Assume, without loss of generality, from Lemma 4.3 that $D_{\mathsf{Z}}$ is not too concentrated on any cluster $B^i_{\mathsf{Z}}$. Recall that $D_{\mathsf{Z}}(\bigcup_i B^i_{\mathsf{Z}}) \ge 199/200$. Therefore, there exist disjoint sets $M$ and $M'$ such that

$$D_{\mathsf{Z}}\left(\bigcup_{i \in M} B^i_{\mathsf{Z}}\right) \ge \frac{1}{400} \quad \text{and} \quad D_{\mathsf{Z}}\left(\bigcup_{i \in M'} B^i_{\mathsf{Z}}\right) \ge \frac{1}{400}. \tag{4.8}$$

This is because we can build the set $M$ greedily by adding terms until the mass exceeds $1/400$. Upon adding the final term to overcome the threshold, the total mass is at most $397/400$ since no term is larger than $99/100$. Therefore, the remainder of terms not included in $M$ must have a mass of at least $199/200 - 397/400 = 1/400$.

Furthermore, recall that since the distance between any two clusters is at least $c_2 n$, the same distance lower bound holds for the union of clusters over $M$ and $M'$ as well. This proves that the distribution $D_{\mathsf{Z}}$ is *well-spread* which implies a circuit depth lower bound due to Lemma 3.18: An immediate application of this lemma gives a circuit depth lower bound of $\Omega(\log n)$ for $D_{\mathsf{Z}}$ since $\mathrm{dist}(S_1, S_2) \ge c_2 n$ and $\mu = \frac{1}{400}$. Since the circuit depth of $D_{\mathsf{Z}}$ is at most one more than the circuit depth of $\psi$, the lower bound is proven.

**Theorem 4.5 (Formal statement of the NLTS theorem)** *Consider a $[[n, k, d]]$ CSS code satisfying Property 4.2 with parameters $\delta_0, c_1, c_2$ as stated. Let $\mathbf{H}$ be the corresponding local Hamiltonian. Then for*

$$\epsilon < \frac{1}{400c_1}\left(\frac{\min\{m_{\mathsf{X}}, m_{\mathsf{Z}}\}}{n}\right) \cdot \min\left\{\left(\frac{k-1}{4n}\right)^2, \delta_0, \frac{c_2}{2}\right\}, \tag{4.9}$$

*and every state $\psi$ such that* $\mathrm{tr}(\mathbf{H}\psi) \le \epsilon n$, *the circuit depth of $\psi$ is at least* $\Omega(\log n)$. *For linear-rate and linear-distance codes satisfying*[1] *Property 4.2, the bound on $\epsilon$ is a constant.*

**Proof of Fact 4.4:** Consider a purification of the state $\psi$ as $|\psi\rangle$ on a potentially larger Hilbert space. Write $|\psi\rangle$ as $\sum_{z\in\{0,1\}^n} |\psi_z\rangle \otimes |z\rangle$ where the second register is the original $n$ qubit code space. Define $C \overset{\mathsf{def}}{=} \sum_{z\in S} \| \, |\psi_z\rangle \, \|^2$ and

$$|\psi'\rangle = \frac{1}{\sqrt{C}} \sum_{z\in S} |\psi_z\rangle \otimes |z\rangle \overset{\mathsf{def}}{=} \sum_{z\in S} \left|\psi'_z\right\rangle \otimes |z\rangle \,. \tag{4.10}$$

Since $C = D_{\mathsf{Z}}(S) \overset{\mathsf{def}}{=} 1-\eta$, by the gentle measurement lemma [70] we have $\frac{1}{2}\| \, |\psi\rangle\langle\psi| - |\psi'\rangle\langle\psi'| \, \|_1 \le 2\sqrt{\eta}$. Measuring $|\psi'\rangle$ in the computational basis, we obtain a string $z \in S$ with probability $\| \left|\psi'_z\right\rangle \|^2$. Measuring $|\psi'\rangle$ in the Hadamard basis, we obtain a string $x$ with probability

$$p(x) \overset{\mathsf{def}}{=} \frac{1}{2^n} \left\| \sum_z (-1)^{x.z} \left|\psi'_z\right\rangle \right\|^2 = \frac{1}{2^n} \left( \sum_{z,w} (-1)^{x.(z\oplus w)} \left\langle \psi'_w \middle| \psi'_z \right\rangle \right). \tag{4.11}$$

Then we can compute the collision probability of $p(x)$:

$$\sum_x p(x)^2 = \frac{1}{2^{2n}} \sum_x \left( \sum_{z,w} (-1)^{x.(z\oplus w)} \left\langle \psi'_w \middle| \psi'_z \right\rangle \right)^2 \tag{4.12a}$$

$$= \frac{1}{2^{2n}} \left( \sum_x \sum_{s,t,z,w} (-1)^{x.(z\oplus w\oplus s\oplus t)} \left\langle \psi'_s \middle| \psi'_t \right\rangle \left\langle \psi'_w \middle| \psi'_z \right\rangle \right) \tag{4.12b}$$

$$= \frac{1}{2^n} \left( \sum_{s,t,z,w : z\oplus w\oplus s\oplus t=0} \left\langle \psi'_s \middle| \psi'_t \right\rangle \left\langle \psi'_w \middle| \psi'_z \right\rangle \right) \tag{4.12c}$$

$$= \frac{1}{2^n} \left( \sum_{s,t,w} \left\langle \psi'_s \middle| \psi'_t \right\rangle \left\langle \psi'_w \middle| \psi'_{s\oplus t\oplus w} \right\rangle \right) \tag{4.12d}$$

$$\le \frac{1}{2^n} \left( \sum_{s,t} \| |\psi'_s\rangle \| \, \| |\psi'_t\rangle \| \cdot \left( \sum_w \| |\psi'_w\rangle \| \, \| |\psi'_{s\oplus t\oplus w}\rangle \| \right) \right) \tag{4.12e}$$

$$\le \frac{1}{2^n} \left( \sum_{s,t} \| |\psi'_s\rangle \| \, \| |\psi'_t\rangle \| \cdot \left( \sqrt{\sum_w \| |\psi'_w\rangle \|^2} \sqrt{\sum_w \| |\psi'_{s\oplus t\oplus w}\rangle \|^2} \right) \right) \tag{4.12f}$$

$$= \frac{1}{2^n} \left( \sum_{s,t} \| |\psi'_s\rangle \| \, \| |\psi'_t\rangle \| \right) = \frac{1}{2^n} \left( \sum_{s\in S} \| |\psi'_s\rangle \| \right)^2 \tag{4.12g}$$

[1] While the distance parameter $d$ does not appear in the bound on $\epsilon$, Property 4.2 for $\delta = 0$ implies constant distance.

$$\leq \frac{1}{2^n} \cdot |S| \cdot \left( \sum_s \||\psi'_s\rangle\|^2 \right) = \frac{|S|}{2^n}. \tag{4.12h}$$

The previous line follows by an application of the Cauchy-Schwarz inequality. Apply it again to calculate that

$$\sum_{x \in T} p(x) \leq \sqrt{|T| \sum_x p(x)^2} \leq \sqrt{\frac{|S| \cdot |T|}{2^n}}. \tag{4.13}$$

Since $\frac{1}{2}\| \, |\psi\rangle\langle\psi| - |\psi'\rangle\langle\psi'| \, \|_1 \leq 2\sqrt{\eta}$, we conclude that $D_{\mathsf{X}}(T) \leq 2\sqrt{\eta} + \sqrt{\frac{|S|\cdot|T|}{2^n}}$. □

## 4.2 Proof that Property 4.2 holds for quantum Tanner codes

It remains to show that there exist codes that satisfy the conditions of Theorem 4.1. The codes of Leverrier and Zémor [43] (defined in Section 2.3.1) satisfy.

We start with the following claim which is stated along the same lines as [43, Theorem 1], and proved next. [43, Theorem 1] effectively proves Claim 4.6 for $\delta = 0$, which is the statement for distance. We use the expansion of the graphs to improve this to the small $\delta$ regime. We have changed some constants from the version in [43], for consistency purposes.

**Claim 4.6** *Fix $\lambda \in (0, \frac{1}{2})$, $\gamma \in (\frac{1}{2} + \lambda, 1)$ and $\kappa > 0$. Suppose $C_A, C_B$ have distance at least $\kappa\Delta$ and $C_0^\perp, C_1^\perp$ are $\Delta^{\frac{3}{2}-\lambda}$-robust with $\Delta^\gamma$ resistance to puncturing. Then there exist constants $c_1, c_2, \delta_0$ such that the following holds when $\delta \leq \delta_0$.*

1. *For any $x \in G_{\mathsf{X}}^\delta$ with $c_1\delta m_{\mathsf{X}} \leq |x| \leq c_2 n$, there is a $y \in C_{\mathsf{Z}}^\perp$ satisfying $|x \oplus y| < |x|$.*
2. *For any $z \in G_{\mathsf{Z}}^\delta$ with $c_1\delta m_{\mathsf{Z}} \leq |z| \leq c_2 n$, there is a $w \in C_{\mathsf{X}}^\perp$ satisfying $|z \oplus w| < |z|$.*

*Note that $\delta_0$ is chosen simply to ensure that $c_1\delta_0 m_{\mathsf{X}} \leq c_2 n$ and $c_1\delta_0 m_{\mathsf{Z}} \leq c_2 n$.*

We will now establish Property 4.2 using this claim. For $x \in G_{\mathsf{X}}^\delta$, if $c_1\delta m_{\mathsf{X}} \leq |x|_{C_{\mathsf{Z}}^\perp} \leq c_2 n$, then there is a $y' \in C_{\mathsf{Z}}^\perp$ such that $c_1\delta m_{\mathsf{X}} \leq |x \oplus y'| \leq c_2 n$. Note that $x \oplus y' \in G_{\mathsf{X}}^\delta$, since $H_{\mathsf{X}} y' = 0$. Thus, we can invoke Claim 4.6 (many times) to conclude that there is a $y \in C_{\mathsf{Z}}^\perp$ such that $|x \oplus y \oplus y'| < c_1\delta m_{\mathsf{X}}$. But $|x|_{C_{\mathsf{Z}}^\perp} \leq |x \oplus y \oplus y'| < c_1\delta m_{\mathsf{X}}$, leading to a contradiction. Thus, either $|x|_{C_{\mathsf{Z}}^\perp} \geq c_2 n$ or $|x|_{C_{\mathsf{Z}}^\perp} \leq c_1\delta m_{\mathsf{X}} = c_1\delta\frac{m_{\mathsf{X}}}{n} \cdot n$. We can argue similarly for $G_{\mathsf{Z}}^\delta$. Thus, Property 4.2 is satisfied with modified constant $\delta_0 \longrightarrow \delta_0 \cdot \frac{\min\{m_{\mathsf{X}}, m_{\mathsf{Z}}\}}{n}$.

**Proof of Claim 4.6:** We prove the first part of the claim. The second part follows along the same lines. Following [43], we define $\mathcal{G}_{1,x}^\square$ as the sub-graph of $\mathcal{G}_1^\square$ that is induced by $x \in G_{\mathsf{X}}^\delta$ (in other words, we only consider those edges of $\mathcal{G}_1^\square$ for which the corresponding squares have a '1' assigned by $x$). Let $S \subset V_1$ be the set of vertices in $\mathcal{G}_{1,x}^\square$. Most vertices $v$ in $S$ have their local view according to $C_1^{\perp}$. But, $x$ is an approximate codeword from $G_{\mathsf{x}}^{\delta}$. So there are no restrictions on the local views of at most $\delta m_{\mathsf{x}}$ vertices in $S$. We now modify the definition of 'exceptional vertices' from [43]. Let $S_e \subset S$ be the set of vertices $v$ which satisfy one of the two conditions:

- The degree is at least $\Delta^{\frac{3}{2}-\lambda}$ in $\mathcal{G}_{1,x}^{\square}$.
- The local view of $x$ at $v$ violates a check in $C_1^{\perp}$.

Since $|S| \geq \frac{2|x|}{\Delta^2}$, we choose $c_1 \overset{\mathsf{def}}{=} \frac{\Delta^{3-2\lambda}}{256}$ to conclude that $|S| \geq \frac{\Delta^{1-2\lambda}}{128}\delta m_{\mathsf{x}}$. Now, we establish the following bound on $|S_e|$, which modifies [43, Claim 9].

$$|S_e| \leq \frac{256|S|}{\Delta^{1-2\lambda}} + 2\delta m_{\mathsf{x}} \leq \frac{512|S|}{\Delta^{1-2\lambda}}. \tag{4.14}$$

To establish this bound, we proceed the same as [43]. Note that all the vertices in $S$ that are not 'violated' by $x$ have degree at least $\kappa\Delta$ (distance of the local code $C_1^{\perp}$). Thus, setting $c_2 \overset{\mathsf{def}}{=} \frac{\kappa\Delta^{\frac{1}{2}-\lambda}}{16} \cdot \frac{|V_1|}{n}$ and noting that $|S| \geq 2\delta m_{\mathsf{x}}$ for large constant $\Delta$, we obtain

$$\frac{|S|}{2} \leq (|S| - \delta m_{\mathsf{x}}) \leq \frac{2|x|}{\kappa\Delta} \implies |S| \leq \frac{4|x|}{\kappa\Delta} \leq \frac{|V_1|}{4\Delta^{\frac{1}{2}+\lambda}}. \tag{4.15}$$

If $|S_e| \leq 2\delta m_{\mathsf{x}}$, eq. (4.14) is verified. Otherwise, by using the expander mixing lemma and Lemma 2.12, we have

$$\frac{\Delta^{\frac{3}{2}-\lambda}}{2}|S_e| \leq \Delta^{\frac{3}{2}-\lambda}(|S_e| - \delta m_{\mathsf{x}}) \leq E(S_e, S), \tag{4.16a}$$

$$E(S_e, S) \leq \frac{\Delta^2|S_e||S|}{|V_1|} + 4\Delta\sqrt{|S_e||S|} \leq \frac{\Delta^{\frac{3}{2}-\lambda}}{4}|S_e| + 4\Delta\sqrt{|S_e||S|}, \tag{4.16b}$$

which implies $|S_e| \leq \frac{256|S|}{\Delta^{1-2\lambda}}$.

Having established eq. (4.14), which modifies a similar expression in [43] by a constant factor of 8, we proceed further in a very similar manner. We define the normal vertices ($S \setminus S_e$), heavy edges, and the set $T$ in the same manner. The upper bound on $|T|$ in [43, Claim 11] remains unchanged. To arrive at [43, Claim 12], the definition of $\alpha$ is slightly modified according to eq. (4.14). We need a vertex in $T$ that is not adjacent to a large number of vertices in $S_e$. For this, [43] upper bound $|E(S_e, T)|$ using the expander mixing lemma. The modified constants lead to a new upper bound

$$|E(S_e, T)| \leq \frac{256}{\Delta^{\frac{1}{2}-\lambda}}|T| + 128\Delta^{\lambda}\sqrt{|S||T|} \overset{\mathsf{def}}{=} \beta\Delta^{\frac{1}{2}+\lambda}|T|, \quad \beta = 256 + \frac{512}{\Delta}. \tag{4.17}$$

The rest of the argument from [43, Theorem 1] remains unchanged with the modified constants $\alpha, \beta$.

□

# Chapter 5

# The path to the quantum PCP conjecture

> The certitude that some shelf in some hexagon held precious books and that these precious books were inaccessible, seemed almost intolerable.
>
> Jorge Luis Borges, *The Library of Babel*

## 5.1 Lower bounds for one classical ansatz

We must first answer, did resolving the NLTS conjecture make any tangible progress towards answering the quantum PCP conjecture? The best answer to that question — we believe — is that it did considerable damage to *refuting* the quantum PCP conjecture. Low-depth or trivial circuits were a potential classical *ansatz* by which solutions to local Hamiltonians could be approximately described. At the very least, we have proven that low-depth circuits are not a generic enough ansatz with which we can classically approximate *all* local Hamiltonians. But perhaps, there exists another ansatz that can describe the low-energy states of all local Hamiltonians.

To be more specific, a classical ansatz is a classical string $w$, such as the description of a low-depth circuit, with which a *classical* verifier can verify the statement $\lambda_{\min}(\mathbf{H}) < b$. In the case of $w$ being the description of the low-depth circuit with low-energy, the verifier is given by Lemma 3.11. A different classical ansatz worth noting is the description of Clifford circuits (equivalently, stabilizer states). For any local Hamiltonian $\mathbf{H} = \sum h_i$ and a Clifford circuit $C$, the energy

$$\langle 0^{n'} | C^\dagger \mathbf{H} C | 0^{n'} \rangle = \sum_{ij} \langle 0^{n'} | C^\dagger P_{ij} C | 0^{n'} \rangle \tag{5.1}$$

where $h_i = \sum_j P_{ij}$ is a decomposition of each local Hamiltonian term into a linear combination of Paulis. Calculating, $\langle 0^{n'} | C^\dagger P_{ij} C | 0^{n'} \rangle$ is easy due to the Gottesman-Knill theorem [71]. Therefore, if every local Hamiltonian, had a low-energy state describable by a Clifford circuit then this would also disprove the quantum PCP conjecture (a la Lemma 3.11).

Since our proof of the NLTS theorem (Theorem 4.1) holds for stabilizer codes, then the ground space is a stabilizer subspace and has a Clifford circuit classical ansatz. This is one indication that

Theorem 4.1 is far from a proof of the quantum PCP conjecture; what it does say is that the two classical ansatzes, low-depth quantum circuits and Clifford circuits, are apples and oranges — i.e. incomparable — even in an approximate sense.

And, therefore, it is unclear if the circuit depth lower bounding techniques developed in this thesis directly apply to a potential quantum PCP construction. Furthermore, we suspect that stronger arguments than lightcone-based correlation arguments will be necessary for a quantum PCP construction.

## 5.2 The role of codes in quantum PCPs

It is tempting to say that the recent progress in constructions of good quantum LDPC codes [36–43] is simultaneously progress towards the quantum PCP conjecture. Such an argument could be made based on Theorem 4.1 or the fact that classical LTCs play an integral role in both the algebraic and combinatorial constructions of the PCP theorem. However, we believe that the relationship between quantum codes and a potential quantum PCP construction is not so clear.

For one, the construction of some optimal quantum code (be it LTC or some other property) is *not* known to be sufficient for quantum PCPs; instead, we suspect that a myriad of new mathematical insights will be required to construct quantum PCPs. But perhaps, the most damning reason to suspect that quantum codes may not be directly useful for quantum PCPs is the very boon that gives us Theorem 4.1, local indistinguishability.

At the highest level of generality, one could summarize the classical PCP theorem as an elegant locally testable code wrapped around the satisfying assignment for the formula [10, 11]. The very first PCP constructions were exponentially-long and came from exactly this inspiration; a simple construction involves encoding the solution to the NP-complete problem of "system of quadratic equations over $\mathbb{F}_2$" in a Hadamard code.

A reasonable first attempt to construct an exponentially-long quantum PCP would be to produce the same construction but with the two analogs: quantum locally testable codes and ground states of local Hamiltonians. Consider converting our original local Hamiltonian $\mathbf{H}$ instance into a new Hamiltonian $\mathbf{H}'$ such that the ground space of $\mathbf{H}'$ consists only of $\mathrm{Enc}(|\psi\rangle)$ where $|\psi\rangle$ is a ground state of $\mathbf{H}$. A possible attempt would be to have $\mathbf{H}' = \mathbf{H}'_{\text{code}} + \mathbf{H}'_{\text{test}}$ where $\mathbf{H}'_{\text{code}}$ would verify that the state was a code state of the quantum locally testable code and $\mathbf{H}'_{\text{test}}$ would check that the encoded state corresponded to a ground state of $\mathbf{H}$. Let $d$ be the distance of the locally testable code and $\ell$ is the locality of $\mathbf{H}'_{\text{test}}$. Then, due to local indistinguishability (Fact 3.17), if $d < \ell$ in order then $\mathbf{H}'_{\text{test}}$ cannot distinguish code states. Since $\ell = O(1)$, then $d = O(1)$ making the use of the code moot.

Therefore, we don't achieve an immediate construction of exponentially long quantum PCPs even if we were able to construct a family of quantum locally testable codes (even with low-rate analogous to the Hadamard code). The only known construction of exponential length quantum PCPs is due to the inclusion $\mathsf{QMA} \subseteq \mathsf{P}^{\#\mathsf{P}}$ and classical exponential length PCPs for $\mathsf{P}^{\#\mathsf{P}}$.

A more fundamental issue with the use of error correction in a construction of quantum PCPs is that the classical PCP constructions rely on the repetition of information in classical error correction. Concretely, in Dinur's construction [19], during the "graph powering" step, information is duplicated and consistency checks are added (which are later simplified during the "alphabet reduction" step). We do not know of any techniques for overcoming this fundamental barrier.

## 5.3 Future problems worth pondering

### A better ansatz for classical proofs

As we stated earlier, for every choice of classical ansatz, we could form an NLTS-like conjecture for lower bounds on the complexity of that ansatz. This would lead to, for example, a "Clifford-NLTS" conjecture, a "contractible Tensor Network-NLTS" conjecture, and so on. The true challenge would be to construct a family of local Hamiltonians for which each of the ansatzes was simultaneously futile. This would be a necessary consequence of the quantum PCP conjecture. Although not sufficient, the hope is that it would provide a more clear picture of what a quantum PCP Hamiltonian might look like.

Gharabian and Le Gall [72] recently introduced a variation of the NLTS theorem, coined the No-low energy samplable states (NLSS) conjecture, to broaden the class of quantum ansatzes.

**Conjecture 5.1 (NLSS Conjecture [72])** *There exists a fixed constant $\epsilon > 0$ and an explicit family of $O(1)$-local frustration-free commuting Hamiltonians $\{\mathbf{H}^{(n)}\}_{n=1}^{\infty}$ where $\mathbf{H}^{(n)} = \sum_{i=1}^{m} h_i^{(n)}$ acts on n particles and consists of $m = \Theta(n)$ local terms such that every state $|v\rangle$ with a succinct representation allowing perfect-sampling access,*

$$\langle v|\mathbf{H}|v\rangle \geq \lambda_{\min}(\mathbf{H}) + \epsilon m. \tag{5.2}$$

In [72], a state $|v\rangle$ with a succinct representation allowing perfect-sampling access is any state for which there is an efficient classical algorithm for $Q(x) = \langle x|v\rangle$ and an efficient sampler for the distribution $P(x) = |\langle x|v\rangle|^2$. In a similar vein to Lemma 3.11, [72] shows that the NLSS conjecture is a necessary consequence of the quantum PCP conjecture assuming $\mathsf{MA} \neq \mathsf{QMA}$. Since, it is unlikely that low-depth quantum circuits exhibit perfect-sampling access [73], the NLTS and NLSS conjecture appear somewhat orthogonal, a first indication that the NLSS conjecture is an interesting next step.

The NLSS conjecture is a strengthening of an alluded-to "Clifford-NLTS" conjecture. This is because of a characterization of the output of any Clifford circuit $C$ as

$$C\,|0^n\rangle \propto \sum_{Ax=b} i^{q(x)}\,|x\rangle \tag{5.3}$$

where $Ax = b$ is an affine subspace of $\{0,1\}^n$ and $q : \{0,1\}^n \to \mathbb{F}_4$ is a quadratic function [74]. Since it is easy to sample uniformly from $Ax = b$, these states have perfect-sampling access. This generalizes[1] to another class of well-studied states, phase states $|\psi_f\rangle$ (see Appendix A.3) for

[1] We call this a generalization due to the results of [4] which show that the restriction to a subspace of $\{0,1\}^n$ can be dropped through randomized reductions.

polynomial-time computable functions $f : \{0,1\}^n \to \{0,1\}$:

$$\left|\psi_f\right\rangle = \frac{1}{\sqrt{2^n}} \sum_{x\in\{0,1\}^n} (-1)^{f(x)} \left|x\right\rangle . \tag{5.4}$$

Since phase states for polynomial-time computable functions have perfect-sampling access, the NLSS conjecture that these states are also not low-energy. We hope that studying the NLSS conjecture or even just phase states might shed a more *algebraic* light on the quantum PCP conjecture and require a different set of tools than those used in Theorem 4.1. In particular, it is worth noting that ground states of local Hamiltonians are effectively characterized by phase states for functions $f \in$ PP (see Theorem A.11 in Appendix A.3); therefore the NLSS conjecture is asking for a P lower bound on characterizing the low-energy states. Perhaps our better understanding of the classical complexity theory pantheon will aid in proving the NLSS conjecture.

**A quantum proof for a classical problem** There are two natural definitions for non-deterministic quantum computation, QCMA or QMA, where QCMA requires that the proof be classical whereas the QMA proof can be quantum. One concrete way to interpret QCMA = QMA is that every local Hamiltonian problem has a ground state which can be described as the output of a *polynomial*-depth quantum circuit (a much more powerful class than the low-depth circuits considered for NLTS); conversely, QCMA ≠ QMA is equivalent to the existence of a family of QMA-hard local Hamiltonian problems with no ground states describable by polynomial-depth circuits.

If the quantum PCP conjecture is true and QCMA ≠ QMA, then there exists a family of QMA-hard local Hamiltonian problems with no low-energy states describable by polynomial-depth circuits. Our techniques for proving the NLTS theorem do not extend to this regime as they are based on lightcones. There are some natural complete problems for QCMA [75], most of which are QMA problems restricted to proofs describable by polynomial-depth circuits.

We ask then, what is the best construction of a probabilistically checkable proof for QCMA? Is there a polynomial-sized proof that can be efficiently checked by reading a constant number of terms? In other words, is there a reduction from QCMA to a promised-gap local Hamiltonian — i.e. Conjecture 3.9 but with QCMA-hardness? Note that the reduction must convert the problem into a local Hamiltonian problem and not a CSP as we suspect NP ≠ QCMA.

Our rationale behind considering QCMA over QMA is that an initial classical witness might sidestep many of the aforementioned issues about no-cloning and indistinguishability. Furthermore, promise-gapped local Hamiltonians may be ∈ QCMA; we have no strong evidence to suggest otherwise. Therefore, this is a more modest starting problem than the full quantum PCP conjecture.

**The relationship to multi-prover entangled proofs** A natural generalization of proofs is multi-prover interactive proofs (also known as games) [76]. In this model, multiple provers are trying to convince a verifier of the validity of a statement; the verifier asks them questions and they respond in turn in with answers. When the provers are not allowed to communicate nor share any entanglement (only classical correlations), and the questions and answers are restricted to poly$(n)$ bit length, this class is called MIP and is equal to NEXP due to [76]. Equivalently, let the value of a game be the maximal probability with which the verifier accepts where the maximum is taken over all strategies for the provers. Then, the problem of deciding, for a game with poly$(n)$ bit questions and answers, if the value of a game is either 1 or at most $1/2$ is $\mathsf{NEXP}$-complete. The PCP theorem can be rephrased in the language of games as it is $\mathsf{NP}$-complete to decide if the value of a multi-prover game is either 1 or at most $1/2$ when the question length is $O(\log n)$ bits and answer length is $O(1)$ bits — i.e. there exists a reduction from every $n$-bit $\mathsf{NP}$ problem to such a game.

The quantum generalization of multi-prover interactive proofs asks about deciding the *entangled* value of a game where the entangled value is the maximum over entangled strategies. Quantum games are, at their core, about the complexity of *bipartite* entanglement and bipartite quantum correlations since it has been shown that every game can be reduced to an equivalent game with two provers. This is in contrast to studying $\mathsf{QMA}$ which is about many-body entanglement. Furthermore, a priori, there is no bound on the entanglement necessary for an ideal strategy. In [77], a landmark result showed that when questions and answers are of poly$(n)$ bit length, the equivalent class for entangled games, $\mathsf{MIP}^\star$, equals $\mathsf{RE}$, the class of recursively enumerable languages. Since, $\mathsf{MIP}^\star$ is considerably more powerful than its classical counterpart, the status of a "games version" of the quantum PCP conjecture is unclear; it could sit anywhere between $\mathsf{NP}$ and $\mathsf{RE}$.

**Conjecture 5.2 (Quantum games PCP conjecture [78])** *There is an efficient quantum polynomial time reduction from every* $\mathsf{QMA}$ *problem of size $n$ to the problem of deciding the entangled value of a multi-prover game with question length of $O(\log n)$ bits and answer length of $O(1)$ bits.*

Note the previous conjecture only asks about the hardness of deciding the entangled value of the game. Progress towards the quantum games PCP conjecture was made by Natarajan and Vidick [78] but the validity of those arguments requires verification due to issues in the quantum sound low-degree test (see [79] for details on the issue and subsequent rectification). Therefore, we believe that the quantum games PCP conjecture is still open. Furthermore, the relationship between the quantum games PCP conjecture and the standard quantum "Hamiltonian" PCP conjecture (Conjecture 3.9) is unclear; [78] hints that a quantum "Hamiltonian" PCP conjecture for Hamiltonians of the $XX/ZZ$ type may imply the quantum games PCP conjecture. That implication is due to the low-degree test [79] and classical PCP techniques.

# Bibliography

[1] Anurag Anshu, Nikolas Breuckmann, and Chinmay Nirkhe. Nlts hamiltonians from good quantum codes, 2022.

[2] Anurag Anshu and Chinmay Nirkhe. Circuit Lower Bounds for Low-Energy States of Quantum Code Hamiltonians. In Mark Braverman, editor, *13th Innovations in Theoretical Computer Science Conference (ITCS 2022)*, volume 215 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 6:1–6:22, Dagstuhl, Germany, 2022. Schloss Dagstuhl – Leibniz-Zentrum für Informatik.

[3] Chinmay Nirkhe, Umesh Vazirani, and Henry Yuen. Approximate Low-Weight Check Codes and Circuit Lower Bounds for Noisy Ground States. In *45th International Colloquium on Automata, Languages, and Programming (ICALP 2018)*, Leibniz International Proceedings in Informatics (LIPIcs), pages 91:1–91:11, 2018.

[4] Sandy Irani, Anand Natarajan, Chinmay Nirkhe, Sujit Rao, and Henry Yuen. Quantum Search-To-Decision Reductions and the State Synthesis Problem. In Shachar Lovett, editor, *37th Computational Complexity Conference (CCC 2022)*, volume 234 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 5:1–5:19, Dagstuhl, Germany, 2022. Schloss Dagstuhl – Leibniz-Zentrum für Informatik.

[5] Anand Natarajan and Chinmay Nirkhe. A classical oracle separation between qma and qcma. 2022.

[6] Adam Bouland, Bill Fefferman, Chinmay Nirkhe, and Umesh Vazirani. On the complexity and verification of quantum random circuit sampling. *Nature Physics*, 2018.

[7] Thomas C. Bohdanowicz, Elizabeth Crosson, Chinmay Nirkhe, and Henry Yuen. Good approximate quantum ldpc codes from spacetime circuit hamiltonians. In *Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing*, STOC 2019, page 481–490, New York, NY, USA, 2019. Association for Computing Machinery.

[8] Srinivasan Arunachalam, Sergey Bravyi, Chinmay Nirkhe, and Bryan O'Gorman. The Parametrized Complexity of Quantum Verification. In François Le Gall and Tomoyuki Morimae, editors, *17th Conference on the Theory of Quantum Computation, Communication and Cryptography (TQC 2022)*, volume 232 of *Leibniz International Proceedings*

*in Informatics (LIPIcs)*, pages 3:1–3:18, Dagstuhl, Germany, 2022. Schloss Dagstuhl – Leibniz-Zentrum für Informatik.

[9] Robert W. Floyd. Nondeterministic algorithms. *J. ACM*, 14(4):636–644, oct 1967.

[10] Stephen A. Cook. The complexity of theorem-proving procedures. In *Proceedings of the Third Annual ACM Symposium on Theory of Computing*, STOC '71, pages 151–158, New York, NY, USA, 1971. ACM.

[11] L. A. Levin. Universal sequential search problems. *Problems of Information Transmission*, 9(3):265–266, 1973.

[12] Christopher Umans. Pseudo-random generators for all hardnesses. *Journal of Computer and System Sciences*, 67(2):419–440, 2003. Special Issue on STOC 2002.

[13] Alexei Kitaev. Lecture notes in computer assisted diagnosis, 1999.

[14] Alexei Yu Kitaev, Alexander Shen, Mikhail N Vyalyi, and Mikhail N Vyalyi. *Classical and quantum computation*. Number 47. American Mathematical Soc., 2002.

[15] Julia Kempe, Alexei Kitaev, and Oded Regev. The complexity of the local Hamiltonian problem. *SIAM Journal on Computing*, 35(5):1070–1097, 2006.

[16] Ryan O'Donnell, Dana Moshkovitz, and Irit Dinur. A history of the pcp theorem.

[17] Sanjeev Arora and Shmuel Safra. Probabilistic checking of proofs: A new characterization of np. *J. ACM*, 45(1):70–122, January 1998.

[18] Sanjeev Arora, Carsten Lund, Rajeev Motwani, Madhu Sudan, and Mario Szegedy. Proof verification and the hardness of approximation problems. *J. ACM*, 45(3):501–555, May 1998.

[19] Irit Dinur. The pcp theorem by gap amplification. *J. ACM*, 54(3):12–es, June 2007.

[20] Dorit Aharonov and Tomer Naveh. Quantum NP - A Survey, 2002.

[21] Dorit Aharonov, Itai Arad, and Thomas Vidick. Guest Column: The Quantum PCP Conjecture. *SIGACT News*, 44(2):47–79, June 2013.

[22] Alexei Kitaev and John Preskill. Topological entanglement entropy. *Phys. Rev. Lett.*, 96:110404, Mar 2006.

[23] Robert Raussendorf and Hans J. Briegel. A one-way quantum computer. *Phys. Rev. Lett.*, 86:5188–5191, May 2001.

[24] Robert Raussendorf, Daniel E. Browne, and Hans J. Briegel. Measurement-based quantum computation on cluster states. *Phys. Rev. A*, 68:022312, Aug 2003.

[25] Steven R. White. Density matrix formulation for quantum renormalization groups. *Phys. Rev. Lett.*, 69:2863–2866, Nov 1992.

[26] Yudong Cao, Jonathan Romero, Jonathan P. Olson, Matthias Degroote, Peter D. Johnson, Mária Kieferová, Ian D. Kivlichan, Tim Menke, Borja Peropadre, Nicolas P. D. Sawaya, Sukin Sim, Libor Veis, and Alán Aspuru-Guzik. Quantum chemistry in the age of quantum computing. *Chemical Reviews*, 119(19):10856–10915, Oct 2019.

[27] B. Apolloni, C. Carvalho, and D. de Falco. Quantum stochastic optimization. *Stochastic Processes and their Applications*, 33(2):233 – 244, 1989.

[28] Tadashi Kadowaki and Hidetoshi Nishimori. Quantum annealing in the transverse ising model. *Phys. Rev. E*, 58:5355–5363, Nov 1998.

[29] Alberto Peruzzo, Jarrod McClean, Peter Shadbolt, Man-Hong Yung, Xiao-Qi Zhou, Peter J. Love, Alán Aspuru-Guzik, and Jeremy L. O'Brien. A variational eigenvalue solver on a photonic quantum processor. *Nature Communications*, 5(1):4213, Jul 2014.

[30] Ryan LaRose, Arkin Tikku, Étude O'Neel-Judy, Lukasz Cincio, and Patrick J. Coles. Variational quantum state diagonalization. *npj Quantum Information*, 5(1):57, Jun 2019.

[31] Edward Farhi, Jeffrey Goldstone, and Sam Gutmann. A quantum approximate optimization algorithm. *arXiv preprint arXiv:1411.4028*, 2014.

[32] Michael H. Freedman and Matthew B. Hastings. Quantum systems on non-k-hyperfinite complexes: A generalization of classical statistical mechanics on expander graphs. *Quantum Info. Comput.*, 14(1–2):144–180, January 2014.

[33] Sergey Bravyi and Mikhail Vyalyi. Commutative version of the local hamiltonian problem and common eigenspace problem. *Quantum Info. Comput.*, 5(3):187–215, May 2005.

[34] Fernando G.S.L. Brandao and Aram W. Harrow. Product-state approximations to quantum ground states. In *Proceedings of the Forty-Fifth Annual ACM Symposium on Theory of Computing*, STOC '13, page 871–880, New York, NY, USA, 2013. Association for Computing Machinery.

[35] Dorit Aharonov and Lior Eldar. The commuting local hamiltonian problem on locally expanding graphs is approximable in np. *Quantum Information Processing*, 14(1):83–101, January 2015.

[36] A.Yu. Kitaev. Fault-tolerant quantum computation by anyons. *Annals of Physics*, 303(1):2 – 30, 2003.

[37] Michael H Freedman, David A Meyer, and Feng Luo. Z2-systolic freedom and quantum codes. *Mathematics of quantum computation, Chapman & Hall/CRC*, pages 287–320, 2002.

[38] Jean-Pierre Tillich and Gilles Zémor. Quantum ldpc codes with positive rate and minimum distance proportional to the square root of the blocklength. *IEEE Transactions on Information Theory*, 60(2):1193–1202, 2014.

[39] Matthew B. Hastings, Jeongwan Haah, and Ryan O'Donnell. Fiber bundle codes: Breaking the $n^{1/2}$ polylog$(n)$ barrier for quantum ldpc codes, 2020.

[40] Nikolas P. Breuckmann and Jens N. Eberhardt. Balanced Product Quantum Codes. *IEEE Transactions on Information Theory*, 67(10):6653–6674, 2021.

[41] Pavel Panteleev and Gleb Kalachev. Quantum ldpc codes with almost linear minimum distance, 2020.

[42] Pavel Panteleev and Gleb Kalachev. Asymptotically Good Quantum and Locally Testable Classical LDPC Codes, 2021.

[43] Anthony Leverrier and Gilles Zémor. Quantum Tanner codes, 2022.

[44] Lior Eldar. Robust quantum entanglement at (nearly) room temperature. *arXiv: Quantum Physics (to appear in ITCS 2021)*, 2019.

[45] Anurag Anshu and Nikolas P. Breuckmann. A construction of Combinatorial NLTS, 2022.

[46] Michael A Nielsen and Isaac Chuang. Quantum computation and quantum information, 2002.

[47] Sanjeev Arora and Boaz Barak. *Computational Complexity: A Modern Approach*. Cambridge University Press, New York, NY, USA, 1st edition, 2009.

[48] Peter W. Shor. Scheme for reducing decoherence in quantum computer memory. *Phys. Rev. A*, 52:R2493–R2496, Oct 1995.

[49] Emanuel Knill and Raymond Laflamme. Theory of quantum error-correcting codes. *Phys. Rev. A*, 55:900–911, Feb 1997.

[50] A. M. Steane. Error correcting codes in quantum theory. *Phys. Rev. Lett.*, 77:793–797, Jul 1996.

[51] A. R. Calderbank and Peter W. Shor. Good quantum error-correcting codes exist. *Phys. Rev. A*, 54:1098–1105, Aug 1996.

[52] L. Eldar and A. W. Harrow. Local hamiltonians whose ground states are hard to approximate. In *2017 IEEE 58th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 427–438, 2017.

[53] Dorit Aharonov and Lior Eldar. Quantum locally testable codes. *SIAM Journal on Computing*, 44(5):1230–1262, 2015.

[54] R Tanner. A recursive approach to low complexity codes. *IEEE Transactions on information theory*, 27(5):533–547, 1981.

[55] Michael Sipser and Daniel A Spielman. Expander codes. *IEEE transactions on Information Theory*, 42(6):1710–1722, 1996.

[56] Alexander Lubotzky, Ralph Phillips, and Peter Sarnak. Ramanujan graphs. *Combinatorica*, 8(3):261–277, 1988.

[57] Grigorii Aleksandrovich Margulis. Explicit group-theoretical constructions of combinatorial schemes and their application to the design of expanders and concentrators. *Problemy peredachi informatsii*, 24(1):51–60, 1988.

[58] Libor Caha, Zeph Landau, and Daniel Nagaj. Clocks in feynman's computer and kitaev's local hamiltonian: Bias, gaps, idling, and pulse tuning. *Phys. Rev. A*, 97:062306, Jun 2018.

[59] Johannes Bausch and Elizabeth Crosson. Analysis and limitations of modified circuit-to-Hamiltonian constructions. *Quantum*, 2:94, September 2018.

[60] Itai Arad, Zeph Landau, and Umesh Vazirani. Improved one-dimensional area law for frustration-free systems. *Physical Review B*, 85:195145, May 2012.

[61] Anurag Anshu, Itai Arad, and David Gosset. An area law for 2d frustration-free spin systems, 2021.

[62] Jeff Kahn, Nathan Linial, and Alex Samorodnitsky. Inclusion-exclusion: Exact and approximate. *Combinatorica*, 16(4):465–477, Dec 1996.

[63] H. Buhrman, R. Cleve, R. De Wolf, and C. Zalka. Bounds for small-error and zero-error quantum algorithms. pages 358–368, 1999.

[64] Shachar Lovett and Emanuele Viola. Bounded-depth circuits cannot sample good codes. *Electronic Colloquium on Computational Complexity (ECCC)*, 17:115, 10 2010.

[65] Rui Chao, Ben W. Reichardt, Chris Sutherland, and Thomas Vidick. Overlapping Qubits. In Christos H. Papadimitriou, editor, *8th Innovations in Theoretical Computer Science Conference (ITCS 2017)*, volume 67 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 48:1–48:21, Dagstuhl, Germany, 2017. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik.

[66] Sergey Bravyi, David Poulin, and Barbara Terhal. Tradeoffs for reliable quantum information storage in 2d systems. *Phys. Rev. Lett.*, 104:050503, Feb 2010.

[67] Scott Aaronson and Daniel Gottesman. Improved simulation of stabilizer circuits. *Phys. Rev. A*, 70:052328, Nov 2004.

[68] Irit Dinur, Min-Hsiu Hsieh, Ting-Chun Lin, and Thomas Vidick. Good Quantum LDPC Codes with Linear Time Decoders, 2022.

[69] Max Hopkins and Ting-Chun Lin. Explicit lower bounds against omega(n)-rounds of sum-of-squares, 2022.

[70] A. Winter. Coding theorem and strong converse for quantum channels. *IEEE Transactions on Information Theory*, 45(7):2481–2485, 1999.

[71] Daniel Gottesman. The heisenberg representation of quantum computers. *Group22: Proceedings of the XXII International Colloquium on Group Theoretical Methods in Physics*, 1998.

[72] Sevag Gharibian and François Le Gall. Dequantizing the quantum singular value transformation: Hardness and applications to quantum chemistry and the quantum pcp conjecture, 2021.

[73] Barbara M Terhal and David P DiVincenzo. Adaptive quantum computation, constant depth quantum circuits and Arthur-Merlin games. *Quantum Information & Computation*, 4(2):134–145, 2004.

[74] Maarten Van den Nest. Classical simulation of quantum computation, the gottesman-knill theorem, and slightly beyond. *Quantum Inf. Comput.*, 10:258–271, 2010.

[75] Pawel Wocjan, Dominik Janzing, and Thomas Beth. Two qcma-complete problems, 2003.

[76] L. Babai, Lance Fortnow, and Carsten Lund. Nondeterministic exponential time has two-prover interactiveprotocols. pages 16–25 vol.1, 11 1990.

[77] Zhengfeng Ji, Anand Natarajan, Thomas Vidick, John Wright, and Henry Yuen. Mip*=re, 2020.

[78] Anand Natarajan and Thomas Vidick. Low-degree testing for quantum states, and a quantum entangled games pcp for qma. In *2018 IEEE 59th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 731–742, 2018.

[79] Zhengfeng Ji, Anand Natarajan, Thomas Vidick, John Wright, and Henry Yuen. Quantum soundness of the classical low individual degree test, 2020.

[80] A. Uhlmann. The "transition probability" in the state space of a *-algebra. *Rep. Math. Phys.*, 9:273–279, 1976.

[81] Matthew B Hastings. An area law for one-dimensional quantum systems. *Journal of Statistical Mechanics: Theory and Experiment*, 2007(08):P08024, 2007.

[82] Itai Arad, Alexei Kitaev, Zeph Landau, and Umesh Vazirani. An area law and sub-exponential algorithm for 1D systems, 2013. arXiv preprint arXiv: 1301.1162.

[83] Anurag Anshu, Aram W. Harrow, and Mehdi Soleimanifar. From communication complexity to an entanglement spread area law in the ground state of gapped local hamiltonians. *https://arxiv.org/abs/2004.15009*, 2020.

[84] Chris Beck, Russell Impagliazzo, and Shachar Lovett. Large deviation bounds for decision trees and sampling lower bounds for ac0-circuits. In *2012 IEEE 53rd Annual Symposium on Foundations of Computer Science*, pages 101–110, 2012.

[85] Scott Aaronson. Open problems related to quantum query complexity. *ACM Transactions on Quantum Computing*, 2(4), dec 2021.

[86] Scott Aaronson and Greg Kuperberg. Quantum versus classical proofs and advice. In *Twenty-Second Annual IEEE Conference on Computational Complexity (CCC'07)*, pages 115–128, 2007.

[87] Bill Fefferman and Shelby Kimmel. Quantum vs. classical proofs and subset verification. In Igor Potapov, Paul G. Spirakis, and James Worrell, editors, *43rd International Symposium on Mathematical Foundations of Computer Science, MFCS 2018, August 27-31, 2018, Liverpool, UK*, volume 117 of *LIPIcs*, pages 22:1–22:23. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2018.

[88] Andrew Lutomirski. Component mixers and a hardness result for counterfeiting quantum money, 2011.

[89] Atul Singh Arora, Alexandru Gheorghiu, and Uttam Singh. Oracle separations of hybrid quantum-classical circuits. 2022.

[90] Andris Ambainis, Andrew M. Childs, and Yi-Kai Liu. Quantum property testing for bounded-degree graphs. In Leslie Ann Goldberg, Klaus Jansen, R. Ravi, and José D. P. Rolim, editors, *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques*, pages 365–376, Berlin, Heidelberg, 2011. Springer Berlin Heidelberg.

[91] Adam D. Bookatz. Qma-complete problems. 2012.

[92] Charles H. Bennett, Ethan Bernstein, Gilles Brassard, and Umesh V. Vazirani. Strengths and weaknesses of quantum computing. *SIAM J. Comput.*, 26(5):1510–1523, 1997.

[93] Dexter Kozen. Lower bounds for natural proof systems. In *Proceedings of the 18th Annual Symposium on Foundations of Computer Science*, SFCS '77, page 254–266, USA, 1977. IEEE Computer Society.

[94] Andris Ambainis. Quantum lower bounds by quantum arguments. *J. Comput. Syst. Sci.*, 64(4):750–767, jun 2002.

[95] Honghao Fu. Personal Communication, Oct 2022.

[96] Hartmut Klauck and Supartha Podder. Two results about quantum messages. In *International Symposium on Mathematical Foundations of Computer Science*, pages 445–456. Springer, 2014.

[97] Chinmay Nirkhe. NLTS Hamiltonians from codes, 2022. Simons Institute for the Theory of Computing Quantum Colloquium. Panel Umesh Vazirani, Dorit Aharanov, Matthew Hastings, Anand Natarajan, and Chinmay Nirkhe.

[98] Sergey Bravyi and Barbara Terhal. Complexity of stoquastic frustration-free hamiltonians. *SIAM J. Comput.*, 39(4):1462–1485, nov 2009.

[99] Stephen P Jordan, David Gosset, and Peter J Love. Quantum-merlin-arthur-complete problems for stoquastic hamiltonians and markov matrices. *Physical Review. A*, 81(3), 3 2010.

[100] Sergey Bravyi, David P. Divincenzo, Roberto Oliveira, and Barbara M. Terhal. The complexity of stoquastic local hamiltonian problems. *Quantum Info. Comput.*, 8(5):361–385, may 2008.

[101] Scott Aaronson. The complexity of quantum states and transformations: from quantum money to black holes. *arXiv preprint arXiv:1607.05256*, 2016.

[102] William Kretschmer. Quantum pseudorandomness and classical complexity. *arXiv preprint arXiv:2103.09320*, 2021.

[103] Gregory Rosenthal and Henry Yuen. Interactive proofs for synthesizing quantum states and unitaries. *arXiv preprint arXiv:2108.07192*, 2021.

[104] Bill Fefferman and Cedric Lin. Quantum Merlin Arthur with Exponentially Small Gap, 2016.

[105] Abhinav Deshpande, Alexey V. Gorshkov, and Bill Fefferman. The importance of the spectral gap in estimating ground-state energies, 2020.

[106] Dorit Aharonov, Michael Ben-Or, Fernando G. S. L. Brandao, and Or Sattath. The pursuit of uniqueness: Extending Valiant-Vazirani theorem to the probabilistic and quantum settings, 2008.

# Appendix A

# Additional lower bounds on description complexity

> Whether or not you find your own way, you're bound to find some way. If you happen to find my way, please return it, as it was lost years ago. I imagine by now it's quite rusty.
>
> Norton Juster, *The Phantom Tollbooth*

The circuit depth lower bounds used in proving the NLTS theorem are an example of what we call *description complexity* lower bounds. Description complexity is the minimal length of a *useful* classical description of an object. To illustrate, what we mean by useful, consider the following two examples: (1) the description of a classical circuit $U$ such that its output is purported to be the ground state of some local Hamiltonian $\mathbf{H}$, and (2) the statement "the minimum eigenvector of $\mathbf{H}$". Both may describe the same object but the first is useful for a $\mathsf{BQP}$ or stronger device (as the $\mathsf{BQP}$ device can verify that it is the ground state) while the second is only useful for a $\mathsf{QCMA}$ or stronger device. Therefore, there is a subtlety here between description complexity and Kolmogorov complexity as implicitly there is a verification (computational problem) associated with the description.

The NLTS theorem is one example of description complexity lower bounds (when restricted to descriptions that are circuit descriptions). Appendix A.1 provides additional circuit depth lower bounds from quantum error correction. In Appendix A.2, we discuss superpolynomial lower bounds on the complexity of ground states in the classical oracle model. In Appendix A.3, we discuss the minimum description complexity of $\mathsf{QMA}$-search problems through the lens of search-to-decision reductions and state synthesis.

## A.1 Intermediate results leading to the NLTS theorem

*This section is based on [2] by Anshu and Nirkhe.*

### A.1.1 Omitted proofs from earlier sections

**Proof of Lemma 3.22:** Since $\mathrm{tr}(\mathbf{H}\phi) \leq \epsilon m$, Markov's inequality ensures that $\mathrm{tr}(D_{\leq 2\epsilon m}\phi) \geq \frac{1}{2}$, where $D_{\leq 2\epsilon m}$ is the subspace (and projector onto) of energy $\leq 2\epsilon m$. Since

$$D_{\leq 2\epsilon m} = \sum_{s\in\{0,1\}^m : |s|\leq 2\epsilon m} D_s \tag{A.1}$$

and the number of $s$ satisfying $|s| \leq 2\epsilon m$ is[1]

$$\binom{m}{2\epsilon m} \leq 2^{2\sqrt{2\epsilon}m}, \tag{A.2}$$

there exists a $s^\star \in \{0,1\}^m$ such that

$$\mathrm{tr}(D_{s^\star}\phi) \geq 2^{-2\sqrt{2\epsilon}m-1}. \tag{A.3}$$

Now, Fact 2.4 ensures that $D_{s^\star}$ is also an error correcting code of distance $d$. Apply Lemma 3.21 (assuming $2^t \leq \frac{d}{2}$) and using $n' \leq 2^t n$ (using Fact 3.4) and $m \leq n\ell$, we conclude

$$2^{-2\sqrt{2\epsilon}n\ell-1} \leq 2\exp\left(-\frac{d^2}{400\cdot 2^{3t}n}\right). \tag{A.4}$$

Solving for $t$ and simplifying constants gives us the desired statement. □

**Proof of Lemma 3.23:** Let $\psi$ be a state on $n'$ qubits such that $\psi = U\left|0^{\otimes n'}\right\rangle$ where $U$ is a circuit of depth $t$. Suppose $2^t < d$. Further assume that $\psi$ is $\delta$-close to the code $C$ in trace distance, meaning that there exists a state $\rho_{\text{code}} \in C$ such that $\|\psi_{\text{code}} - \rho_{\text{code}}\|_1 \leq \delta$. Thus, Uhlmann's theorem [80] ensures that there is a purification $|\rho\rangle$ on $n'$ qubits such that $\||\psi\rangle\langle\psi| - |\rho\rangle\langle\rho|\|_1 \leq \delta$.

Let Enc be any encoding CPTP map from $(\mathbb{C}^2)^{\otimes k} \to (\mathbb{C}^2)^{\otimes n}$ mapping $k$ qubits to the $k$ qubit code space. Define $\mathcal{E}$ as the maximally decohering channel as follows

$$\mathcal{E}(\cdot) \stackrel{\text{def}}{=} \frac{1}{4^k}\sum_{a,b\in\{0,1\}^k}\left(X^aZ^b\right)(\cdot)\left(X^aZ^b\right)^\dagger. \tag{A.5}$$

Then let $\Theta$ be the encoding of $\rho$ defined as

$$\Theta \stackrel{\text{def}}{=} \mathrm{Enc}\circ\mathcal{E}\circ\mathrm{Enc}^{-1}(\rho). \tag{A.6}$$

This state is well-defined and has entropy $S(\Theta) \geq k$ since $S(\mathcal{E}(\rho)) \geq k$.

**Fact A.1 (Extended local indistinguishability property)** *For any region $R_1 \cup R_2$ where $R_1$ is contained in the code qubits and $R_2$ in the ancilla qubits with $|R_1| < d$, $\rho_{R_1\cup R_2} = \Theta_{R_1\cup R_2}$.*

[1] A tighter bound can improve the $\epsilon$ dependence to $\epsilon\log(1/\epsilon)$ [2].

We prove this fact after the lemma. Let $R \subset [n']$ be any region of the qubits of size $< d$. Using this fact,

$$\|\psi_R - \Theta_R\|_1 \le \delta. \tag{A.7}$$

Applying Fact 3.3 here, we have

$$\mathrm{tr}_{-\{i\}}(U^\dagger \Theta U) = \mathrm{tr}_{-\{i\}}(U^\dagger(\Theta_{L_i} \otimes \nu_{-L_i})U). \tag{A.8}$$

Since the size of $L_i$ is $< d$, we can combine eq. (A.7) and eq. (A.8) to achieve

$$\left\|\mathrm{tr}_{-\{i\}}(U^\dagger(\psi_{L_i} \otimes \nu_{-L_i})U) - \mathrm{tr}_{-\{i\}}(U^\dagger(\Theta_{L_i} \otimes \nu_{-L_i})U)\right\|_1 \le \delta. \tag{A.9}$$

However, $U^\dagger \psi U = |0^{\otimes n'}\rangle\langle 0^{\otimes n'}|$ and so

$$\left\||0\rangle\langle 0| - \mathrm{tr}_{-\{i\}}(U^\dagger(\Theta_{L_i} \otimes \nu_{-L_i})U)\right\|_1 \le \delta. \tag{A.10}$$

Using standard entropy bounds, we can bound the entropy of the $i$th qubit of the rotated state $\Theta$:

$$S\left(\mathrm{tr}_{-\{i\}}(U^\dagger \Theta U)\right) = S\left(\mathrm{tr}_{-\{i\}}(U^\dagger(\Theta_{L_i} \otimes \nu_{-L_i})U)\right) \le H_2(\delta) \le 2\delta \log(1/\delta). \tag{A.11}$$

Notice that $S(U^\dagger \Theta U) = S(\Theta) = k$. We can, therefore, bound $k$ by

$$k \le S(\Theta) \le \sum_{i \in [n']} S\left(\mathrm{tr}_{-\{i\}}(U^\dagger \Theta U)\right) \le 2\delta \log(1/\delta) n'. \tag{A.12}$$

This leads to a contradiction since we assumed $k > 2\delta \log(1/\delta) n'$. □

**Proof of Fact A.1:** Let $R_1$ be a subset of the code qubits and $R_2$ be a subset of the ancilla qubits such that $|R_1| < d$. We can express any code state $|\psi\rangle$ over the $n'$ qubits as

$$|\psi\rangle = \sum_{x \in \{0,1\}^k} |\overline{x}\rangle |\psi_x\rangle \tag{A.13}$$

where $\{|\overline{x}\rangle\}$ is a basis for the code and $|\psi_x\rangle$ are un-normalized. Let $U$ be any logical operator (i.e. one that preserves the code space). Then,

$$\mathrm{tr}_{-(R_1 \cup R_2)}\left(U \psi U^\dagger\right) = \sum_{x,y \in \{0,1\}^k} \mathrm{tr}_{-R_1}(U |\overline{x}\rangle\langle\overline{y}| U^\dagger) \otimes \mathrm{tr}_{-R_2}(|\psi_x\rangle\langle\psi_y|) \tag{A.14}$$

In the summation, if $x = y$, then the first component is $\phi_{R_1}$ for some fixed state $\phi_{R_1}$ by local indistinguishability. Furthermore, if $x \neq y$, then the first component is 0 by orthogonality of $U|\overline{x}\rangle$ and $U|\overline{y}\rangle$ despite the erasure of $|R_1| < d$ qubits. Therefore,

$$\mathrm{tr}_{-(R_1 \cup R_2)}\left(U \psi U^\dagger\right) = \phi_{R_1} \otimes \sum_{x \in \{0,1\}^k} \mathrm{tr}_{-R_2}(\psi_x). \tag{A.15}$$

which is an invariant of $U$, which means $\mathrm{tr}_{-(R_1 \cup R_2)}(\rho) = \mathrm{tr}_{-(R_1 \cup R_2)}(U \rho U^\dagger)$ . Since (a) $\Theta$ is a mixture over applications of *logical* Paulis to $\rho$ and (b) a logical operator applied to a code state is another code state and therefore is locally indistinguishable, then it follows that $\rho_{R_1 \cup R_2} = \Theta_{R_1 \cup R_2}$. □

### A.1.2 Lower bounds for high-rate codes

The following is a simple corollary of Lemma 3.23.

**Corollary A.2** *Let $C$ be a $[[n,k,d]]$ code and $|\psi\rangle$ a pure-state and the trace-distance between $|\psi\rangle$ and $C$ is $0<\delta<1/2$ such that $k>2\delta\log(1/\delta)n$. Then, the circuit complexity $\mathsf{cc}(|\psi\rangle)$ satisfies*

$$\mathsf{cc}(|\psi\rangle) \geq \log\left(\min\left\{d, \frac{k}{2\delta\log(1/\delta)n}\right\}\right). \tag{A.16}$$

**Proof:** By Lemma 3.23, either $2^{\mathsf{cc}(|\psi\rangle)} \geq d$ or $k \leq 2\delta\log(1/\delta)2^{\mathsf{cc}(|\psi\rangle)}n$ since $m \leq 2^{\mathsf{cc}(|\psi\rangle)}n$. Rearranging this is equivalent to the corollary. □

Corollary A.2 shows that if we are given a $[[n,k,d]]$ code with linear rate $k=\Omega(n)$, then a state generated by a depth $t\leq\gamma\log d$ circuit must be $\Omega\left(2^{-2t}\right)=\Omega\left(d^{-2\gamma}\right)$ far from the code space in trace distance. An even stronger separation holds if the code is the zero-eigenspace (ground-space) of a commuting local Hamiltonian.

Consider a $[[n,k,d]]$ QLDPC code $C$ which is the common zero-eigenspace of commuting checks $\{\Pi_j\}_{j=1}^m$ of locality $\ell$ each (this includes, but is not restricted to, the stabilizer code defined earlier). Consider a state $|\psi\rangle = U|0^{\otimes n'}\rangle$ obtained by applying a depth $t$ circuit $U$ on $n'$ qubits. Suppose there is a state $\rho_0\in C$ having good fidelity with the code space, that is, $F \overset{\mathsf{def}}{=} F(\psi_{\text{code}},\rho_0)$. Fact 3.4 ensures that we can choose $n'\leq 2^t n$. We prove the following lemma.

**Lemma A.3** *For the state $|\psi\rangle$ as defined above, it holds that*

$$2^{2t} \geq \min\left(d, \frac{1}{64\sqrt{\ell}\log^2 d\ell}\cdot\frac{k\sqrt{d}}{n\cdot\sqrt{\log\frac{1}{F}}}\right). \tag{A.17}$$

The proof of this lemma appears in [2, Appendix B]. It uses the tool of approximate ground-space projectors (AGSP) and the principle that low min-entropy for gapped ground states implies low entanglement entropy [60, 81–83]. The lemma shows that if the code has linear rate $k=\Omega(n)$, then any state generated by a depth $t\leq\gamma\log d$ circuit must be $1-\exp\left(-\widetilde{\Omega}\left(d^{1-4\gamma}\right)\right)$ far from the code space in trace distance. Here, the $\tilde{\Omega}$ notation hides some polylog factors. This bears some resemblance with the results of [64, 84], which show that the distributions sampled from depth $t$ (and size $e^{n^{1/t}}$) classical $\mathsf{AC0}$ circuits are $1-e^{-n^{1/t}}$ far from the uniform distribution over a good classical code (linear rate and linear distance). A comparison with Lemma A.3 is largely unclear, due to the differences between classical and quantum codes, as well as $\mathsf{AC0}$ circuits and quantum circuits.

We now prove Theorem 3.24 by showing how the entropy-based bounds from the previous results can be improved from handling states physically near the code space to all low-energy states, once we assume that the code is a stabilizer code. The key property we exploit is that the local indistinguishability property of the code space $C$ also holds for each eigenspace $D_s$ in the case of stabilizer codes. We make this precise in the following facts; all facts are proven after the proof of the theorems.

The following fact argues that logical operators not only preserve the code space $C$ but rather any eigenspace $D_s$.

**Fact A.4** *Fix a stabilizer code C on n qubits with generator set* $\{C_i\}_{i\in[m]}$*. For any string* $s \in \{0,1\}^n$*, a state* $\rho$ *such that* $\rho_{\text{code}} \in D_s$*, and a logical operator* $P \in \mathcal{L}$*, we have* $(P\rho P)_{\text{code}} \in D_s$*.*

Each pair of Pauli operators either commute or anti-commute. The following fact imposes constraints on non-logical and non-stabilizer Pauli operators.

**Fact A.5** *Let P be a Pauli operator such that for some* $i \in [m]$*,* $PC_i = -C_iP$*. For any* $s \in \{0,1\}^m$ *and any quantum state* $\rho$ *such that* $\rho_{\text{code}} \in D_s$*, we have* $\mathrm{tr}(P\rho) = 0$*.*

The third crucial fact we use is Fact 2.4, the local indistinguishability of every eigenspace $D_s$.

Since logical operators act like single-qubit Pauli operators within the code space, they can be used for randomization. Define the following quantum "completely depolarizing in the logical basis" channel that acts on code qubits, analogous to the channel defined in eq. (A.6):

$$\mathcal{E}(\cdot) \overset{\text{def}}{=} \frac{1}{4^k} \sum_{a,b\in\{0,1\}^k} \left(\overline{X}^a \overline{Z}^b\right)(\cdot)\left(\overline{Z}^b \overline{X}^a\right) \tag{A.18}$$

where $\overline{X}^a = \prod_i \overline{X_i}^{a_i}$ is a product of logical $X$ operators defined by $a$ and likewise $\overline{Z}^b$ is a product of logical $Z$ operators defined by $b$. We will utilize the following two properties of this channel, analogous to Fact A.1.

**Fact A.6** *It holds that*

1. *For any quantum state* $\rho$*, the entropy* $S(\mathcal{E}(\rho)) \geq k$*.*
2. *For any quantum state* $\rho$ *such that* $\rho_{\text{code}} \in D_s$ *for some s,* $\mathcal{E}(\rho)_{\text{code}} \in D_s$*. Furthermore, for any set* $T \subset [n']$ *of size less than d,* $\rho_T = \mathcal{E}(\rho)_T$*.*

The next fact describes how all stabilizer terms of the code can be measured simultaneously using a short-depth circuit if the code has small locality. Let $m$ be the number of checks for an $\ell$-local code; recall that then $n/\ell \leq m \leq \ell n$.

**Fact A.7** *Let C be a stabilizer code of locality* $\ell$ *on n qubits with m checks* $\{C_i\}_{i\in[m]}$*. Then, there is a circuit V of depth* $\leq 2\ell^3$ *which coherently measures the value of each stabilizer term into m ancilla.*

Lastly, consider a state $|\phi\rangle = U|0^{\otimes n'}\rangle$ where $U$ is a circuit of depth $t$. From Fact 3.4, we can assume $n' \leq n2^t$ without loss of generality. We are now ready to state and prove the following theorem for codes of large rate.

**Proof of Theorem 3.24:** All stated intermediate claims are proven in the next sub-section. By assumption, $|\phi\rangle = U|0^{\otimes n'}\rangle$ for a circuit of depth $t < \log(d) - 2\ell^3$. Define the energy of each local Hamiltonian term $h_i$ as

$$\epsilon_i \overset{\text{def}}{=} \operatorname{tr}(h_i\phi) = \frac{1}{2} - \frac{1}{2}\operatorname{tr}(C_i\phi). \tag{A.19}$$

Add $m \leq \ell n$ new syndrome-measurement ancilla (SMA) qubits each with initial state $|0\rangle$ and coherently measure the entire syndrome using the depth $2\ell^3$ circuit $V$ from Fact A.7. Then the state

$$|\psi\rangle = V\left(|\phi\rangle \otimes \left|0^{\otimes m}\right\rangle\right) = VU|0^{\otimes(n'+m)}\rangle \overset{\text{def}}{=} W|0^{\otimes(n'+m)}\rangle \tag{A.20}$$

with $W = VU$ a circuit of minimum circuit depth $\overset{\text{def}}{=}$ depth$(W) \leq t + 2\ell^3$. Define the state obtained by incoherently measuring all the SMA qubits of $|\psi\rangle$ as

$$\Psi = \sum_{s\in\{0,1\}^m} D_s |\phi\rangle\langle\phi| D_s \otimes |s\rangle\langle s| \tag{A.21}$$

where we abuse notation slightly and use $D_s$ both as the eigenspace and the projector onto it. Since we assume that $C$ is a stabilizer code, the Hamiltonian terms $h_i$ all mutually commute, and therefore so do the measurements of the SMA qubits. Therefore, the order of measurement used is irrelevant.

Define the state $\Theta = \mathcal{E}(\Psi)$ obtained by applying the logical completely depolarizing channel $\mathcal{E}$ from eq. (A.18). Then, we have

$$\Theta = \sum_s \operatorname{tr}(D_s\phi D_s)\mu_s \otimes |s\rangle\langle s| \qquad \text{for } \mu_s \overset{\text{def}}{=} \mathcal{E}\left(\frac{D_s\phi D_s}{\operatorname{tr}(D_s\phi D_s)}\right). \tag{A.22}$$

**Claim A.8** *Fix any region $R \subset [n'+m]$. Let $S_R$ be the set of all indices $i \in [m]$ such that the ith SMA qubit belongs to $R$. It holds that*

$$F(\psi_R, \Psi_R) \geq 1 - \sum_{i\in S_R} \epsilon_i. \tag{A.23}$$

*Further, if $|R| < d$, then $\Psi_R = \Theta_R$.*

For every $j \in [n'+m]$, let $L_j$ be the support of the lightcone of $j$ with respect to the unitary $W^\dagger$. Note that $|L_j| \leq 2^{\text{depth}(W)} < d$. Since $W^\dagger|\psi\rangle$ is $|0\rangle^{\otimes(n'+m)}$, we have that for any qubit $j \in [n'+m]$,

$$\operatorname{tr}_{-\{j\}}\left(W^\dagger\psi W\right) = |0\rangle\langle 0|. \tag{A.24}$$

However, Fact 3.3 allows us to equate

$$\mathrm{tr}_{-\{j\}}\left(W^\dagger \psi W\right) = \mathrm{tr}_{-\{j\}}\left(W^\dagger (\psi_{L_j} \otimes \nu_{-L_j}) W\right), \tag{A.25a}$$

$$\mathrm{tr}_{-\{j\}}\left(W^\dagger \Theta W\right) = \mathrm{tr}_{-\{j\}}\left(W^\dagger (\Theta_{L_j} \otimes \nu_{-L_j}) W\right). \tag{A.25b}$$

Using eq. (A.23), we find that for all $j \in [n' + m]$,

$$\begin{aligned}
&F\left(\mathrm{tr}_{-\{j\}}\left(W^\dagger \psi W\right), \mathrm{tr}_{-\{j\}}\left(W^\dagger \Theta W\right)\right) && \text{(A.26a)}\\
&\quad = F\left(\mathrm{tr}_{-\{j\}}\left(W^\dagger (\psi_{L_j} \otimes \nu_{-L_j}) W\right), \mathrm{tr}_{-\{j\}}\left(W^\dagger (\Theta_{L_j} \otimes \nu_{-L_j}) W\right)\right) && \text{(A.26b)}\\
&\quad \geq F\left(\psi_{L_j}, \Theta_{L_j}\right) && \text{(A.26c)}\\
&\quad \geq 1 - \sum_{i \in S_{L_j}} \epsilon_i. && \text{(A.26d)}
\end{aligned}$$

We now infer from eq. (A.24) and eq. (A.26d) that[2]

$$S\left(\mathrm{tr}_{-\{j\}}\left(W^\dagger \Theta W\right)\right) \leq 2\left(\sum_{i \in S_{L_j}} \epsilon_i\right) \log \frac{1}{\min\left(\sum_{i \in S_{L_j}} \epsilon_i, \frac{1}{4}\right)}. \tag{A.27}$$

Using the concavity of the function $x \mapsto x \log \frac{1}{\min\left(x, \frac{1}{4}\right)}$ in the interval $x \in (0, 2^{\mathrm{depth}(W)})$, we can average over all $j \in [n' + m]$ to conclude

$$\begin{aligned}
&\mathop{\mathbf{E}}_{j \in [n'+m]} S\left(\mathrm{tr}_{-\{j\}}\left(W^\dagger \Theta W\right)\right) && \text{(A.28a)}\\
&\qquad \leq 2 \mathop{\mathbf{E}}_{j \in [n'+m]} \left(\left(\sum_{i \in S_{L_j}} \epsilon_i\right) \log \frac{1}{\min\left(\sum_{i \in S_{L_j}} \epsilon_i, \frac{1}{4}\right)}\right) && \text{(A.28b)}\\
&\qquad \leq 2 \cdot \left(\mathop{\mathbf{E}}_{j \in [n'+m]} \sum_{i \in S_{L_j}} \epsilon_i\right) \log \frac{1}{\min\left(\mathbf{E}_{j \in [n'+m]} \sum_{i \in S_{L_j}} \epsilon_i, \frac{1}{4}\right)}. && \text{(A.28c)}
\end{aligned}$$

The next claim helps upper and lower bound this expression.

**Claim A.9** *It holds that*

$$\frac{\epsilon m}{n' + m} \leq \mathop{\mathbf{E}}_{j \in [n'+m]} \sum_{i \in S_{L_j}} \epsilon_i \leq 2^{2\,\mathrm{depth}(W)} \frac{\epsilon m}{n' + m}. \tag{A.29}$$

[2] Given a binary distribution $(p, 1 - p)$, we can upper bound its entropy as follows. If $p \geq \frac{1}{4}$, then an upper bound is 1. Else the upper bound is $2p \log \frac{1}{p}$. The combined upper bound is $2p \log \frac{1}{\min(p, \frac{1}{4})}$.

We now upper bound the entropy of $\Theta$. For this, let us assume $2^{\mathrm{depth}(W)} \leq \frac{1}{\epsilon}$, else the proof is immediate.

$$S(\Theta) = S(W^\dagger \Theta W) \leq \sum_{j \in [n'+m]} S\left(\mathrm{tr}_{-\{j\}}\left(W^\dagger \Theta W\right)\right) \tag{A.30a}$$

$$\leq 2^{1+2\,\mathrm{depth}(W)} \epsilon m \log \frac{1}{\min\left(\frac{\epsilon m}{n'+m}, \frac{1}{4}\right)} \tag{A.30b}$$

$$\leq 2^{1+2\,\mathrm{depth}(W)} \epsilon \ell n \log \frac{2^{\mathrm{depth}(W)}}{\epsilon} \tag{A.30c}$$

$$\leq \left(2^{2+2\,\mathrm{depth}(W)} \ell n\right) \cdot \epsilon \log \frac{1}{\epsilon}. \tag{A.30d}$$

The inequality in eq. (A.30a) comes from the subadditivity of entropy; the inequality in eq. (A.30b) uses eq. (A.28b) and then substitutes the upper and lower bounds given in Claim A.9; the inequality in eq. (A.30c) uses $\frac{n}{\ell} \leq m \leq \ell n$ and $\frac{\epsilon m}{n'+m} \geq \frac{\epsilon}{\ell 2^t + 1} \geq \frac{\epsilon}{2^{\mathrm{depth}(W)}}$; the inequality in eq. (A.30d) uses $2^{\mathrm{depth}(W)} \leq \frac{1}{\epsilon}$. Furthermore, $\Theta$ is the output of $\mathcal{E}$ acting on $\Psi$. By Fact A.6 (Item 2), $S(\Theta) \geq k$. Combining the lower and upper bounds on the entropy of $\Theta$, the proof concludes. □

## A.2 A classical oracle separation between QMA and QCMA

*This section is based on [5] by Natarajan and Nirkhe.*

There are two natural *quantum* analogs of the computational complexity class NP. The first is the class QMA previously described and the second is the class QCMA in which the quantum polynomial-time decision algorithm is given access instead to a poly$(n)$ *bit classical* state. While it is easy to prove that QCMA $\subseteq$ QMA as the quantum witness state can be immediately measured to yield a classical witness string, the question of whether QCMA $\overset{?}{=}$ QMA, first posed by Aharonov and Naveh [20], remains unanswered. If QCMA = QMA, then every local Hamiltonian would have an efficient classical witness of its ground energy; morally, this can be thought of as an efficient classical description of its ground state. The relevance of local Hamiltonians to condensed matter physics makes this question a central open question in quantum complexity theory [85].

Because P $\subseteq$ QCMA $\subseteq$ QMA $\subseteq$ PSPACE, any unconditional separation of the two complexity classes would imply P $\neq$ PSPACE and seems unlikely without remarkably ingenious new tools. A more reasonable goal is an oracle separation between the two complexity classes. The first oracle separation, by Aaronson and Kuperberg [86], showed that there exists a black-box unitary problem for which quantum witnesses suffice and yet no polynomial sized classical witness and algorithm can solve the problem with even negligible success probability. A second black-box separation was discovered a decade later by Fefferman and Kimmel [87]. The Fefferman and Kimmel oracle is a completely positive trace perseving (CPTP) map called an "in-place" permutation oracle. Both oracles [86, 87] are inherently quantum[3]. Whereas, the "gold-standard" of oracle separations — namely black-box function separations (also known as classical oracle separations) — only require access to a *classical function* that can be queried in superposition[4].

### A.2.1 Graph oracles

The major result of [5] is that there exists a distribution over black-box function problems separating QMA and QCMA. Each black-box function corresponds to the adjacency list of a $N \overset{\text{def}}{=} 2^n$ vertex constant-degree colored graphs[5] $G = (V, E)$. Roughly speaking, a graph is a YES instance if the second eigenvalue of its normalized adjacency matrix is 1 (equivalently, if it has at least two connected components) and a graph is a NO instance if it second eigenvalue is at most $1 - \alpha$ for some fixed constant $\alpha$ (equivalently, the graph has one connected component and is expanding). We call this problem the *expander distinguishing problem*.

**Distribution oracles** A distribution over functions (equivalently, a distribution over graphs) is a YES instance if it is entirely supported on YES graphs and a distribution over functions is a NO instance if it is entirely supported on NO graphs.

In [5], we construct YES and NO distributions over graphs and an efficient algorithm with access to a quantum witness (an $n$-qubit state) can distinguish YES instances from NO instances. Furthermore, we prove that every query-efficient algorithm with access to even a super-polynomial length *classical* witness cannot distinguish YES instances from NO instances.

Our work is not the first to consider oracles that sample from distributions over functions. The in-place oracle separation of [87] between QMA and QCMA used oracles that sampled random permutations. For a somewhat different problem, of separating bounded-depth quantum-classical circuits, [89] introduced a related notion called a "stochastic oracle"—the main difference between this and our model is that a stochastic oracle resamples an instance every time it is queried.

**Comparison with previous oracle separations between QMA and QCMA** Figure A.1 summarizes our work in relation to previous oracle separations. In terms of results, we take a further step towards the standard oracle model—all that remains is to remove the randomness from our oracle. In terms of techniques, we combine the use of counting arguments and the adversary method from previous works with a BQP lower bound for a similar graph problem, due to [90]. This lower bound was shown using the polynomial method. We view the judicious combination of these lower bound techniques—as simple as it may seem—as one of the conceptual contributions of this paper.

[3] It might be reasonable to wonder if the unitary oracles can be converted into classical oracles by providing oracle access to the exponentially long classical descriptions of the respective matrices. This is not known to be true because it is unclear how to use access to the classical description to solve the QMA problem.

[4] One reason this model is natural is that if we were given a circuit of size $C$ to implement this classical function, then we would automatically get a quantum circuit of size $C$ to implement the oracle, simply by running the classical circuit coherently. This is not true for the "in-place" permutation oracle model, assuming that one-way functions exist.

[5] A similar problem was previously conjectured to be an oracle separation for these complexity classes by Lutomirski [88].

| Reference | Separating black box object | Proof techniques used |
|---|---|---|
| [86] | $n$-qubit unitaries | Adversary method |
| [87] | $n$-qubit CPTP maps | Combinatorial argument, Adversary method |
| Our work [5] | Distributions over $n$-bit boolean functions | Combinatorial argument, Adversary method, Polynomial method |
| Conjectured | $n$-bit boolean function | ? |

Figure A.1: List of known oracle separations

**Intuition for hardness** The expander distinguishing problem is a natural candidate for a separation between QMA and QCMA because it is an "oracular" version of the sparse Hamiltonian problem, which is complete for QMA [91, Problem H-4]. To see this, we recall some facts from spectral graph theory. The top eigenvalue of the normalized adjacency matrix $A$ for regular graphs is always 1 and the uniform superposition over vertices is always an associated eigenvector. If the graph is an expander (the NO case of our problem), the second eigenvalue is bounded away from 1, but if the graph is disconnected (the YES case of our problem), then the second eigenvalue is exactly 1. Thus, our oracle problem is exactly the problem of estimating the minimum eigenvalue of $\mathbb{I} - A$ (a sparse matrix for a constant-degree graph), on the subspace orthogonal to the uniform superposition state. Viewing $\mathbb{I} - A$ as a sparse Hamiltonian, we obtain the connection between our problem and the sparse Hamiltonian problem.

One reason to show oracle separations between two classes is to provide a *barrier* against attempts to collapse the classes in the "real" world. We interpret our results as confirming the intuition that any QCMA protocol for the sparse Hamiltonian must use more than just black-box access to entries of the Hamiltonian: it must use some nontrivial properties of the ground states of these Hamiltonians. In this sense, it emulates the original quantum adversary lower bound of [92] which showed that any BQP-algorithm for solving NP-complete problems must rely on some inherent structure of the NP-complete problem as BQP-algorithms cannot solve unconstrained search efficiently.

### A.2.2 Overview of proof techniques

**Quantum witnesses and containment in oracular** QMA A quantum witness for any YES instance graph is any eigenvector $|\xi\rangle$ of eigenvalue 1 that is orthogonal to the uniform superposition over vertices. The verification procedure is simple: project the witness into the subspace orthogonal to the uniform superposition over vertices, and then perform one step of a random walk along the graph, by querying the oracle for the adjacency matrix in superposition. Verify that the state after the walk step equals $|\xi\rangle$. This is equivalent to a 1-bit phase estimation of the eigenvalue. If a graph is a NO instance, then there does not exist any vector orthogonal to the uniform superposition (the unique eigenvector of value 1) that would pass the previous test.

Whenever, the graph has a connected component of $S \subsetneq V$, then an eigenvector orthogonal to the uniform superposition of eigenvalue 1 exists. When $|S| \ll N$, this eigenvector is very close to $|S\rangle$, the uniform superposition over basis vectors $x \in S$. Notice that this state only depends on the connected component $S$ and not the specific edges of the graph. Furthermore, the state $|S'\rangle$ for any subset $S'$ that approximates $S$ forms a witness that is accepted with high probability.

**Lower bound on classical witnesses** The difficulty in this problem lies in proving a *lower bound* on the ability for classical witnesses to distinguish YES and NO instances. To prove a lower bound, we argue that any quantum algorithm with access to a polynomial length classical witness must make an exponential number of (quantum) queries to the adjacency list of the graph in order to distinguish YES and NO instances. This, in turn, lower bounds the time complexity of any QCMA algorithm distinguishing YES and NO instances but is actually slightly stronger since we don't consider the computational complexity of the algorithm between queries.

Proving lower bounds when classical witnesses are involved is difficult because the witness could be based on any property of the graph. For example, the classical witness could describe cycles, triangles, etc. contained in the graph — while it isn't obvious why such a witness would be helpful, proving that any such witness is insufficient is a significant challenge. One way to circumvent this difficulty is to first show a lower bound *assuming* some structure about the witness[6], and then "remove the training wheels" by showing that the assumption holds for any good classical witness.

**Lower bound against "subset witnesses"** One structure we can assume is that the witness only depends on the set of vertices contained in the connected component $S$. This is certainly the case for the quantum witness state. Our result shows that any polynomial-length witness only depending on the vertices in $S$ requires an exponential query complexity to distinguish YES and NO instance graphs.

The starting point for this statement is the exponential query lower bound *in the absence of a witness* (i.e. for BQP) for the expander distinguishing problem proven by Ambainis, Childs and Liu [90], using the polynomial method. In [90], the authors define two distributions over constant-degree regular colored graphs: the first is a distribution $P_1$ over random graphs with overwhelming probability of having a second normalized eigenvalue at most $1 - \epsilon_0$. The second is a distribution $P_\ell$ over random graphs with overwhelming probability of having $\ell$ connected components. Since, almost all graphs in $P_1$ are NO graphs and all graphs in $P_\ell$ are YES graphs, any algorithm distinguishing YES and NO instances must be able to distinguish the two distributions. We first show that a comparable query lower bound still holds even when the algorithm is given a witness consisting of polynomially many random points $F$ from any one connected component.

Next, we show that if there were a QCMA algorithm where the optimal witness depends only on the set of vertices $S$ in one of the connected components, by a counting argument, there must exist a combinatorial *sunflower* of subsets $S$ that correspond to the same witness string. A *sunflower*,

[6] Assuming structure about a witness is a common technique in theoretical computer science and in particular lower bounds for classical witnesses of quantum statements. For example, lower bounds against natural proofs [93]. Another example is the NLTS statement [1] which is about lower bounds for classical witnesses for the ground energy of a quantum Hamiltonian of a particular form: constant-depth quantum circuits.

in this context, is a set of subsets such that each subset contains a core $F \subset V$ and every vertex of $V \setminus F$ occurs in a small fraction of subsets. This implies that there exists a BQP algorithm which distinguishes YES instances corresponding to the sunflower from all NO instances. Next, we show using an adversary bound [94], a quantum query algorithm cannot distinguish the distribution of YES instances corresponding to the sunflower from the uniform distribution of YES instances such that the core $F$ is contained in a connected component (the ideal sunflower).

This indistinguishability, along with the previous polynomial method based lower bound, proves that QCMA algorithm — whose witness only depends on the vertices in the connected component — for the expander distinguishing problem must make an exponential number of queries to the graph.

**Removing the restriction over witnesses** Our proof, thus far, has required the restriction that the witness only depends on the vertices in the connected component. In some sense, this argues that there is an oracle separation between QMA and QCMA if the prover is restricted to being "near-sighted": it cannot see the intricacies of the edge-structure of the graph, but can notice the separate connected components of the graph. If the near-sighted prover was capable of sending quantum states as witnesses, then she can still aid a verifier in deciding the expander distinguishing problem, whereas if she could only send classical witnesses, then she cannot aid a verifier.

It now remains to remove the restriction that the witness can only depend on the vertices in the connected component. We do this by introducing *randomness* into the oracle, precisely designed to "blind" the prover to the local structure of the graph. In the standard oracle setting, the verifier and prover both get access to an oracle $x \in \{0,1\}^N$, and the prover provides either a quantum witness, $|\xi(x)\rangle \in (\mathbb{C}^2)^{\otimes \mathrm{poly}(n)}$ or a classical witness, $\xi(x) \in \{0,1\}^{\mathrm{poly}(n)}$. The verifier then runs an efficient quantum algorithm $V^x$ which takes as input $|\xi(x)\rangle$ (or $\xi(x)$, respectively) and consists of quantum oracle gates applying the unitary transform defined as the linear extension of

$$|i\rangle \mapsto (-1)^{x_i} |i\rangle \text{ for } i \in [N]. \tag{A.31}$$

We now extend modify this setup slightly. Instead of a single oracle $x$, we consider a distribution $\mathcal{B}$ over oracles. The prover constructs a quantum witness $|\xi(\mathcal{B})\rangle$ (or a classical witness $\xi(\mathcal{B})$, respectively) based on the distribution $\mathcal{B}$. The verifier then samples a classical oracle $x \leftarrow \mathcal{B}$ from the distribution, and then runs the verification procedure $V^x$ which takes as input $|\xi(\mathcal{B})\rangle$ (or $\xi(\mathcal{B})$, respectively) and applies quantum oracle gates corresponding to $x$. The success probability of the verifier is taken over the distribution $\mathcal{B}$ and the randomness in the verification procedure.

From our previous observations, graphs with the same connected component $S$ have the same ideal witness state. So, if the distribution $\mathcal{B}$ is supported on all graphs with the same connected component $S$, then the witness state corresponding to $S$ will suffice. Furthermore, in the case of the classical witness system, the witness can only depend on $S$ and the previously stated lower bound applies. This motivates the oracle problem of distinguishing distributions, marked either YES or NO, over $2^n$ bit strings (or equivalently $n$-bit functions).

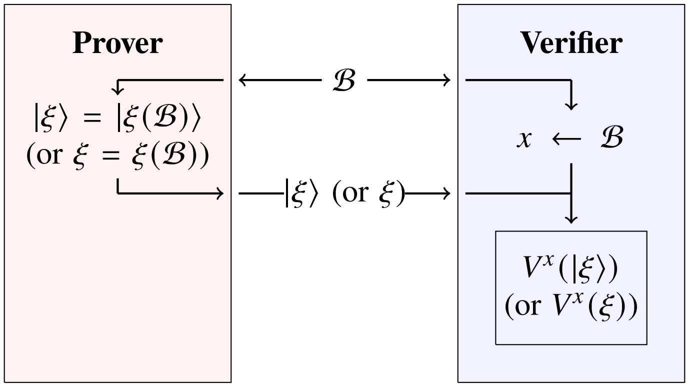
Prover
$|\xi\rangle = |\xi(\mathcal{B})\rangle$
(or $\xi = \xi(\mathcal{B})$)
$\mathcal{B}$
$|\xi\rangle$ (or $\xi$)
Verifier
$x \leftarrow \mathcal{B}$
$V^x(|\xi\rangle)$
(or $V^x(\xi)$)

Figure A.2: *C*artoon of interaction between Prover and Verifier for a distribution over classical boolean functions.

### A.2.3 Statement of the result

**Theorem A.10 (Classical oracle separation [5])** *For every sufficiently large integer $n$ that is a multiple of 200, there exist distributions over* 100*-regular* 100*-colored graphs on $N = 2^n$ vertices labeled either* YES *or* NO *such that*

- *Each* YES *distribution is entirely supported on* YES *instances of the expander-distinguishing problem and, likewise, each* NO *distribution is entirely supported on* NO *instance of the expander-distinguishing problem.*

- *There exists a* poly$(n)$ *time quantum algorithm $V_q$ taking a witness state $|\xi\rangle$ as input and making $O(1)$ queries to the quantum oracle such that*

  1. *For every* YES *distribution $\mathcal{B}$, there exists a quantum witness $|\xi\rangle \in (\mathbb{C}^2)^{\otimes n}$ such that*

  $$\mathop{\mathbf{E}}_{x \leftarrow \mathcal{B}} \mathbf{Pr}[V_q^x(|\xi\rangle) \textit{ accepts}] \geq 1 - O(N^{-3}). \quad \text{(A.32)}$$

  2. *For every* NO *distribution $\mathcal{B}$, for all quantum witnesses $|\xi\rangle \in (\mathbb{C}^2)^{\otimes n}$,*

  $$\mathop{\mathbf{E}}_{x \leftarrow \mathcal{B}} \mathbf{Pr}[V_q^x(|\xi\rangle) \textit{ accepts}] \leq 0.01. \quad \text{(A.33)}$$

- *Any quantum algorithm $V_c$ accepting a classical witness of length $q(n)$ satisfying the following two criteria either requires $q(n)$ to be exponential or must make an exponential number of queries to the oracle.*

  1. *For every* YES *distribution $\mathcal{B}$, there exists a classical witness $\xi = \xi(\mathcal{B}) \in \{0, 1\}^{q(n)}$*

  $$\mathop{\mathbf{E}}_{x \leftarrow \mathcal{B}} \mathbf{Pr}[V_c^x(\xi) \textit{ accepts}] \geq 0.99. \quad \text{(A.34)}$$

2. *For every* NO *distribution* $\mathcal{B}$, *for all classical witnesses* $\xi \in \{0,1\}^{q(n)}$,

$$\mathop{\mathbf{E}}_{x \leftarrow \mathcal{B}} \mathbf{Pr}[V_c^x(\xi) \textit{ accepts}] \leq 0.01. \tag{A.35}$$

Although our main theorem is formulated as a query lower bound, it can be converted to a separation between the relativized classes of QMA and QCMA via a standard diagonalization argument. Similarly, it was pointed out to us [95] that it proves a separation between the relativized classes of BQP/qpoly and BQP/poly, following the technique of [86].

### A.2.4 Implications and future directions

There are several future questions raised by [5] that we find interesting:

**Oracle and communication separations** The most natural question is of course whether the randomness in the oracle can be removed, to obtain a separation in the standard model. We conjecture that our problem yields such a separation, but a new technique seems necessary to prove it.

Another natural question is to show a *communication complexity* separation between QMA and QCMA. This has been shown for one-way communication by Klauck and Podder [96] but their problem does not yield a separation for two-way communication. Could our query separation be lifted to the communication world by use of the appropriate gadget?

The class QMA(2) is another relative of QMA which is perhaps even more enigmatic than QCMA. In QMA(2), the witness state is promised to be an unentangled between the first and second half of the qubits. We do not even know of a quantum (unitary) oracle separation between QMA(2) and QMA, nor do we have a natural candidate problem. Could we at least formulate such a candidate by considering "oracular" versions of QMA(2)-complete problems, in analogy to what we do in [5] for QCMA.

**Implications for Quantum PCPs** In a recent panel [97] on the quantum PCP conjecture and the NLTS theorem [1], an interesting question was posed of whether MA or QCMA (lower or upper) bounds can be placed on the complexity of the promise-gapped local Hamiltonian problem. Because the oracle presented in this result corresponds to a sparse Hamiltonian with a problem of deciding if the second eigenvalue of the Hamiltonian is 1 or $< 1-\alpha/d = 1-\Omega(1)$, one might wonder if this provides oracular evidence that quantum PCPs are at least QCMA-hard. Unfortunately, to the best of our knowledge, this is not a reasonable conclusion. While we give evidence that the promise-gapped *sparse* Hamiltonian problem is likely QCMA-hard, the reduction from the sparse Hamiltonian problem to the local Hamiltonian problem does not imply that the promise-gapped local Hamiltonian problem is likely QCMA-hard. The only algorithm known for checking a witness for the sparse Hamiltonian problem is applying Hamiltonian simulation on the witness; this is not a local algorithm.

**Connections to stoquastic Hamiltonians** Since the oracles studied in [5] correspond to the adjacency lists of graphs, they can be viewed as sparse access to a Hamiltonian $\mathbf{H}$ which is the Laplacian of a graph (recall that if the adjacency matrix is $A$, then the Laplacian is $\mathbb{I} - A/d$). Such Hamiltonians have a special structure not present in general Hamiltonians: they are *stoquastic*, meaning that the off-diagonal entries are nonpositive. The local Hamiltonian (LH) problem for stoquastic Hamiltonians is significantly easier than the general LH problem, and in some cases is even contained in $\mathsf{MA}$ as shown by Bravyi and Terhal [98]. It is worth noticing why this is not in tension with our result—in particular, why this does not imply that our oracle problem is contained in oracular $\mathsf{MA}$.

- Crucially, the $\mathsf{MA}$-containment for stoquastic LH holds *only* for the ground state: this is because of the Perron-Frobenius theorem, which implies that ground states of such Hamiltonians have nonnegative coefficients. However, in our case, we want the first excited state: the state of minimum energy for $\mathbf{H}$ restricted to the subspace orthogonal to the uniform superposition. It was shown by [99] that all excited state energies are $\mathsf{QMA}$-hard to calculate for a stoquastic Hamiltonian.
- The $\mathsf{MA}$ containment also uses the locality of the Hamiltonian, which in turn imposes a strong structure on the adjacency matrix of the graph. The random graphs we consider will not have this structure. (While it was shown by [100] showed an $\mathsf{AM}$ algorithm for calculating the ground energy *stoquastic and sparse* Hamiltonians, again this does not apply to higher excited states.)
- At an intuitive level, in graph language, the LH problem for stoquastic Hamiltonians is to find a component of the graph where the average value of some *potential function* (given by the diagonal entries of $\mathbf{H}$) is minimized. An $\mathsf{MA}$ verifier can solve this by executing a random walk, given the right starting point by Merlin. In contrast, our problem is to determine whether the graph as a whole is connected—a global property that an $\mathsf{MA}$ verifier cannot determine.

## A.3 Quantum search-to-decision and state synthesis

*This section is based on [4] by Irani, Natarajan, Nirkhe, Rao, and Yuen.*

It is a useful fact in classical computer science that *search* problems are often efficiently reducible to *decision* problems. For example, the canonical way of constructing a satisfying assignment of a given 3SAT formula $\varphi$ (if there exists one) using an oracle for the decision version of 3SAT is to adaptively query the oracle for the satisfiability of $\varphi$ conditioned on some partial assignment to the variables of the formula. Based on the oracle answers, the partial assignment can be extended bit-by-bit to a full assignment. Each oracle query reveals an additional bit of the assignment. This strategy generally works for any problem in $\mathsf{NP}$. Likewise, the optimal value of an optimization problem can be calculated to exponential accuracy using binary search. The main consequence of this is that complexity theory often focuses on decision problems (without losing generality) and less on the complexity of search problems.

Quantum information and computation have shifted our perspective on these traditional notions of classical complexity theory. We now consider *quantum* search problems, where the goal is to output a quantum state (as opposed to a classical bit string) satisfying some condition. In the quantum setting, it is no longer apparent that search-to-decision reductions still hold, and thus it is unclear whether the complexity of quantum search problems can be directly related to the complexity of corresponding quantum decision problems.

Is there an efficient search-to-decision reduction for the Local Hamiltonians problem, or more generally for the class QMA? In other words, given quantum query access to an oracle deciding the Local Hamiltonians problem, can a polynomial-time quantum algorithm (i.e. BQP machine) efficiently *prepare* a low-energy state $|\psi\rangle$ of a given local Hamiltonian?

The classical strategy of incrementally building a partial assignment does not appear to work in the QMA setting. First, there does not appear to be a natural way of "conditioning" a quantum state on a partial assignment. Second, quantum states are exponentially complex: the description size (complexity) of a general quantum state on $n$ qubits is exponential in $n$, and this is suspected to remain true even when considering ground states of local Hamiltonians[7]. This complexity of quantum states poses a significant challenge to finding a search-to-decision reduction for QMA; it is not clear how yes/no answers to QMA decision problems (even when obtained in superposition) can be used to construct exponentially-complex QMA witnesses.

On the other hand, there *is* a natural quantum analog of the bit-by-bit search-to-decision algorithm for NP that works for constructing *general* quantum states. This is due to a general algorithm for *state synthesis* described by Aaronson in [101] (for which we give an overview of in Appendix A.3.1): there exists a polynomial-time quantum algorithm $A$ such that every $n$-qubit state $|\psi\rangle$ can be encoded into a classical oracle $f$ where, by making $O(n)$ superposition queries to the oracle $f$, the algorithm $A$ will output a state that is exponentially close to $|\psi\rangle$. One can observe that for states $|\psi\rangle$ that QMA witnesses (such as ground states of local Hamiltonians), the oracle $f$ corresponds to a PP function (which is at least as powerful as a QMA oracle). This yields a search-to-decision reduction for QMA, albeit with a decision oracle of higher complexity.

We explore the complexity of search-to-decision procedures in the quantum setting, where the goal is a quantum *state synthesis* algorithm that outputs a *target* quantum state (e.g. a ground state of a local Hamiltonian) by making quantum queries to a classical decision oracle. We investigate how the complexity of the state synthesis algorithm and the complexity of the decision oracle depend on the type of states we want to generate. We consider both the generalized state synthesis problem for arbitrary states in the Hilbert space $(\mathbb{C}^2)^{\otimes n}$ as well as the specific task of generating solutions to QMA problems.

We construct state synthesis and search-to-decision procedures for the quantum setting using only one or two superposition queries as opposed to $O(n)$ superposition queries; for QMA witnesses, the synthesis procedure requires only one query to a PP oracle. Simultaneously, we prove results

[7] Due to the QMA $\neq$ QCMA conjecture [86]. Formally, there is no known poly-sized description of a witness (proof) for every local Hamiltonian problem.

suggesting the impossibility of any search-to-decision reduction for QMA. More precisely, there exists a *quantum* oracle $\mathcal{O}$ relative to which *all* efficient query algorithms fail to be a good search-to-decision reduction for $\mathsf{QMA}^{\mathcal{O}}$, the relativization of QMA. This stands in contrast to classes such as NP, MA, and QCMA, which all have efficient search to decision reductions, relative to any oracle. As a consequence, proving the impossibility of QMA search-to-decision without an oracle is at least as hard as separating QCMA and QMA which is at least as hard as separating P and PP. We believe that the juxtaposition of our results lends further weight to the view that the complexity of tasks where the outputs (and inputs) are quantum states cannot be directly explained by the traditional study of decision problems (which has been the main focus of quantum complexity theory to date). In particular, we believe our results suggest that the relationship between search and decision problems is much more mysterious in the quantum setting. As suggested by Aaronson in [101] and others in some recent works [102, 103], the complexity of quantum states (and more generally, quantum state transformations) deserves to be studied more deeply as a subject in its own right.

### A.3.1 Starting point

Before describing our results in more detail, we first explain the starting point for our investigations, which is a simple state synthesis algorithm described by Aaronson [101] in his lecture notes. He shows that there exists a $\mathrm{poly}(n)$-time quantum algorithm $A$ which makes $O(n)$ quantum queries to a classical oracle such that for every $n$-qubit state $|\psi\rangle = \sum_x \alpha_x |x\rangle$, there exists a classical oracle $f$ for which the algorithm $A^{O_f}$ will output a state that is $\exp(-n)$-close to $|\psi\rangle$. In [101], Aaronson raises the question as to whether his protocol can be improved to a sublinear number of queries. That 1 query is sufficient to achieve a polynomially small error in synthesizing arbitrary states and 2-queries are sufficient for an exponentially small error. Both the 1-query and the 2-query algorithms given here require exponential time and polynomial space.

To understand Aaronson's state synthesis algorithm, we first observe that we can write any quantum state in the form

$$|\psi\rangle = \sum_{x\in\{0,1\}^n} e^{i\theta_x} \sqrt{\mathbf{Pr}[X = x]}\; |x\rangle \tag{A.36}$$

where $\mathbf{Pr}[X = x]$ is the probability distribution of some $n$-bit random variable $X$ and $\{\theta_x\}_{\{0,1\}^n}$ are a set of phases. The synthesis algorithm performs $2n$ queries to synthesize the "QSample state".

$$\sum_{x\in\{0,1\}^n} \sqrt{\mathbf{Pr}[X = x]}\; |x\rangle \tag{A.37}$$

and then performs two additional queries at the end to apply the phases $e^{i\theta_x}$ to each basis state $|x\rangle$.

The $2n$-query procedure to build the QSample state works in $n$ stages. Inductively assume that after the $k$th stage, for $k < n$, the intermediate state of the algorithm is the $k$-qubit state

$$\sum_{y\in\{0,1\}^k} \sqrt{\mathbf{Pr}[X_{\le k} = y]}\; |y\rangle \tag{A.38}$$

where $\mathbf{Pr}[X_{\le k} = y]$ denotes the marginal probability of the first $k$ bits of $X$ are equal to $y$. Controlled on the prefix $|y\rangle$ the algorithm queries the oracle $f$ to obtain a (classical description of) the conditional probabilities $\mathbf{Pr}[X_{k+1} = 0 \mid X_{\le k} = y]$ and $\mathbf{Pr}[X_{k+1} = 1 \mid X_{\le k} = y]$, and prepares a $(k+1)$st qubit in the state

$$\sqrt{\mathbf{Pr}[X_{k+1} = 0 \mid X_{\le k} = y]}\ |0\rangle + \sqrt{\mathbf{Pr}[X_{k+1} = 1 \mid X_{\le k} = y]}\ |1\rangle\ . \tag{A.39}$$

The algorithm performs another query to $f$ to uncompute the descriptions of the conditional probabilities. The resulting $k+1$ qubit state is then equal to

$$\sum_{y\in\{0,1\}^{k+1}} \sqrt{\mathbf{Pr}[X_{\le k} = y_{\le k}]} \cdot \sqrt{\mathbf{Pr}[X_{k+1} = y_{k+1} \mid X_{\le k} = y_{\le k}]}\ |y\rangle \tag{A.40}$$

$$= \sum_{y\in\{0,1\}^{k+1}} \sqrt{\mathbf{Pr}[X_{\le k+1} = y]}\ |y\rangle \tag{A.41}$$

which maintains the desired invariant. After the $n$th stage, a similar process applies the phases $\{\theta_x\}$ to generate the output state. The approximations come in when the conditional probabilities and phases are specified with $\mathrm{poly}(n)$ bits of precision, which result in the final state being at most $\exp(-n)$ far from the ideal target state $|\psi\rangle$. With this $O(n)$-query state synthesis algorithm in mind, we now proceed to describe our results.

### A.3.2 Results

**A one-query search-to-decision algorithm for QMA with a PP oracle.** In the case of generating physically relevant states, i.e. solutions to QMA problems, such as the low-energy states of local Hamiltonians, there exists a one-query search-to-decision algorithm using a PP oracle. While one would hope to find a search-to-decision reduction in which the oracle complexity is only QMA, PP is the smallest complexity class containing QMA for which we can construct an oracular algorithm for search problems. Furthermore, given our no-go result for QMA search-to-decision (see below), this may be the optimal search-to-decision algorithm.

**Theorem A.11 (QMA-search to PP-decision reduction [4])** *There exists a probabilistic polynomial time quantum algorithm making a single query to a PP phase oracle such that, given as input a QMA problem, either aborts or outputs a witness $|\phi\rangle$. The algorithm will succeed in outputting a witness (i.e. not abort) with all but inverse exponential (in the system size) probability.*

To start sketching the proof, it is fruitful to notice that a single oracle query $|x\rangle \overset{O_f}{\mapsto} (-1)^{f(x)}\,|x\rangle$ for $x \in \{0,1\}^n$ potentially contains $2^n$ bits of information and a quantum state requires $2^n$ complex numbers to describe. Furthermore, the collection of $2^{2^n}$ states

$$|p_f\rangle \overset{\text{def}}{=} O_f H^{\otimes n}\,|0^n\rangle = \sum_{x\in\{0,1\}^n} (-1)^{f(x)}\,|x\rangle \tag{A.42}$$

defined for any function $f : \{0, 1\}^n \to \{0, 1\}$ are a diverse set of states in the Hilbert space. These states, referred to as *phase states* henceforth, despite not forming an $\epsilon$-net for $(\mathbb{C}^2)^{\otimes n}$, turn out to provide a good approximation for $(\mathbb{C}^2)^{\otimes n}$ when considering the Haar-random distribution[8]. It follows that if we wanted to synthesize the witness to a QMA-complete problem, such as a low-energy state $|\tau\rangle$ for a local Hamiltonian problem, it suffices to build phase state $|p_f\rangle$ with constant overlap with the low-energy subspace. Finding a state with constant overlap with the target state is sufficient because QMA is efficiently verifiable, and given a state with constant overlap with the low-energy subspace, it is possible to distill a low-energy state with constant probability (by performing an energy measurement). However, it is not necessarily the case that a low-energy state of the QMA problem will have a good approximation by a phase state. To solve this issue, we prove that for any state $|\tau\rangle$, with high probability $C|\tau\rangle$ will have a good approximation by a phase state where $C$ is a random Clifford unitary. Therefore, we can instead attempt to synthesize $C|\tau\rangle$ which is the result of Theorem A.11. In particular, if we can synthesize a phase state $|p\rangle$ that has constant overlap with $C|\tau\rangle$, then $C^\dagger|p\rangle$ will have constant overlap with the target $|\tau\rangle$.

Furthermore, using a slight modification of the same algorithm, we can perform a somewhat weaker one-query search-to-decision reduction for $\mathsf{QMA}_{\text{exp}}$, the class of non - deterministic quantum computations with only an inverse exponential gap between completeness and soundness. $\mathsf{QMA}_{\text{exp}}$ is known to equal PSPACE [104, 105], and our algorithm prepares a witness state with constant overlap with a low-energy state with one query to a PSPACE oracle (note that here, we cannot efficiently amplify the overlap with an energy measurement due to the inverse-exponential energy gap). As a further observation, we also show that quantum query access to a classical oracle gives one-query search-to-decision reductions when the witness is classical: in particular, for QCMA and NP. The one-query algorithm preparing the witness first reduces QCMA to *unique* QCMA (UQCMA) using the Valiant-Vazirani reduction [106] and then uses the Bernstein-Vazirani algorithm to extract the unique polynomial length witness with a single query.

**A no-go result for search-to-decision for QMA.** The previous result shows that search-to-decision reductions for QMA are possible with a PP decision oracle. However, the optimal search-to-decision reduction for QMA is with a QMA decision oracle (rather than a stronger PP oracle). We provide evidence that this is unlikely to exist: we prove that there is a quantum oracle relative to which QMA search-to-decision is impossible. This stands in contrast to classes such as NP, MA, and QCMA, which all have efficient search to decision reductions, relative to any oracle.

More precisely, there exists a *quantum* oracle $\mathcal{O}$ relative to which *all* efficient query algorithms fail to be a good search-to-decision reduction for $\mathsf{QMA}^{\mathcal{O}}$, the relativization of QMA. The oracle $\mathcal{O}$ is a reflection $\mathbb{I} - 2|\psi\rangle\langle\psi|$ about a Haar-random state $|\psi\rangle$; we rely on the concentration of measure phenomenon of the Haar measure to prove this oracle no-go result.

[8] Recall, the Haar-measure is the unique left- and right- invariant distribution over unitary matrices over $(\mathbb{C}^2)^{\otimes n}$ and the Haar-random distribution is the distribution over quantum states $U|0^n\rangle$ where $U$ is sampled according to the Haar-measure.

**Theorem A.12 (Oracle impossibility for $\mathsf{QMA}$ search-to-decision [4])** *There exists a quantum oracle $\mathcal{O}$ relative to which* all $\mathrm{poly}(n)$*-time query algorithms fail to be a good search-to-decision reduction for* $\mathsf{QMA}^{\mathcal{O}}$.

The proof of this theorem uses an oracle identical to that of Aaronson and Kuperberg [86] used to separate $\mathsf{QMA}$ and $\mathsf{QCMA}$. We conjecture that this is a noncoincidence and that *any* $\mathsf{QMA}$ and $\mathsf{QCMA}$ separating oracle yields a $\mathsf{QMA}$ search-to-decision impossibility result. Similar to the reasons for why the gold-standard of oracle separation between $\mathsf{QMA}$ and $\mathsf{QCMA}$ is a $n$-bit boolean function, the ideal oracle for proving $\mathsf{QMA}$ search-to-decision impossibility is also a $n$-bit boolean function. Does the oracle presented in Section A.2 also yield a search-to-decision impossibility?

**Open questions.** What is the power of a $\mathsf{QMA}$ decision oracle? In particular, what states can be synthesized with queries to a $\mathsf{QMA}$ oracle in superposition? Is there a weaker oracle class than $\mathsf{PP}$ that can achieve search-to-decision for $\mathsf{QMA}$ witnesses?