

Quantum Stochastic Processes and Quantum Many-Body Physics



Johannes Karl Richard Bausch

June 2017

St John's College
University of Cambridge

This dissertation is submitted for the degree of
Doctor of Philosophy

Für meine Familie.

Quantum Stochastic Processes and Quantum Many-Body Physics

Johannes Karl Richard Bausch

This dissertation investigates the theory of quantum stochastic processes and its applications in quantum many-body physics. The main goal is to analyse complexity-theoretic aspects of both static and dynamic properties of physical systems modelled by quantum stochastic processes. The thesis consists of two parts: the first one addresses the computational complexity of certain quantum and classical divisibility questions, whereas the second one addresses the topic of Hamiltonian complexity theory.

In the divisibility part, we discuss the question whether one can efficiently sub-divide a map describing the evolution of a system in a noisy environment, i.e. a CPTP- or stochastic map for quantum and classical processes, respectively, and we prove that taking the n^{th} root of a CPTP or stochastic map is an NP-complete problem. Furthermore, we show that answering the question whether one can divide up a random variable X into a sum of n iid random variables Y_i , i.e. $X = \sum_{i=1}^n Y_i$, is poly-time computable; relaxing the iid condition renders the problem NP-hard.

In the local Hamiltonian part, we study computation embedded into the ground state of a many-body quantum system, going beyond “history state” constructions with a linear clock. We first develop a series of mathematical techniques which allow us to study the energy spectrum of the resulting Hamiltonian, and extend classical string rewriting to the quantum setting. This allows us to construct the most physically-realistic QMA_{EXP} -complete instances for the LOCAL HAMILTONIAN problem (i.e. the question of estimating the ground state energy of a quantum many-body system) known to date, both in one- and three dimensions. Furthermore, we study weighted versions of linear history state constructions, allowing us to obtain tight lower and upper bounds on the promise gap of the LOCAL HAMILTONIAN problem in various cases. We finally study a classical embedding of a Busy Beaver Turing Machine into a low-dimensional lattice spin model, which allows us to dictate a transition from a purely classical phase to a Toric Code phase at arbitrarily large and potentially even uncomputable system sizes.

Declaration

This dissertation is the result of my own work and includes nothing which is the outcome of work done in collaboration except as declared in the Preface and specified in the text.

It is not substantially the same as any that I have submitted, or, is being concurrently submitted for a degree or diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the Preface and specified in the text. I further state that no substantial part of my dissertation has already been submitted, or, is being concurrently submitted for any such degree, diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the Preface and specified in the text.

It does not exceed the prescribed word limit for the relevant Degree Committee.

Acknowledgements

First and foremost I would like to express my deepest gratitude to my supervisor, Toby Cubitt, for guiding me through my PhD. Coming from a high-energy background, I had to learn almost all of quantum computation, quantum information and complexity theory from scratch. Toby's invaluable insights and excitement for our field brought me up to speed in no time, and it's through him that I learned to appreciate what our research has to offer. I'm especially grateful for our endless discussions and the time he invested in answering my countless questions. I would also like to thank the Department of Pure and Applied Mathematics and Theoretical Physics, as well as our group, CQIF, and especially Richard Jozsa, for hosting my research, and for providing such an encouraging workplace. Furthermore, I would like to thank St. John's College for offering an excellent social hub, where I could always connect with friends after a long day of study. I'm immensely grateful for the hospitality that I've encountered during my stay there. I'm also in debt to the various funding bodies which enabled me to pursue my doctoral studies in Cambridge: foremost the German National Academic Foundation and the EPSRC, as well as my college, which also enabled me to attend numerous conferences and workshops all over the world.

My research would not have been possible without the tremendous efforts of my collaborators: Toby Cubitt, Māris Ozols, Angelo Lucia, David Pérez-Garcia, Michael Wolff, Stephen Piddock, and Elizabeth Crosson. I feel honoured to have worked and learned from them, and hope for many more joint projects in the future. There are many more people that I've had the pleasure to discuss research questions with, and to whom I'd like to express my gratitude: Steve Brierley, Nilanjana Datta, András Gilyén, Carlos Guillén, Felix Leditzky, Josh Lockart, Will Matthews, Graeme Mitchison, Jenish Mehta, Imdad Sardharwalla, Sergii Strelchuk, and Thomas Vidick. Special thanks goes to Richard Jozsa and Aram Harrow who have kindly agreed to examine my thesis, and for their excellent questions raised during the viva. Finally, I want to express my gratitude towards those who have shown me how beautiful and fun physics and mathematics can be; foremost my thanks goes to Prof. Manfred Salmhofer, Prof. Eberhard Freitag, Prof. Matthias Bartelmann, and my tutor Kambis Veshgini in Heidelberg, and Prof. Liam McAllister at Cornell University.

I'm extremely lucky to have shared my time in Cambridge with so many great people and friends. Above all, I want to say thank you to Jean Maillard, Sandro Bauer, and Conrad Koziol for being the best company and support one could possibly hope for; this also goes for my other fellow Johnians Mike Keebler, Chris Hooton, Mubeen Goolam, Laura Keating, and Natacha Crooks, as well as all the other great people whom I only got to know over the last few months. Moreover, I want to thank all my friends from home, who have held contact despite my living away numerous hours on a plane: Sebastian Gast, Michael Wiedemann, Timo Besenreuther, Claude Huober, Wolfgang Mack, Andreas Müller, Daniel Zeifang, and Laura Mangold, as well as everyone else I forgot at this point. Cambridge is a special place with people that have grown close to me, and I am especially grateful for having been taught that true stories seldom take the straightest way.

Last but not least I want to express my gratitude to my family, my parents Jutta and Carl-Erich, my sister Elena, and my grandmother Johanna, who have always been there for me with encouraging words and support. Thank you!

Contents

Introduction	xi
1 Many-Body Quantum Physics	xi
2 Quantum Stochastic Processes	xv
3 Complexity Theory	xx
4 Putting it all Together	xxiv
5 Structure of Thesis	xxix
6 Authors and Contribution	xxx
1 Divisibility	I
1.1 CPTP and Stochastic Matrix Divisibility	4
1.1.1 Preliminaries	5
1.1.2 NP-Toolbox	8
1.1.3 Equivalence of Computational Questions	11
1.1.4 Reduction of STOCHASTIC ROOT to CPTP ROOT	13
1.1.5 Reduction of NONNEGATIVE ROOT to STOCHASTIC ROOT	14
1.1.6 Reduction of 1-IN-3SAT to NONNEGATIVE ROOT	16
1.1.7 Orthonormalisation and Handling the Unwanted Inequalities	18
1.1.8 Lifting Singularities	20
1.1.9 Complete Embedding	21
1.1.10 Bit Complexity of Embedding	22
1.2 Distribution Divisibility	24
1.2.1 Preliminaries	25
1.2.2 Equivalence to Polynomial Factorisation	28
1.2.3 Divisibility	29
1.2.4 Decomposability	34
1.3 Chapter Summary	45

2	Hamiltonian Complexity: Turing’s Wheelbarrow	47
2.1	Extended Introduction and Overview of Results	49
2.1.1	Historical Context	49
2.1.2	Main Result	51
2.1.3	Proof Ideas and Techniques	53
2.1.4	Structure of the Chapter	60
2.2	Turing’s Wheelbarrow	61
2.2.1	Preliminaries	61
2.2.2	Quantum Ring Machine	71
2.2.3	Unitary Labelled Graphs	78
2.2.4	Quantum Thue Systems	84
2.2.5	Hardness Result	89
2.2.6	Turing’s Wheelbarrow	92
2.2.7	Final Dimension Reduction	107
2.3	Chapter Summary	108
3	Hamiltonian Complexity: Lattice Crystals	111
3.1	Overview of Results	112
3.1.1	History State Construction	113
3.1.2	Tiling Construction	115
3.1.3	Hard Instances for the LOCAL HAMILTONIAN Problem	116
3.2	Turing’s Cube	119
3.2.1	Single Gate Universality	119
3.2.2	Circuit Encoding	120
3.2.3	Static Lattice Constraints	124
3.2.4	Dynamic Lattice Constraints	129
3.2.5	Analysis of History State Branching	134
3.3	QMA_{EXP} -Hardness Proof	135
3.4	Chapter Summary	139

4	Size-Driven Quantum Phase Transitions	141
4.1	Preliminaries	143
4.1.1	Embedding a Generalised Tiling into a Hamiltonian Spectrum	143
4.1.2	The Toric Code	145
4.1.3	Combining Hamiltonian Spectra	145
4.2	Size-Driven Quantum Phase Transitions	146
4.2.1	Hamiltonian Construction	146
4.2.2	Prime Period Tiling	149
4.2.3	Turing Machine Tiling	154
4.2.4	Thermal Stability	157
4.3	Chapter Summary	164
5	Beyond History States	167
5.1	Results and Overview	168
5.2	Improving Circuit Hamiltonians	172
5.2.1	Partial Diagonalisation of Weighted History States	172
5.2.2	Kitaev’s Geometrical Lemma for Weighted History States	172
5.2.3	Symmetrised Metropolis Hamiltonians with target ground state distributions	174
5.2.4	Explicit Construction of an $\Omega(T^{-2})$ UNSAT Penalty Circuit Hamiltonian	175
5.3	Tightness of the Geometrical Lemma for the UNSAT Penalty	176
5.3.1	A Tight Bound for the Clock Hamiltonian	176
5.3.2	A Tight Bound for the Full Circuit Hamiltonian	178
5.4	Limitations on further improvement	183
5.5	Universal adiabatic computation	186
5.6	Chapter Summary	190
6	Conclusion	193

Introduction

The freedom of human thought is very limited. We all live in a very narrow cage, the “zeitgeist”, in which we have very little freedom of motion. If, in different ages, people thought differently, this was not because the cage got wider, but because the cage moved.

—Albert Szent-Györgyi, *The Crazy Ape*

1 Many-Body Quantum Physics

If $|\Psi_t\rangle$ describes the state of a quantum mechanical system at time t , its time evolution is dictated by the time-dependent Schrödinger equation

$$\frac{\partial}{\partial t} |\Psi_t\rangle = -i\mathbf{H} |\Psi_t\rangle, \quad (1)$$

where \mathbf{H} is the Hamiltonian operator, i.e. a Hermitian matrix acting on state vectors such as $|\Psi_t\rangle$ in some Hilbert space \mathcal{H} . We generally assume that all Hilbert spaces are finite-dimensional, if not specified otherwise, and we use natural units ($c = k_B = \hbar = 1$). For time-independent problems, the eigenvalue equation

$$\mathbf{H} |\Psi\rangle = E |\Psi\rangle$$

yields the energy levels E of the system described by the Hamiltonian \mathbf{H} , where $|\Psi\rangle$ are the corresponding eigenstates at that energy level. The time-evolution of a quantum system can be obtained by solving the differential equation 1; one finds that $|\Psi_t\rangle = \exp(-i\mathbf{H}t) |\Psi_0\rangle$ for some initial state $|\Psi_0\rangle$ is a valid solution. Since the Hamiltonian \mathbf{H} is Hermitian, the operator $\mathbf{U}_t = \exp(-i\mathbf{H}t)$ is unitary for all times t , i.e. $\mathbf{U}_t \mathbf{U}_t^\dagger = \mathbf{1}$.

Many-body problems address the case where the system contains more than one distinguishable constituent, e.g. when there are two independent electrons, coupled by \mathbf{H} . In this case, the Hilbert space factorises into a product space $\mathcal{H} = \mathcal{H}_1 \otimes \dots \otimes \mathcal{H}_n$, where n is the number of particles, and $\dim \mathcal{H}_i$ is the local dimension of the i^{th} particle. In case \mathbf{H} describes particles which are non-interacting, the Schrödinger equation can be solved by a simple combinatorial reconstruction from the energy levels of the local systems.

The by far more interesting case is when at least some of the spins interact, i.e. when they are coupled by some non-trivial terms in the Hamiltonian \mathbf{H} . Such interacting quantum many-body systems appear ubiquitously in all areas of physics, and an important subclass are *local* Hamiltonians: if

$$\mathbf{H} = \sum_i \mathbf{h}_i \tag{2}$$

where \mathbf{h}_i acts non-trivially on at most k subsystems for some k , we call \mathbf{H} k -local. Observe that this definition makes no assumption on the geometric layout of the system, or how long-range the interaction is. Local Hamiltonians can be nearest-neighbour, but we may also encounter highly-contrived interaction graphs without any nice geometric structure.

In nuclear physics, for example, protons and neutrons in the nuclei are held together by an attractive interaction (the strong force), counteracting the electromagnetic repulsion of the positively-charged protons. Various descriptions and effective Hamiltonians exist that model nuclei, see [ERS75]. A well-known example in this context, widely-used in quantum chemistry, is the famous Born-Oppenheimer approximation [ERS75, appdx. L], factoring the wave function of the molecules into nuclei and electrons, i.e. $|\Psi_t\rangle = |\Psi_t^{\text{electrons}}\rangle \otimes |\Psi_t^{\text{nucleus}}\rangle$. Under this assumption, one can write down the interaction Hamiltonian \mathbf{H} which contains terms that bind the electrons to a static Coulomb potential arising from the nuclei’s positions, but neglecting the nuclear kinetic energy itself; in this way the electronic wave function can be approximated independently.

Many-body physics is also the foundation for condensed matter physics, i.e. the study of condensed phases of matter, such as liquids or solids. An important class of solids are crystals, where the particles sit at the intersections of a fixed lattice, and with interactions that are translationally-invariant under a shift by a lattice vector (i.e. the interactions are “the same” irrespective of the position of the lattice, where the boundaries are potentially excluded).

The Schrödinger equation describes the evolution of a quantum system under the assumption that there is no outside world interacting with it, e.g. in the form of noise; such systems are called *closed*. But what about a more realistic scenario, where environmental effects cannot be neglected—i.e. what about open quantum dynamics? Mathematically, we describe open quantum systems with so-called quantum operations. An excellent overview can be found in [NC10, ch. III]; let us summarise the major concepts briefly. Quantum operations can be used in a broad array of cases, both to describe weakly-coupled systems—where the environment is a mere perturbation to the closed system dynamics—as well as strongly-coupled systems. We denote the action of a quantum operation \mathcal{E} on a quantum state ρ simply via $\mathcal{E}(\rho)$. But which class of maps \mathcal{E} is a valid quantum operation? For closed dynamics, $\mathcal{E}(\rho) = \mathbf{U}\rho\mathbf{U}^\dagger$ represents closed-system evolution under the unitary matrix \mathbf{U} ; measurements can similarly be represented via $\mathcal{E}(\rho) = \mathbf{M}\rho\mathbf{M}^\dagger$, where for example \mathbf{M} is now a projective measurement operator.

A convenient way of thinking about the more general case of open system dynamics is to consider the environment ρ' as part of the system ρ that we aim to model, perform a unitary evolution, and then trace out the environment. Mathematically, and under the assumption that system and environment start out in a product state, we can describe the dynamics via

$$\mathcal{E}(\rho) = \text{tr}_{\text{environment}} (\mathbf{U}\rho \otimes \rho' \mathbf{U}^\dagger). \quad (3)$$

This characterisation of a quantum operation is also known as Stinespring dilation [Sti55]. It can be shown that for all possible quantum operations, it suffices to assume that $\dim \rho' = (\dim \rho)^2$, and that the environment starts out in a pure state¹, i.e. $\rho' = |e\rangle\langle e|$. If we expand the partial trace in eq. (3)—denoting with $\{|i\rangle\}_i$ a basis for the environment, we obtain

$$\mathcal{E}(\rho) = \sum_i \langle i| \mathbf{U}(\rho \otimes |e\rangle\langle e|) \mathbf{U}^\dagger |i\rangle =: \sum_i \mathbf{E}_i \rho \mathbf{E}_i^\dagger, \quad (4)$$

which is also known as the operator-sum representation of the quantum operation \mathcal{E} , and the \mathbf{E}_i are called Krauss operators [Cho75]. It is straightforward to show that $\sum_i \mathbf{E}_i^\dagger \mathbf{E}_i = \mathbb{1}$, as long as $\text{tr} \mathcal{E}(\rho) = 1$ —i.e. for operations which are *trace-preserving*—in [NC10] the authors include the case when $\text{tr} \mathcal{E}(\rho) \leq 1$, which e.g. describes a measurement on the joint system and environment; for the purposes of this thesis we will not need to consider this case. There is also a natural way to translate an operator-sum expression as in eq. (4) into a physical system-environment model; we refer the interested reader to [NC10, sec. 8.2.3].

Given some quantum operation $\mathcal{E}(\rho) = \sum_i \mathbf{E}_i \rho \mathbf{E}_i^\dagger$ and a positive operator \mathbf{A} on some arbitrary *extended* system (referring to \mathbf{A} 's eigenvalues being nonnegative), it is immediate to show that for an identity operation \mathcal{I} on the extended system, the joint operation $(\mathcal{I} \otimes \mathcal{E})(\mathbf{A})$ yields a positive map as well—a property of \mathcal{E} which we call complete positivity. In brief, a quantum operation is thus a CPTP map—completely positive and trace preserving.

A bottom-up classification for CPTP maps is given e.g. in [Wolo8]; [NC10, Th. 8.1]: if we want to construct the most general map from density operators to density operators $\mathcal{T} : \mathcal{B}(\mathcal{H}) \rightarrow \mathcal{B}(\mathcal{H}')$ —where $\mathcal{B}(\mathcal{H})$ denotes the bounded linear operators on \mathcal{H} —we need to require three properties that make such a mapping reasonable.

1. Linearity: $\forall c \in \mathbb{C}$ and $\rho, \rho' \in \mathcal{B}(\mathcal{H}) : \mathcal{T}(\rho + a\rho') = \mathcal{T}(\rho) + a\mathcal{T}(\rho')$.
2. Trace-Preservation: $\text{tr} \mathcal{T}(\rho) = 1$ if $\text{tr} \rho = 1$, which together with linearity implies $\forall A \in \mathcal{B}(\mathcal{H}) : \text{tr}(\mathcal{T}(A)) = \text{tr} A$.
3. Complete Positivity: preserves nonnegative eigenvalues (i.e. $\mathcal{T}(A^\dagger A) \geq 0 \forall A \in \mathcal{B}(\mathcal{H})$); this

¹We can always enlarge the environment to purify it.

necessarily has to hold true if the channel describes part of the evolution of a bigger system, which does not itself interact. More explicitly, $\mathcal{T} \otimes \mathcal{I}_n \geq 0$ for all $n \in \mathbb{N}$, where \mathcal{I}_n is the identity on an n -dimensional auxiliary system.

These assumptions together precisely specify CPTP maps.

In contrast to closed system dynamics, time evolution under a quantum operation is not necessarily unitary. Yet under the assumption that the evolution is *Markovian* (see [Pre15, sec. 3.5]), there exists an analogue to the evolution described by the Schrödinger equation. More specifically, if $\rho(t)$ is to be governed by a differential equation comparable to the Schrödinger equation, we know that $\rho(t + dt)$ can *only* depend on $\rho(t)$.² In this case, and if we want to describe the open quantum evolution for all times $t \geq 0$, we can then use a so-called Markovian quantum master equation [CEW12b], i.e. we describe the dynamics of a density matrix ρ via the Liouvillian \mathcal{L} in Lindblad form

$$\frac{d\rho}{dt} = -i[\mathbf{H}, \rho] + \sum_i \left(2\mathbf{L}_i \rho \mathbf{L}_i^\dagger - \{\mathbf{L}_i^\dagger \mathbf{L}_i, \rho\} \right). \quad (5)$$

Here, \mathbf{H} is the usual Hamiltonian with the system’s internal interactions (including, potentially, the unitary part of the open system dynamics generated by \mathcal{L}); the \mathbf{L}_i are the Lindblad operators describing the coupling of the system with the environment. The \mathbf{L}_i form an orthogonal linear basis of the operator space; $\{\cdot, \cdot\}$ denotes the anti-commutator. Such a Lindblad equation naturally generates a semi-group of CPTP maps \mathcal{E}_t , which describes the open quantum evolution of a system for time $t \geq 0$, as a solution of the form $\rho(t) = \mathcal{E}_t(\rho(0))$. The converse direction is not true in general, and deciding whether or not some CPTP map \mathcal{E}' is Markovian or not is computationally hard [CEW12b]. In chapter 1, we study a related question, i.e. in which cases one can efficiently sub-divide a CPTP evolution into smaller but discrete sub-steps, *without* requiring semi-group structure.

The spectral gap of a Hamiltonian \mathbf{H} defined via

$$\Delta(\mathbf{H}) := \lambda_1(\mathbf{H}) - \lambda_{\min}(\mathbf{H}),$$

i.e. as the energy difference between \mathbf{H} ’s ground and first excited state energies. It plays a central role in studying the physical properties of the system it describes. “Gapped” and “gapless” are, strictly speaking, only well-defined properties in the large system limit: if for a family of Hamiltonians (\mathbf{H}_n) , where n is the system’s size, has a finite gap in the limit $n \rightarrow \infty$, then the system is gapped; otherwise it is gapless—to avoid ambiguity, we often demand that the spectrum above the ground state is continuous in the latter case. We for example know that the transverse Ising model has an exponentially-closing gap; [LSM61] showed

²If the system is strongly coupled to the environment, Markovianity is often not the case, as information can be exchanged with the environment for a period of time—the history of the evolution starts to play a role, i.e. the evolution deviates from the case where the initial conditions are product.

that this extends to the half-integer spin Heisenberg model for spins aligned on a chain, which [Haso4] extended to arbitrary-dimensional lattices. The case of integer spins remains open; for the one-dimensional anti-ferromagnetic Heisenberg model this problem is known as the Haldane conjecture [Hal83].

2 Quantum Stochastic Processes

A Markov chain is probably the most intuitive-to-understand example of a stochastic process [Bas11]. Assume you stand on a tiled floor, and with a certain probability—decided e.g. through rolling a few dice—you either take a step forward, or backward. Such a discrete random walk is an example of a Markov chain, where your position on the tiles is the state $s \in S$, and the probabilities associated with moving in a certain direction are the transition probabilities p_{ij} between states $i \mapsto j$, with $i, j \in S$. The transition matrix $\mathbf{P} = (p_{ij})_{i,j \in S}$ is thus such that the row sums are 1, and every entry satisfies $0 \leq p_{ij} \leq 1$.

Markov chains are used ubiquitously in a broad range of fields. In chemistry, equilibria of reactions can be computed by analysing the steady states of the corresponding Markov chain; an example are Michaelis–Menten kinetics of enzymes, $E + S \rightleftharpoons ES \rightarrow E + P$ for some enzyme E , substrate S and product P . The equilibrium between $E + S$ and the enzyme-substrate-complex ES can be calculated from the individual transition probabilities $p_1 = \mathbb{P}(E + S \rightarrow ES)$ and $p_2 = \mathbb{P}(ES \rightarrow E + S)$ for a specific time interval. In the language of Markov chains and transition matrices,

$$\mathbf{P} := \begin{pmatrix} 1 - p_1 & p_1 \\ p_2 & 1 - p_2 \end{pmatrix}$$

describes this process; the steady state is the stationary vector—the left eigenvector of \mathbf{P} corresponding to the maximal eigenvalue of 1—which in this case is simply $\langle s | = (p_2/p_1 \ 1)$. Adding in the reaction $ES \rightarrow E + P$ with probability q and without inverse reaction, the steady state of the matrix

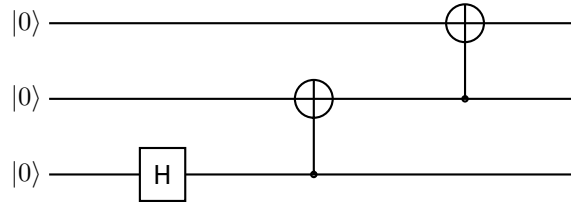
$$\mathbf{P}' := \begin{pmatrix} 1 - p_1 & p_1 & 0 \\ p_2 & 1 - p_2 - q & q \\ 0 & 0 & 1 \end{pmatrix}$$

is simply $\langle s' | = (0 \ 0 \ 1)$, as the product reaction acts as a sink from which there is no escape.

Stochastic processes also play a major role in data compression (such as the Lempel-Ziv-Markov chain algorithm LZMA, see [Sal07]), search (Google’s PageRank™ algorithm for weighting search results [Fero8]) or for modelling other random processes in finance, biology or statistics, see chapter 1.

Such classical Markov chains also play a fundamental role in encoding quantum computation into the ground state of a (local) Hamiltonian operator. A detailed and rigorous introduction can be found in

chapter 2, but let us give a short overview over this technique. Assume you are given the following short quantum circuit that can be used to prepare a GHZ state $(|000\rangle + |111\rangle)/\sqrt{2}$:



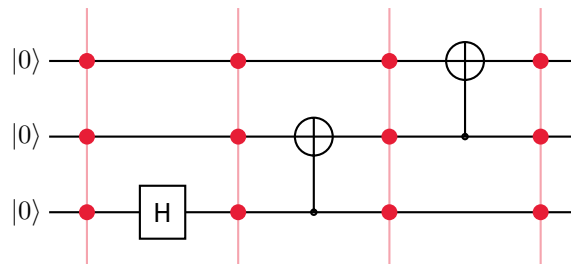
The circuit consists of three unitaries: a Hadamard gate $\mathbf{U}_1 = \text{HADAMARD}$, and two controlled not gates $\mathbf{U}_2 = \mathbf{U}_3 = \text{CNOT}$. Our goal is to write a Hamiltonian \mathbf{H} , which in its ground state $|\psi_0\rangle$ “encodes” the action of the quantum circuit. It is an interesting question to ask what “encoding” means precisely in this context: here, it would supposedly be sufficient to be able to access the output of the quantum circuit, say by measuring the ground state in a specific basis, and to some precision that we are satisfied with. This immediately opens the question, however, whether we can trust the ground state: how do we guarantee that the output is the one given by the quantum circuit?

A trivial way of ensuring this is by writing $\mathbf{U} = (\mathbb{1}_4 \otimes \mathbf{U}_1)(\mathbb{1}_2 \otimes \mathbf{U}_2)(\mathbf{U}_3 \otimes \mathbb{1}_2)$ for the overall quantum circuit, and then defining a Hamiltonian

$$\mathbf{H} := \mathbf{U}(\mathbb{1}_8 - (|0\rangle\langle 0|)^{\otimes 3})\mathbf{U}^\dagger. \quad (6)$$

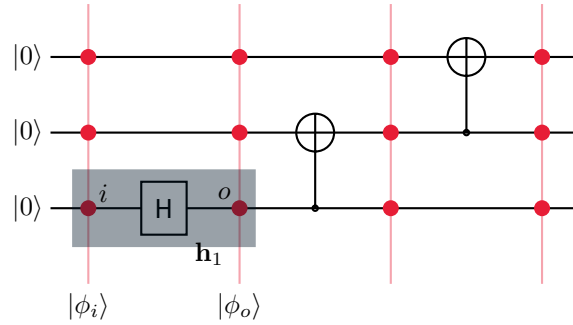
It is straightforward to verify that the state $\mathbf{U}|0\rangle$ is the unique ground state with eigenvalue 0, and all other states have eigenvalue 1. However, \mathbf{H} will not, in general, be a local operator as in eq. (2), even though the gates \mathbf{U}_1 , \mathbf{U}_2 and \mathbf{U}_3 are local in the sense that they only touch at most two qubits at the same time.

Following an argument from [Cub15], a second attempt would be to partition the circuit into intermediate steps:



Each dot \bullet now represents one qubit with Hilbert space \mathbb{C}^2 , and we can hope to define local Hamiltonian operators acting on the in- and output qubits of each gate that ensure that the overall Hamiltonian ground state still encodes the computation output. This, however, is not possible in general. To prove this, consider

the HADAMARD gate only, for which we want to define a local Hamiltonian term \mathbf{h}_1 acting on spins labelled i and o :

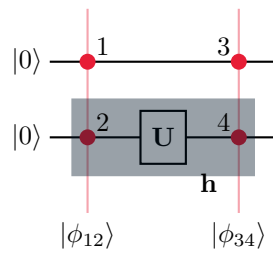


The states $|\phi_i\rangle$ and $|\phi_o\rangle$ represent the initial state for the circuit—i.e. $|\phi_i\rangle = |0\rangle^{\otimes 3}$ —and the state after the first gate application, i.e. $|\phi_o\rangle = (\mathbb{1}_4 \otimes \mathbf{U}_1) |\phi_i\rangle = |0\rangle^{\otimes 2} \otimes \mathbf{U}_1 |0\rangle$. What the Hamiltonian term \mathbf{h}_1 thus has to accomplish is that

$$\ker(\mathbb{1}_4^{\otimes 2} \otimes \mathbf{h}_1) = \text{span}\{|\phi_i\rangle \otimes |\phi_o\rangle\}, \quad (7)$$

where we assumed without loss of generality that \mathbf{h}_1 is positive semi-definite³. One can verify that $\mathbf{h}_1 = \mathbb{1}_4 - |0\rangle\langle 0| \otimes \mathbf{U}_1 |0\rangle\langle 0| \mathbf{U}_1^\dagger$ satisfies eq. (7); but note how this Hamiltonian term depends on the input to the computation step (i.e. the state $|0\rangle$). We could go ahead and define a different Hamiltonian for every input, which would not be particularly interesting. A *general* term \mathbf{h}_1 which works *independent of the input* is what we ultimately seek—and this is where the construction breaks down. In fact, one can prove a no-go theorem forbidding an embedding in this fashion in generality⁴.

To do this, assume you want to encode an even simpler circuit as eq. (7), i.e. for two general input qubits $|\phi_{12}\rangle \mapsto \mathbb{1} \otimes \mathbf{U} |\phi_{12}\rangle =: |\phi_{34}\rangle$, where \mathbf{U} is some single-qubit gate; in the following, subsets on states and operators will denote the qubits that the objects are defined on, according to the following assignment (and for the sake of clarity we will drop the subscript on $\mathbb{1}$ that indicates the identity matrix size, which will be clear from context):



In particular, this should also hold for a Bell state $|\phi_{12}\rangle = |\Phi^+\rangle = (|00\rangle + |11\rangle)/\sqrt{2}$. We pick the

³Substitute \mathbf{h}_1 with $\mathbf{h}_1 - \lambda_{\min}(\mathbf{h}_1)\mathbb{1}$ otherwise.

⁴It is worth pointing out that this approach clearly works when the computation is classical, since the classical gate never sees the entanglement of the input state.

corresponding state from $\text{span}\{|\phi_{12}\rangle \otimes |\phi_{34}\rangle\}$, namely

$$|\Phi^+\rangle_{12} \otimes [(\mathbf{1} \otimes \mathbf{U}) |\Phi^+\rangle]_{34} = ((\mathbf{1})_{123} \otimes (\mathbf{U})_4) |\Phi^+\rangle_{12} \otimes |\Phi^+\rangle_{34} =: |\psi\rangle,$$

where the subscripts mark which qubits the state denotes; observe that the subsystems are ordered 1–2–3–4 here. According to eq. (7), the state should be in the kernel of $(\mathbf{1})_{13} \otimes (\mathbf{h})_{24}$. A direct calculation of the corresponding expectation value then yields

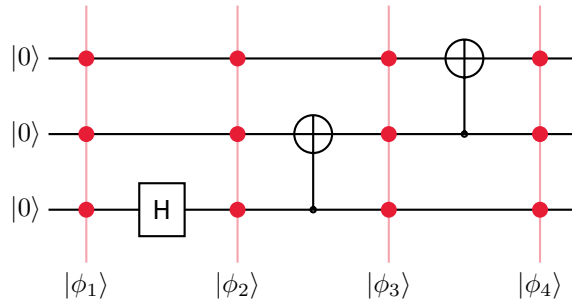
$$\begin{aligned} \langle \psi | (\mathbf{1})_{13} \otimes (\mathbf{h})_{24} | \psi \rangle &= \text{tr} \left((\mathbf{1})_{13} \otimes (\mathbf{1} \otimes \mathbf{U})_{24}^\dagger (\mathbf{h})_{24} (\mathbf{1} \otimes \mathbf{U})_{24} |\Psi^+\rangle\langle\Psi^+|_{12} |\Psi^+\rangle\langle\Psi^+|_{34} \right) \\ &= \text{tr} \left((\mathbf{h}')_{24} \text{tr}_1 |\Psi^+\rangle\langle\Psi^+|_{12} \text{tr}_3 |\Psi^+\rangle\langle\Psi^+|_{34} \right) \\ &= \text{tr} \left((\mathbf{h}')_{24} (\mathbf{1})_2 \otimes (\mathbf{1})_4 \right) \\ &= \text{tr}(\mathbf{h}') = 0 \end{aligned}$$

and thus $\mathbf{h}' = 0$, since a unitary similarity transformation preserves the operators spectrum—i.e. \mathbf{h} and \mathbf{h}' have the same eigenvalues; since the trace is the sum of the eigenvalues of \mathbf{h} , which is positive semi-definite by assumption. We can thus conclude $\mathbf{h} = 0$, which contradicts eq. (7): the span is non-empty, whereas $\ker(\mathbf{1} \otimes 0) = \{0\}$, and the claim follows.

It is a curious property of quantum computation that seems to render it too complicated to be embedded in this “local Hamiltonian checks”-type embedding, while preserving a tensor product structure of the computational history as a witness to the computation, i.e. a ground state of the type $|\phi_1\rangle \otimes |\phi_2\rangle \otimes \dots \otimes |\phi_T\rangle$. But is there an alternative, which still allows us to encode a quantum computation into the ground state of a local Hamiltonian? In fact there is, but it requires a modification of what we deem an embedding of a computation.

In 1986, Feynman proposed to build a Hamiltonian where the computational history is instead preserved as a so-called *history state*, i.e. $\sum_{t=1}^T |t\rangle \otimes |\phi_t\rangle$, which is defined on the bipartite Hilbert space $\mathbb{C}^T \otimes (\mathbb{C}^2)^{\otimes n}$, where n denotes the number of qubits in the circuit [Fey86]. The first subspace acts as a clock register, which keeps track of the current computational step.

To be a bit more precise, consider again the GHZ circuit from before:



Then the Hamiltonian

$$\begin{aligned}
 \mathbf{H} := & - (|2\rangle\langle 1| \otimes \mathbf{U}_1 + |1\rangle\langle 2| \otimes \mathbf{U}_1^\dagger) \\
 & - (|3\rangle\langle 2| \otimes \mathbf{U}_2 + |2\rangle\langle 3| \otimes \mathbf{U}_2^\dagger) \\
 & - (|4\rangle\langle 3| \otimes \mathbf{U}_3 + |3\rangle\langle 4| \otimes \mathbf{U}_3^\dagger)
 \end{aligned} \tag{8}$$

has a ground space spanned by states of the form $\sum_{t=1}^4 |t\rangle \otimes |\psi_t\rangle$, where $|\psi_1\rangle \in (\mathbb{C}^2)^{\otimes n}$ is an arbitrary quantum state; $|\psi_2\rangle = \mathbf{U}_1 |\psi_1\rangle$, $|\psi_3\rangle = \mathbf{U}_2 \mathbf{U}_1 |\psi_1\rangle$ etc., which if we could ensure that $|\psi_1\rangle = |\phi_1\rangle = |0\rangle^{\otimes 3}$ is correctly initialised would precisely be the history state we want⁵.

Luckily, we already know how to properly initialise the ground state computation, see eq. (6): by adding a local term of the form $\mathbf{P} := |1\rangle\langle 1| \otimes (\mathbb{1}_8 - |\phi_1\rangle\langle \phi_1|)$, we ensure that at time $t = 1$ all states in a configuration not equal to $|\phi_1\rangle$ obtain a penalty. The ground state of \mathbf{H} is then non-degenerate, i.e.

$$\ker(\mathbf{H} + \mathbf{P}) = \text{span}\{|1\rangle \otimes |\phi_1\rangle + |2\rangle \otimes |\phi_2\rangle + |3\rangle \otimes |\phi_3\rangle + |4\rangle \otimes |\phi_4\rangle\},$$

as required. This method obviously works for any quantum circuit, and as long as one manages to write the clock in a local fashion—which was accomplished by [KSV02] in 2002. The difficult parts of translating these constructions into statements on how hard it is to estimate the Hamiltonian’s ground state energy (known as the LOCAL HAMILTONIAN problem, see definitions 2.1, 2.16 and 3.1) stems from the fact that if in addition to the input penalty \mathbf{P} , we also give an output penalty $\mathbf{P}' = |4\rangle\langle 4| \otimes \Pi$, the Hamiltonian will in general be frustrated, which complicates a systematic analysis of the spectrum of

$$\mathbf{H}_{\text{total}} := \mathbf{H} + \mathbf{P} + \mathbf{P}'. \tag{9}$$

⁵We can add diagonal terms to \mathbf{H} to make it a positive semi-definite operator; \mathbf{H} then has a similar structure to a graph Laplacian, which is in fact how the full eigenspectrum of \mathbf{H} becomes accessible; for details we refer the reader to section 4 and chapter 2.

Obtaining lower- and upper bounds for the ground state energy of $\mathbf{H}_{\text{total}}$ has been an active field of research ever since Kitaev’s seminal result proving that estimating the ground state energy of a local Hamiltonian is QMA complete.

The connection between this embedding of quantum computation into the ground state of a local Hamiltonian and a classical Markov chain becomes apparent when looking at eq. (8): on the clock register, \mathbf{H} takes the form

$$\mathbf{H}_{|\text{clock}} = - \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} =: 2\mathbf{T},$$

- which is the adjacency matrix for a path graph of length 4; \mathbf{T} is stoquastic, and we could say that this (almost, up to normalization) transition matrix on the clock register “drives” the quantum computation, as a step from say $t = 2$ to $t = 3$ implies an application of the corresponding unitary \mathbf{U}_2 to the quantum register. We will address this connection in greater detail in section 4. We need to point out that this “quantum stochastic process” is not the same—nor technically related—to the notion of a quantum Markov process, which is a generalisation of a classical Markov process where the transition matrix is a CPTP map repeatedly acting on an initial density matrix ρ_0 , see Gudder2008

Since we will discuss the history of this problem in great detail in chapter 2, we instead turn our attention to a survey of (quantum) complexity theory, and how it comes into play in the context of condensed matter theory and quantum stochastic processes.

3 Complexity Theory

Assume you are looking for a book in a library, but the electronic catalogue is out of service, the front desk is closed and the shelves are not sorted. How would you go about finding it? A naïve way would be to just walk by every shelf and scan the back of every book, until you find the one you seek. Assuming the library has N books on the shelves, the worst-case scenario is that you will find the book after looking at every single one in the library, i.e. N . Realistically, the expected number of books you have to check is $N/2$. The runtime scaling—in this case linear—is what complexity theory addresses: in this case, to be more specific, we are talking about the time complexity, i.e. the number of primitive steps (reading book titles) one needs to undertake in order to answer the question.

The computational model used in this procedure is of course important. In fact, the Church-Turing thesis⁶ states that any model of computation should be equivalent to a Turing machine [Woloz, Ch. 12]—which we will introduce rigorously in chapter 2—but only in the sense that if a (deterministic) computation

⁶The Church-Turing thesis is, strictly speaking, a conjecture, and there are many subtleties in proving any such equivalence explicitly.

can be performed on one device, it can also be done on a (deterministic) Turing machine. The runtime might obviously differ, depending on the computational power of the device at hand; assuming one could e.g. read three book titles at once, the worst-case runtime for our book hunt is $N/3$ rather than N ; if one has access to an oracle—a magic black-box that solves a sub-problem without additional cost—the runtime is often reduced.

A Turing machine is a simple computational model introduced by Alan Turing in 1936, mentioned in many standard textbooks on complexity theory [AB09]; [NC10]; [Wat12a] both in a classical and quantum setting. In brief, a Turing machine consists of a two-way infinite tape of symbols from a fixed set S , a head positioned somewhere on the tape and with some internal state $q \in Q$, and a transition function $\delta : S \times Q \rightarrow S \times Q \times \{\text{left}, \text{right}\}$. At each step, the transition function tells the head to read a symbol $s_i \in S$ from the tape at the head's position, and depending on the head's internal state q_i at that point write a symbol $s_o \in S$, update its internal state to $q_o \in Q$, and move in direction $d \in \{\text{left}, \text{right}\}$, where $(s_o, q_o, d) = \delta(s_i, q_i)$. There are obviously some subtleties to consider in this definition, but we defer a precise formulation to section 2.2.1. The runtime of such a Turing machine is then simply the number of steps required from some input written out on the tape, to some halting configuration $q_f \in Q$ —which is well-defined only if the Turing machine terminates, a problem that is undecidable in general [Dav58] and often referred to as the *halting problem*.

From a complexity-theoretic point of view, one tries to capture different runtimes within so-called time complexity classes. Search is clearly *in* linear time⁷, and thus belongs to the class \mathbf{P} , see section 2.2.1.2: \mathbf{P} is the class of algorithms which, on an input of size N , terminates successfully within poly N steps, where it is important to point out that the polynomial itself has to be the same for all inputs. Similar classes can be defined for exponential runtime, denoted $\mathbf{EXPTIME}$, or logarithmic time, which we write as $\mathbf{DLOGTIME}$, see [AB09]. When working with problems in \mathbf{P} (or other super-polynomial classes), one commonly allows so-called Karp- or poly-time reductions for algorithms to show that they are still contained in \mathbf{P} . Reduction, in this context, simply means that one can cast the instances of some problem A into instances of some other problem B , with the property that the answer remains the same; a poly-time reduction then says that the overhead introduced by the cast is upper-bounded by a fixed polynomial for all instances. Whether an algorithm has runtime N or N^{56} has no bearing for this classification into \mathbf{P} (as either are polynomials), even though the difference might be striking for any practical applications.

Examples for recent complexity-theoretic results for deterministic algorithms are, amongst others, the Agrawal–Kayal–Saxena primality test [AKSo4] shown to be in \mathbf{P} , and the graph isomorphism problem (trivially contained in $\mathbf{EXPTIME}$): in 2015 László Babai announced that he discovered a quasi-polynomial time algorithm for this problem (i.e. with a run-time of $O(\exp((\log N)^c))$) for an input of size N , and

⁷Naturally, one can speed up all subsequent searches by first sorting the shelves—e.g. using block merge sort, which has an $O(N \log N)$ worst-case runtime [KK08]. Performing a search through an ordered list then has a logarithmic runtime.

some fixed $c > 0$), a claim that was retracted earlier this year but with a promised amendment which would save the result [Bab15]; [Bab17].

While complexity theory is rarely discussed outside of science, the (in)famous **P** vs. **NP** conjecture has found its way into pop culture⁸. **NP**—non-deterministic poly-time—is the class of problems that can be answered with a non-deterministic Turing machine [AB09, sec. 2.1.2]. Such a Turing machine has two distinct transition functions δ_0 and δ_1 , and, at each step, picks one or the other; the problems that can be answered with such a machine are those for which there exists *at least one* path through the computation—i.e. a sequence of choices—for which the computation accepts the input. Before explaining this further, it is worth spending a few words on how to rigorously define the term “problem”. It is common to talk of *decision problems* in this context, i.e. **YES** or **NO** questions. The question “Where is Nielsen&Chuang in the library?” is not a decision problem, but the question “Is it on the second floor?” qualifies as such. For a Turing machine which could answer this question (e.g. if it has knowledge of the library stored in the internal head configurations), we obviously need to specify the input following a specific encoding on the tape, e.g. by writing the book title in the Cyrillic alphabet in a specific direction, or an UTF-8 encoded little-endian null-terminated binary sequence on an otherwise blank tape. And here is the problem: what if an invalid input is given? Answering whether a Turing machine halts on an arbitrary input is, in general, undecidable, as we already discussed. We can sidestep this problem by talking of *promise problems*, where the promise is that the input is from a set of validly-formatted strings $s \in \Pi$ and such that $\Pi = \Pi_{\text{YES}} \dot{\cup} \Pi_{\text{NO}}$, i.e. s is either a **YES** or a **NO** instance. A promise-**NP** Turing machine is then said to accept s if it outputs **YES** on some path (in which case $s \in \Pi_{\text{YES}}$), and **NO** otherwise ($s \in \Pi_{\text{NO}}$), a notion which is now well-defined⁹.

- Equivalently, we can characterise an **NP** promise problem by saying that there exists a deterministic poly-time Turing machine M , such that for any $s \in \Pi_{\text{YES}}$, there exists a witness or certificate w such that
- $M(s, w) = \text{YES}$, and if $s \in \Pi_{\text{NO}}$, for all strings w' , we have $M(s, w') = \text{NO}$ (it is worth pointing out that when **NP** is *not* defined on promise problems, there is no statement for the **NO** case).

The most fundamental **NP**-complete problem (meaning it is *in* **NP**, and as hard as any other problem in **NP** under poly-time reductions) is Boolean satisfiability, i.e. the question whether a Boolean formula is satisfiable, see section 1.1.2 and [Coo71].

Another important extension to the complexity zoo are probabilistic classes, such as **BPP** or **StoqMA**. Imagine you and your friend Bob both hold binary strings $s, s' \in \{0, 1\}^n$, respectively, and you want to determine whether the strings are equal. Assuming n can be large, you may not want to communicate the entire string s over a network (e.g. due to bandwidth limitations), so a character-to-character comparison is out of question. However, with access to a source of randomness on at least one side, you can still check whether the strings are identical or not, at least with high enough probability [Bel0]. In brief, the protocol

⁸**P** vs. **NP** is e.g. featured in the film *Travelling Salesman*, one episode of *The Simpson's* and an episode of *Elementary*.

⁹Strictly speaking we should use “promise-**NP**” throughout the thesis, but for the sake of brevity we will generally drop it.

works by viewing the string as an integer $s \in \{0, 1, 2, \dots, 2^{n-1}\}$, and building a “fingerprint” (i, f_i) , where i is a random number between 1 and n , and $f_i := s \pmod{p_i}$, where p_i is the i^{th} prime. This fingerprint is sent to Bob, who compares it with the corresponding fingerprint $f'_i := s' \pmod{p_i}$. One can show that in case $s \neq s'$, the number of primes on which this method fails is upper-bounded by $n/2$ —which in turn means the acceptance probability is upper-bounded by $1/2$. Repeating the protocol 100 times gives a failure probability of 2^{-100} , which gives sufficient confidence for any practical purposes.

Such randomised algorithms play a major role in computer science; before the deterministic primality test [AKSo4], a probabilistic BPP algorithm was the state of the art. Polynomial identity testing [AB09, sec. 7.2.2]—i.e. the question whether a polynomial $p \equiv 0$ —is one of the most important open problems in algebraic computing complexity [Sax14]; [AB09, sec. 7.2.3], with implications for circuit depth lower bounds, and which played a crucial role in proving that $\text{IP} = \text{PSPACE}$ [Sha92]¹⁰. Another important example is the question of testing perfect matchings in bipartite graphs, for which randomised algorithms exist that perform particularly well on dense graphs [AB09, sec. 7.2.3]. And in Hamiltonian complexity theory, BPP plays a role e.g. in a recent algorithm by [LVV15] for finding the ground state of a local 1D gapped quantum system. The class StoqMA is then to BPP what NP is to P, used e.g. in the hardness classification of interactions for the LOCAL HAMILTONIAN problem [CM14]; [PM15].

Instead of listing more complexity classes, we want to focus on BPP’s quantum analogue, the class BQP [AB09, ch. 20.3]; In short, BQP is the class of promise problems solvable in poly-time and with high probability using a quantum Turing machine. This definition sounds very intuitive, yet it hides almost all of the subtleties that arise when trying to construct a quantum version of classical Turing machines¹¹, and we refer the reader to [BV97a] for an extensive discussion. For the purpose of this introduction, we just point out that a quantum Turing machine is one with a unitary transition function (i.e. in particular reversible, a property Bernstein and Vazirani call *well-formed*), where the tape and internal configuration live in a Hilbert space, and where by poly-time quantum Turing machine we mean the machine halts with probability 1 (i.e. transitions to a *final superposition*) on all input strings with the head in the same position, and in polynomially many steps (called *well-behaved*).

Assume we consider a promise problem $\Pi \in \text{BQP}$. While we require that the quantum Turing machine M for this problem halts deterministically in poly-time on any $l \in \Pi$, the output will be a quantum state, and we have to measure it in order to obtain an answer. As for the class BPP, we thus demand that for $l \in \Pi_{\text{YES}}$,

$$\Pr(M(l) = \text{YES}) \geq \frac{2}{3}$$

¹⁰IP is the class of interactive poly-time proof systems, and PSPACE the class of problems solvable using a polynomial amount of space, without any explicit runtime restrictions.

¹¹In fact, the way we choose to define BQP rigorously in this work is not through quantum Turing machines, but uniform classes of quantum circuits, see definition 2.11.

and for $l \in \Pi_{\text{NO}}$,

$$\Pr(M(l) = \text{YES}) \leq \frac{1}{3}.$$

These probabilities can be amplified under a poly-time reduction, see fact 2.12. Analogously to above, QMA is to BQP what NP is to P. It is easy to see that $P \subseteq BPP \subseteq BQP$, yet none of the inclusions is known to be strict. While also $P \subseteq NP$, it is not known in which way BQP and NP are related [Ben+97].

4 Putting it all Together

So far, the introduction addressed seemingly disjoint topics. Quantum many body physics on one hand, a theory that aims to describe composite systems comprising our physical reality, and stochastic processes on the other hand, a mathematical tool to model random dynamics and steady state properties. These two topics are in fact tightly linked, and become particularly fruitful when approached from a complexity-theoretic direction.

To exemplify this connection, consider the family of CPTP maps generated by a master equation in Lindblad form eq. (5), which in linear operator representation

$$\mathbf{E}_{(i,j),(k,l)} := \langle i, j | (\mathcal{E}(|k\rangle\langle l|))^{\Gamma}$$

- can be written as $\mathbf{E}_t = e^{\mathbf{L}t}$, where \mathbf{L} is the linear operator representation of \mathcal{L} ; the reshuffling operation \cdot^{Γ} is defined in eq. (1.1), and we direct the interested reader to section 1.1 and [BŽ06] for an extensive discussion. Open quantum evolution that has such a generator \mathbf{L} is called Markovian, and as we learned, deciding whether or not a particular quantum channel arises from such a generator is computationally hard (NP-hard) to answer. In a similar fashion, a classical stochastic process can be generated by an operator, i.e. if \mathbf{P}_t is a semi-group of stochastic maps describing the transition probabilities in a Markov chain, whether or not $\mathbf{P}_t = e^{\mathbf{Q}t}$ for some generator¹² \mathbf{Q} is an old open question known as the embedding problem for stochastic matrices, and NP-hard to answer computationally (see [Elf37]; [CEW12b]).

This also shows how complexity theory enters the picture, and how we can utilise this connection to make statements about the real world. Using quantum process tomography, one can measure a quantum channel \mathcal{E} to within some precision. Yet even the weak membership variant of the problem—meaning deciding whether or not there is a Markovian channel “close” to the one at hand—is NP-hard to answer. It is thus computationally intractable in any practical sense to decide whether the channel \mathcal{E} is Markovian, in general, and assuming $P \neq NP$. In this work, we address a similar question: given this channel \mathcal{E} , can we at least “sub-divide” it into smaller step, in order to interpolate the evolution between two measurements? More

¹² \mathbf{Q} has a number of conditions attached, similar to the Lindblad generator \mathbf{L} : nonnegative off-diagonal, and zero row sums.

specifically, does there exist a channel \mathcal{F} such that

$$\mathcal{E} = \underbrace{\mathcal{F} \circ \dots \circ \mathcal{F}}_{n \text{ times}}$$

for some given $n > 1$ (definition 1.29)? Or in the classical setting, can we write a stochastic map \mathbf{P} as the power of another stochastic map \mathbf{Q} (definition 1.31)? And what about probability distributions: in which cases can a random variable X be written as the sum of independent random variables Y_1, \dots, Y_n (definitions 1.66 and 1.79)? What if we only care about an approximation (definitions 1.67 and 1.80)? There is a plethora of related questions, and we address the computational complexity of them in chapter 1.

Another connection between many-body quantum physics and quantum stochastic processes becomes important when considering the LOCAL HAMILTONIAN problem from a graph-theoretic angle; we will briefly discuss this here and extend upon section 2, but refer the reader to sections 2.2.1 and 2.2.3 for an extensive discussion. If we write the Hamiltonian in eq. (8) in a slightly different form and add diagonal terms,

$$\begin{aligned} \mathbf{H} := \sum_i & \left[(|1\rangle \otimes |i\rangle - |2\rangle \otimes \mathbf{U}_1 |i\rangle) (\text{Hermitian conjugate}) \right. \\ & + (|2\rangle \otimes |i\rangle - |3\rangle \otimes \mathbf{U}_2 |i\rangle) (\text{h. c.}) \\ & \left. + (|3\rangle \otimes |i\rangle - |4\rangle \otimes \mathbf{U}_3 |i\rangle) (\text{h. c.}) \right], \end{aligned}$$

we can define the partially-diagonalising unitary \mathbf{W} via (see lemma 2.41)

$$\begin{aligned} \mathbf{W} & := \prod_{t=1}^3 \left(|t+1\rangle\langle t+1| \otimes \mathbf{U}_1 \cdots \mathbf{U}_t + (\mathbb{1} - |t+1\rangle\langle t+1|) \otimes \mathbb{1} \right) \\ & = |4\rangle\langle 4| \otimes \mathbf{U}_3 \mathbf{U}_2 \mathbf{U}_1 + |3\rangle\langle 3| \otimes \mathbf{U}_2 \mathbf{U}_1 + |2\rangle\langle 2| \otimes \mathbf{U}_1 + |1\rangle\langle 1| \otimes \mathbb{1}, \end{aligned}$$

such that $\mathbf{W}^\dagger \mathbf{H} \mathbf{W} = \Delta_4 \otimes \mathbb{1}$. Here, Δ_4 is the graph Laplacian of a path graph with four vertices

$$G_4 := \begin{array}{cccc} 1 & 2 & 3 & 4 \\ \circ & \text{---} & \circ & \text{---} & \circ & \text{---} & \circ \end{array},$$

see claim 2.27. We can calculate Δ_4 as the difference between degree matrix $\mathbf{D} = \text{diag}(1, 2, 2, 1)$ and adjacency matrix of G_4 ,

$$\Delta_4 = \begin{pmatrix} 1 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 \\ 0 & -1 & 2 & -1 \\ 0 & 0 & -1 & 1 \end{pmatrix}.$$

The ground state of Δ_4 is—as aforementioned—a uniform superposition over all vertices, i.e. $|\Psi_0\rangle = \sum_{t=1}^4 |t\rangle$, with eigenvalue zero. For \mathbf{H} , we can thus calculate the ground state via

$$\mathbf{W}(|\Psi_0\rangle \otimes |i\rangle) = |1\rangle \otimes |i\rangle + |2\rangle \otimes \mathbf{U}_1 |i\rangle + |3\rangle \otimes \mathbf{U}_2 \mathbf{U}_1 |i\rangle + |4\rangle \otimes \mathbf{U}_3 \mathbf{U}_2 \mathbf{U}_1 |i\rangle.$$

Naturally, we can use a Markov chain instead of a graph, and its Markov chain Laplacian instead; for a discrete Markov process on the same graph, the transition matrix

$$\mathbf{T}_4 = \begin{pmatrix} 0 & 1 & 0 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

is not symmetric, but still only has real eigenvalues $\{-1, -1/2, 1/2, 1\}$. Its stationary vector—corresponding to the largest eigenvalue—is $|\Psi_0\rangle$, i.e. identical to the ground state of Δ_4 . The matrix $\mathbb{1} - \mathbf{T}_4$ should thus be equally suited for embedding computation, in a similar fashion as for graph Laplacians; by going to a Markov chain on



we obtain a symmetrised version of \mathbf{T}_4 , denoted \mathbf{T}'_4 , such that in this case—coincidentally— $2(\mathbb{1} - \mathbf{T}'_4) = \Delta_4$. However, the language of Markov chains allows us to go beyond uniformly-weighted history states, an important modification to Feynman’s original construction. We explore such extensions in chapter 5.

Embedding quantum computation into the ground state of a local Hamiltonian can serve multiple purposes. In this thesis, we mainly focus on the **LOCAL HAMILTONIAN** problem: we prove that local, translationally-invariant interactions on open boundary spin lattices in 1D (chapter 2, with local spin dimension 42 and nearest-neighbour couplings) and 3D (chapter 3, with local spin dimension 4 and 4-local couplings) give rise to **QMA_{EXP}**-hard to approximate ground state energies. Foremost, this is a complexity-theoretic result, yet it has implications for the physics of the material described: if we assume that not all easy-to-verify problems on an exp-time quantum system can also be *solved* in the same runtime—i.e. **QMA_{EXP} ≠ BQEXP**, analogous to **NP ≠ P**—then we also assume that we cannot physically *build* a device which solves **QMA_{EXP}**-hard problems efficiently. However, if we have an actual sample described by either the 1D or 3D Hamiltonians we construct, and we cool the system down, there are two possible outcomes.

1. The system quickly assumes its ground state, i.e. in a time that scales polynomially in the system size.
2. The system remains in a meta-stable low-energy local minimum.

In the first case, we actually build a device that solves **QMA_{EXP}**-hard problems efficiently; depending on

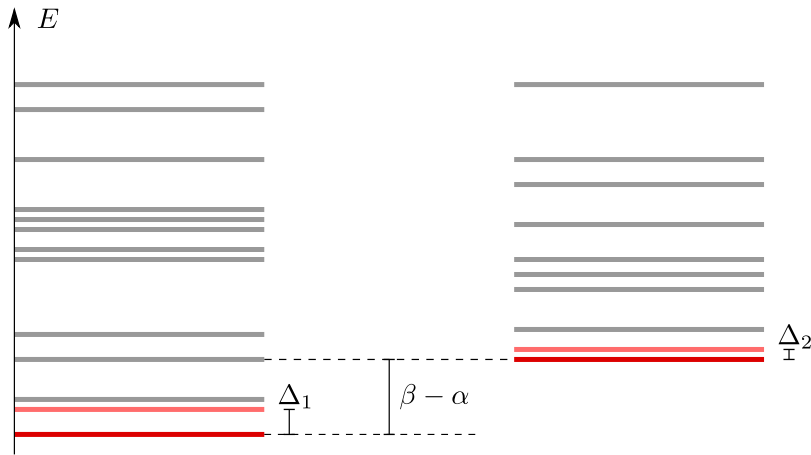


Figure 1: Difference between spectral gap and promise gap. On the left hand side, the Hamiltonian \mathbf{H}_{YES} has spectral gap $\Delta_1 := \lambda_1(\mathbf{H}_{\text{YES}}) - \lambda_{\min}(\mathbf{H}_{\text{YES}})$, and analogously Δ_2 for \mathbf{H}_{NO} . The promise gap $\beta - \alpha$ is the minimal difference between the two ground states for YES and NO instances, respectively, minimised over all Hamiltonians encoding instances of the same size. The higher energy spectrum plays no role in either definition.

the fidelity and specific runtime, this could showcase a complexity-theoretic collapse such as $\text{QMA}_{\text{EXP}} = \text{BQEXP}$ —although one has to be careful to claim empirical evidence as a proof. From a practical perspective, this would of course be a fascinating result; since e.g. $\text{NEXP} \subseteq \text{QMA}_{\text{EXP}}$. More realistic is the second case: systems like this behave in a spin-glass-like fashion, taking substantial time to equilibrate¹³, see [McG14].

Let us be more precise here, and connect this back to stochastic processes. For this we need a slightly more formal definition of the LOCAL HAMILTONIAN problem. Given an integer n and a k -local Hamiltonian \mathbf{H} on a multipartite Hilbert space $(\mathbb{C}^d)^{\otimes n}$, and two real numbers $\beta > \alpha$ such that $\beta - \alpha \geq 1/p(n)$,

¹³The equilibration time until the ground state is reached will generally depend on the spectral gap of the system, which we discuss below. Depending on the type of embedded computation—e.g. a BQP or BQEXP circuit—this gap will generally shrink along with the runtime. It is an interesting question how experimental cooling rates compare to the complexity-theoretic lower bounds on the spectral gap, but it isn't quite as straightforward to answer. In [KO17], for instance, the authors show that rapid cooling in some condensed matter systems takes the system to a meta-stable low energy state different from its ground state, since the procedure allows the system to avoid a first order phase transition (see also [Liu+01]; [Ryl+16] for more examples). The cooling rate thus becomes a parameter of the phase that the system ends up in; a lower cooling rate could, in principle, take the system to its ground state faster. For optical lattices, which have been proposed as a method for simulating condensed matter systems of strongly correlated particles which lies beyond the capabilities of current numerical approaches, another major obstacle seems to lie in the challenge of cooling the system down far enough to e.g. reach magnetic order (see Block2008; [MD11]); such optical lattices would also be excellent candidates for a Feynman-Kitaev type embedding of quantum computation due to the very high control experimentalists have for tuning the spin couplings and the relatively large number of local degrees of freedom. A rough estimate for the temperature necessary to embed, say, 100 gates into the ground state of a Hamiltonian and to be able to distinguish between YES and NO instances with a sufficient fidelity of e.g. ≈ 0.1 can be obtained from [Bau+16, eq. 10]: assuming an effective coupling strength of $g_{\text{eff}} \approx 10^{-14}$ eV (see [CPC17, eq. 1], where we assume $g_{\text{eff}} \propto 10E_R/\Delta_c$ with $E_R = \hbar^2 k^2/2M \approx 10^{-12}$ eV as recoil energy and a detuning of $\Delta_c \approx 20$ MHz, and [Aid16, ch. 4.3] for some experimental numbers) and a promise gap of $\Delta \approx 100^{-2} g_{\text{eff}}$ (see theorem 5.6) we get a temperature $T \approx 10^{-15}$ K. This lies five orders of magnitude below the lowest ever achieved temperature in a Bose Einstein condensate at the time of writing this thesis [Lea03]. Obviously this is just a very rough estimate, but it demonstrates the difficulties arising from a quickly-closing gap and a realization of the Feynman-Kitaev Hamiltonian in an experimental setup for the purposes of extracting computational power.

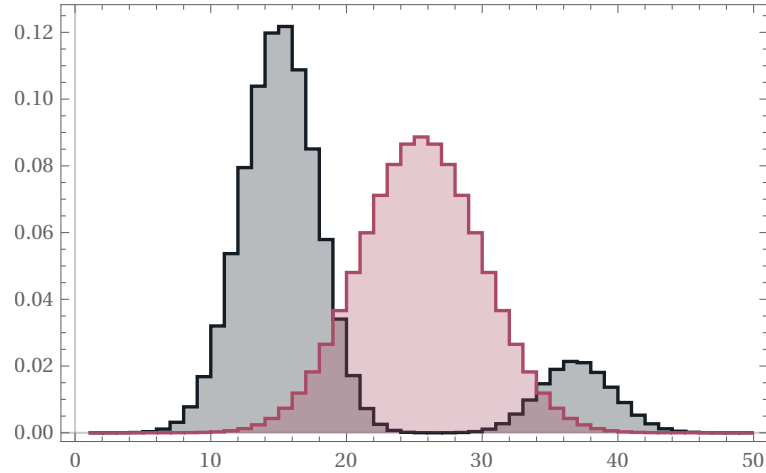


Figure 2: Classical, Gaussian-shaped random walk distribution on a line of length 50 after 20 steps (red), compared to a quantum walk (dark grey)—both with $1/3$ probability for left, right, or no step.

for some fixed polynomial $p(n)$, distinguish between $\lambda_{\min} < \alpha$, or $\lambda_{\min} > \beta$. We are promised that the smallest eigenvalue λ_{\min} of \mathbf{H} is either smaller than α or greater than β (whether or not this promise is hard to satisfy is not important, as mentioned in section 3). The polynomial *promise gap* here is crucial, and part of the definition of the LOCAL HAMILTONIAN problem; it has nothing to do with the *spectral gap* of the Hamiltonian at hand, see fig. 1.

The Feynman-Kitaev type history state embedding has a $1/\text{poly } T$ promise gap, where T is the number of steps of the embedded computation, or, more specifically, the number of vertices that make up the classical clock part of the computation. While, as we have seen, the propagation part of the Hamiltonian is unitarily equivalent to some graph or Markov chain Laplacian, adding both in- and output penalty will perturb the energy spectrum depending on the content of the computational register, see eq. (9).

An interesting question is whether such history state Hamiltonians are suitable for an adiabatic computation approach, i.e. by interpolating

$$\mathbf{H}_\lambda := \lambda \mathbf{H}_{\text{total}} + (1 - \lambda) \mathbf{H}_{\text{trivial}}$$

for λ between 0 and 1. $\mathbf{H}_{\text{trivial}}$ is then a Hamiltonian whose ground state is easy to compute or to reach in an experiment. By increasing λ sufficiently slowly, the adiabatic theorem says that the system described by \mathbf{H}_λ will remain in its ground state, or close to it [Far+oo]; [GTV15]; $\mathbf{H}_{\lambda=1}$ then contains the output to the computation embedded in $\mathbf{H}_{\text{total}}$. The speed of interpolation again depends crucially on the spectral gap along the path $\Delta(\mathbf{H}_\lambda)$, a question that depends on the particular interpolation scheme and the computation Hamiltonian used; we address this in chapter 5.

Another proposed scheme is Hamiltonian quantum computation, where the initial state is prepared, and then time-evolved under $\mathbf{U}_t = \exp(i\mathbf{H}t)$ —in this case there is no need for in- and output penalties [NW08]; [Nag12]; [WL15], but yet again the time until the output can be measured with high probability scales polynomially in the size of the embedded circuit. The connection between computation and complexity, and many-body physics implementing the stochastic process that drives the quantum computation is particularly obvious in this case, as the dynamics of such a system—i.e. the time evolution of the initial state $|\psi_t\rangle := \mathbf{U}_t |\psi_0\rangle$ —is very similar to a quantum diffusion process as illustrated in fig. 2. Beyond chapter 5, the interested reader finds an excellent survey of adiabatic and Hamiltonian quantum computation in [Nago8].

Last but not least, we construct a many-body quantum system on a two-dimensional lattice with a peculiar size-dependent transition from classical phase to topologically ordered (the Toric code, see chapter 4): the encoded “computation” differs from the history state construction, in that it is constructed from diagonal projectors representing a so-called Wang tiling; we present two constructions, one with a fixed but large period, and another one allowing an embedding of a classical Turing machine into the ground state of the system. As we noted before in section 3, such an embedding of classical computation as a tensor product over the computational history—as opposed to a superposition—can indeed be done, and does not contradict the no-go theorem we proved there.

This transition from classical to quantum low-energy spectrum is abrupt: there is no prior warning or level-crossing effect, or different energy scales in the Hamiltonian’s coupling constants, which could give away the length scale at which the transition occurs. We further prove that the transition is thermally robust, in the sense that there exists a finite temperature at which the Gibbs state is sufficiently close to the ground state of the system to observe said effect. Moreover, by taking the large-system limit and then lowering the temperature to zero, we uniquely recover the embedded topological model.

To summarise this introduction, it is worth emphasising that the topic of Hamiltonian complexity theory is intrinsically inter-disciplinary. The results e.g. on the complexity-theoretic side (novel computational models) often stand by themselves, but help our understanding and analysis of quantum many-body systems. The latter, in turn, offer a plethora of paths to go along, if we try to improve how computation can be embedded; and the general theme in this thesis is to work towards “physical relevance” wherever possible. I hope that my research accomplishes this goal, at least to some extent.

5 Structure of Thesis

Every chapter in this thesis (apart from introduction and conclusion) represents an independent paper; since most—if not all—of my work is done in collaborations, I summarise my co-authors contributions in the next section.

I begin this thesis with a discussion of divisibility problems (chapter 1) as outlined in section 1. This chapter is self-contained. Chapters 2 and 3 then contain two QMA_{EXP} -hardness proofs for the LOCAL HAMILTONIAN problem, once for a family of Hamiltonians in 1D, and then on 3D lattice crystals; the latter one heavily depends on the machinery introduced for the 1D case. In chapter 4, I present the novel size-driven phase transitions. The chapter is self-contained. Finally, in chapter 5, I consider modifications to Feynman’s history state construction, and applications to adiabatic quantum computation. The chapter has some links to chapter 2, but is largely independent otherwise.

We conclude the thesis in chapter 6 with a short outlook section.

6 Authors and Contribution

Chapter 1 is co-authored with Dr. Toby Cubitt and published ([BC16b]). Both the maps and distribution divisibility research idea is my supervisor’s; the proof method for embedding 1-IN-3SAT into a matrix (section 1.1.6) is inspired by the infinite divisibility case ([CEW12b]), but the actual proof itself is my work, as is the remainder of the reduction. In the distribution chapter (section 1.2), both proof ideas as well as their rigorous formulation for divisibility and decomposability are my work.

Chapter 2 is joint work with Dr. Toby Cubitt and Dr. Maris Ozols, and accepted for publication ([BCO17]). The idea to reduce the local dimension of Gottesman and Irani was my supervisor’s. The central techniques (Quantum Ring Machines, Unitary Labelled Graphs and Quantum Thue Systems) are my work, and have been refined with the help from both co-authors. The actual QMA-hardness construction (Turing’s Wheelbarrow) emerged from numerous iterations together with Dr. Ozols; the final version including the idea of ghosts and the hardness proof are my work.

- Chapter 3 is joint work with Stephen Piddock, and accepted for publication ([BP17]). Proving a 3D version of the LOCAL HAMILTONIAN problem was my idea, as was using tiling constructions to lower the dimension. Dr. Ozols contributed lemma 3.3. The history state construction emerged from joint discussions, and the final versions of the proofs are my work.

Chapter 4 is a collaboration together with Dr. Toby Cubitt, Dr. Angelo Lucia, Prof. David Perez-Garcia and Prof. Michael M. Wolf ([Bau+16]). The idea of size-driven phase transitions is Prof. Perez-Garcia’s, Prof. Wolf’s, and my supervisor’s. Both the periodic tiling and the Turing machine embedding are my ideas. Dr. Lucia and Prof. Perez-Garcia helped significantly with the thermal stability section.

Finally, chapter 5 is joint work with Dr. Elizabeth Crosson, and most ideas emerged from joint discussions ([BC16a]). The notion of non-uniform history states was introduced by me, as is the $\Omega(T^{-2})$ scaling proof both for the weighted case, as for Kitaev’s original construction (sidestepping the geometrical lemma). The limitations on further improvements is Dr. Crosson’s work.

All figures are my work, with the exception of fig. 2.10 (Ozols) and fig. 5.3 (Crosson).

1 Divisibility

Zeno's arguments about motion, which cause so much disquietude to those who try to solve the problems that they present, are four in number. The first asserts the nonexistence of motion on the ground that that which is in locomotion must arrive at the half-way stage before it arrives at the goal.
— *Aristotle, Physics*

People have pondered divisibility questions throughout most of western science and philosophy. Perhaps the earliest written mention of divisibility is in Aristotle's *Physics* in 350BC, in the form of the Arrow paradox—one of Zeno of Elea's paradoxes (ca. 490–430 BC). Aristotle's lengthy discussion of divisibility (he devotes an entire chapter to the topic) was motivated by the same basic question as more modern divisibility problems in mathematics: can the behaviour of an object—physical or mathematical—be subdivided into smaller parts?

For example, given a description of the evolution of a system over some time interval t , what can we say about its evolution over the time interval $t/2$? If the system is stochastic, this question finds a precise formulation in the *divisibility problem* for stochastic matrices [Kin62]: given a stochastic matrix \mathbf{P} , can we find a stochastic matrix \mathbf{Q} such that $\mathbf{P} = \mathbf{Q}^2$?

This question has many applications. For example, in information theory the stochastic matrices model noisy communication channels, and divisibility becomes important in relay coding, when signals must be transmitted between two parties where direct end-to-end communication is not available [Lor78]. Another direct use is in the analysis of chronic disease progression [CWGo8], where the transition matrix is based on sparse observations of patients, but finer-grained time-resolution is needed. In finance, changes in companies' credit ratings can be modelled using discrete time Markov chains, where rating agencies provide the transition matrix based on annual estimates—for valuation or risk analysis, a transition matrix for a much shorter time periods needs to be inferred [Jar97].

We can also ask about the evolution of the system for *all* times up to time t , i.e. whether the system can be described by some continuous evolution. For stochastic matrices, this has a precise formulation in the *embedding problem*: given a stochastic matrix \mathbf{P} , can we find a generator \mathbf{Q} of a continuous-time Markov

process such that $\mathbf{P} = \exp(\mathbf{Q}t)$? The embedding problem seems to date back further still, and was already discussed by Elfving in 1937 [Elf37]. Again, this problem occurs frequently in the field of systems analysis, and in analysis of experimental time-series snapshots [CEW12a]; [Lju87]; [NKL98].

Many generalisations of these divisibility problems have been studied in the mathematics and physics literature. For example, the question of square-roots of (entry-wise) nonnegative matrices is an old open problem in matrix analysis [Min88]: given an entry-wise nonnegative matrix \mathbf{M} , does it have an entry-wise nonnegative square-root? In quantum mechanics, the analogue of a stochastic matrix is a completely-positive trace preserving (CPTP) map, and the corresponding divisibility problem asks: when can a CPTP map \mathbf{T} be decomposed as $\mathbf{T} = \mathbf{R} \circ \mathbf{R}$, where \mathbf{R} is itself CPTP? The continuous version of this, whether a CPTP can be embedded into a completely-positive semi-group, is sometimes called the *Markovianity problem* in physics [CEW12b]—the latter again has applications to subdivision coding of quantum channels in quantum information theory [MRW15].

Instead of dynamics, we can also ask whether the description of the static state of a system can be subdivided into smaller, simpler parts. Once again, probability theory provides a rich source of such problems. The most basic of these is the classic topic of divisible distributions: given a random variable X , can it be decomposed into $X = Y + Z$ where Y, Z are some other random variables? What if Y and Z are identically distributed?

- If we instead ask for a decomposition into infinitely many random variables, this becomes the question of whether a distribution is infinitely divisible.

In this work, we address two of the most long-standing open problems on divisibility: divisibility of stochastic matrices, and divisibility and decomposability of probability distributions. We also extend our results to divisibility of nonnegative matrices and completely positive maps. Surprisingly little is known about the divisibility of stochastic matrices. Dating back to 1962 [Kin62], the most complete characterisation remains for the case of a 2×2 stochastic matrix [HGo3]. The infinite divisibility problem has recently been solved [CEW12b], but the finite case remains an open problem. Divisibility of random variables, on the other hand, is a widely-studied topic. Yet, despite first results dating back as far as 1934 [CW34], no

- general method of answering whether a random variable can be written as the sum of two or more random variables—whether distributed identically, or differently—is known.

We focus on the computational complexity of these divisibility problems. In each case, we show which of the divisibility problems have efficient solutions—for these, we give an explicit efficient algorithm. For all other cases, we prove reductions to the famous $\mathbf{P} = \mathbf{NP}$ -conjecture, showing that those problems are NP-hard. This essentially implies that—unless $\mathbf{P} = \mathbf{NP}$ —the geometry of the corresponding divisible and non-divisible is highly complex, and these sets have no simple characterisation beyond explicit enumeration. In particular, this shows that any future concrete classification of these NP-hard problems will be at least as hard as answering $\mathbf{P} = \mathbf{NP}$.

The following theorems summarise our main results on maps. Precise formulations and proofs can be

found in section 1.1.

Theorem 1.1. *Given a stochastic matrix \mathbf{P} , deciding whether there exists a stochastic matrix \mathbf{Q} such that $\mathbf{P} = \mathbf{Q}^2$ is NP-complete.*

Theorem 1.2. *Given a CPTP map \mathbf{B} , deciding whether there exists a CPTP map \mathbf{A} such that $\mathbf{B} = \mathbf{A} \circ \mathbf{A}$ is NP-complete.*

In fact, the last two theorems are strengthenings of the following result.

Theorem 1.3. *Given a nonnegative matrix \mathbf{M} , deciding whether there exists a nonnegative matrix \mathbf{N} such that $\mathbf{M} = \mathbf{N}^2$ is NP-complete.*

The following theorems summarise our main results on distributions. Precise formulations and proofs can be found in section 1.2.

Theorem 1.4. *Let X be a finite discrete random variable. Deciding whether X is n -divisible—i.e. whether there exists a random variable Y such that $X = \sum_{i=1}^n Y$ —is in P.*

Theorem 1.5. *Let X be a finite discrete random variable, and $\epsilon > 0$. Deciding whether there exists a random variable Y ϵ -close to X such that Y is n -divisible is in P.* •

Theorem 1.6. *Let X be a finite discrete random variable. Deciding whether X is decomposable—i.e. whether there exist random variables Y, Z such that $X = Y + Z$ —is NP-complete.*

Theorem 1.7. *Let X be a finite discrete random variable, and $\epsilon > 0$. Deciding whether there exists a random variable Y ϵ -close to X such that Y is decomposable is NP-complete.* •

It is interesting to contrast the results on maps and distributions. In the case of maps, the homogeneous 2-divisibility problems are already NP-hard, whereas finding an inhomogeneous decomposition is straightforward. For distributions, on the other hand, the homogeneous divisibility problems are efficiently solvable to all orders, but becomes NP-hard if we relax it to the inhomogeneous decomposibility problem.

This difference is even more pronounced for infinite divisibility. The infinite divisibility problem for maps is NP-hard (shown in [CEW12b]), whereas the infinite divisibility and decomposibility problems for distributions are computationally trivial, since indivisible and indecomposable distributions are both dense—see section 1.2.4.8 and 1.2.3.5.

The chapter is divided into two parts. We first address stochastic matrix and CPTP divisibility in section 1.1, obtaining results on entry-wise positive matrix roots along the way. Divisibility and decomposability of probability distributions is addressed in section 1.2. In both sections, we first give an overview of the history of the problem, stating previous results and giving precise definitions of the problems. We introduce the necessary notation at the beginning of each section, so that each section is largely self-contained.

 •

1.1 CPTP and Stochastic Matrix Divisibility

Mathematically, subdividing Markov chains is known as the finite divisibility problem. The simplest case is the question of finding a stochastic root of the transition matrix (or a CPTP root of a CPTP map in the quantum setting), which corresponds to asking for the evolution over half of the time interval. While the question of divisibility is rather simple to state mathematically, it is not clear a priori whether a stochastic matrix root for a given stochastic matrix exists at all. Historically, this has been a long-standing open question, dating back to at least 1962 [Kin62]. Matrix roots were also suggested early on in other fields, such as economics and general trade theory, at least as far back as 1967 [WA67], to model businesses and the flow of goods. Despite this long history, very little is known about the existence of stochastic roots of stochastic matrices. The most complete result to date is a full characterisation of 2×2 matrices, as given for example in [HGo3]. The authors mention that “...it is quite possible that we have to deal with the stochastic root problem on a case-by-case basis.” This already suggests that there might not be a simple mathematical characterisation of divisible stochastic matrices—meaning one that is simpler than enumerating the exponentially many roots and checking each one for stochasticity.

There are similarly few results if we relax the conditions on the matrix normalisation slightly, and ask for (entry-wise) nonnegative roots of (entry-wise) nonnegative matrices—for a precise formulation, see definitions 1.10 and 1.11. An extensive overview can be found in [Min88]. Following this long history of classical results, quantum channel divisibility recently gained attention in the quantum information literature. The foundations were laid in [WCo8], where the authors first introduced the notion of channel divisibility. A divisible quantum channel is a CPTP map that can be written as a nontrivial concatenation of two or more quantum channels.

A related question is to ask for the evolution under infinitesimal time steps, which is equivalent to existence of a logarithm of a stochastic matrix (or CPTP map) that generates a stochastic (resp. CPTP) semi-group. Classically, the question is known as Elfving’s problem or the embedding problem, and seems to date back even further than the finite case to 1937 [Elf37]. In the language of Markov chains, this corresponds to determining whether a given stochastic matrix can be embedded into an underlying continuous time Markov chain. Analogously, infinite quantum channel divisibility—also known as the Markovianity condition for a CPTP map—asks whether the dynamics of the quantum system can be described by a Lindblad master equation [Lin76]; [Gor76]. The infinite divisibility problems in both the classical and quantum case were recently shown to be NP-hard [CEW12b]. Formulated as weak membership problems, these results imply that it is NP-hard to extract dynamics from experimental data [CEW12a].

However, while related, it is not at all clear that there exists a reduction of the finite divisibility question to the case of infinite divisibility. In fact, mathematically, the infinite divisibility case is a special case of finite divisibility, as a stochastic matrix is infinitely divisible if and only if it admits an n^{th} root for all $n \in \mathbb{N}$

[Kin62].

The finite divisibility problem for stochastic matrices is still an open question, as are the nonnegative matrix and CPTP map divisibility problems. We will show that the question of existence of stochastic roots of a stochastic matrix is NP-hard. We also extend this result to (doubly) stochastic matrices, nonnegative matrices, and CPTP maps.

We start out by introducing the machinery we will use to prove theorem 1.3 and 1.1 in section 1.1.1. A reduction from the quantum to the classical case can be found in section 1.1.4, from the nonnegative to the stochastic case in section 1.1.5 and the main result—in a mathematically rigorous formulation—is then presented as theorem 1.34 in section 1.1.6. Note that in this chapter, in contrast to the rest of the thesis, we will not use bra-ket notation, as the results are more on the pure mathematical side, and keeping with convention, we denote vectors as e.g. e_i instead of $|e_i\rangle$.

1.1.1 Preliminaries

1.1.1.1 Roots of Matrices

In our study of matrix roots we restrict ourselves to the case of square roots. The more general case of p^{th} roots of matrices remains to be discussed. We will refer to square roots simply as roots. To be explicit, we state the following definition.

Definition 1.8. Let $\mathbf{M} \in \mathbb{K}^{d \times d}$, $d \in \mathbb{N}$, where \mathbb{K} is a field which we will always take to be either \mathbb{R} or \mathbb{C} . Then we say that $\mathbf{R} \in \mathbb{K}^{d \times d}$ is a root of \mathbf{M} if $\mathbf{R}^2 = \mathbf{M}$. We denote the set of all roots of \mathbf{M} with $\sqrt{\mathbf{M}}$.

Following the theory of matrix functions—see for example [Hig87]—we remark that in the case of nonsingular \mathbf{M} , $\sqrt{\mathbf{M}}$ is nonempty and can be expressed in Jordan normal form via $\sqrt{\mathbf{M}} = \mathbf{Z}\mathbf{J}\mathbf{Z}^{-1}$ for some invertible \mathbf{Z} , where $\mathbf{J} = \text{diag}(\mathbf{J}_1^\pm, \dots, \mathbf{J}_m^\pm)$. Here \mathbf{J}_i^\pm denotes the \pm -branch of the root function $f(x) = \sqrt{x}$ of the Jordan block corresponding to the i^{th} eigenvalue λ_i ,

$$\mathbf{J}_i^\pm = \begin{pmatrix} \pm f(\lambda_i) & \pm f'(\lambda_i)/1! & \dots & \pm f^{(m_i-1)}(\lambda_i)/(m_i-1)! \\ 0 & \pm f(\lambda_i) & \ddots & \vdots \\ \vdots & \ddots & \ddots & \pm f'(\lambda_i)/1! \\ 0 & \dots & 0 & \pm f(\lambda_i) \end{pmatrix}.$$

In particular, every such Jordan block has precisely two choices for the square root branch, which can be chosen individually for each block (see [Hig87, lem. 1]). If \mathbf{M} is diagonalisable, \mathbf{J} simply reduces to the canonical diagonal form $\mathbf{J} = \text{diag}(\pm\sqrt{\lambda_1}, \dots, \pm\sqrt{\lambda_m})$, where again the signs can be chosen independently for each (nonzero) eigenvalue.

The sign choices are in fact individual.

If \mathbf{M} is derogatory—i.e. there exist multiple Jordan blocks sharing the same eigenvalue λ —it has continuous families of so-called *nonprimary* roots $\sqrt{\mathbf{M}} = \mathbf{Z}\mathbf{U}\mathbf{J}\mathbf{U}^{-1}\mathbf{Z}^{-1}$, where \mathbf{U} is an arbitrary nonsingular matrix that commutes with the Jordan normal form $[\mathbf{U}, \mathbf{J}] = 0$.

Can we always find roots of matrices? For non-singular \mathbf{M} , or singular \mathbf{M} where the algebraic multiplicity of the kernel equals its geometric multiplicity, the answer is always yes (see [HLII, sec. 1.5]). For the sake of completeness, and although we will not need it in this chapter, we mention the following theorem which

- allows us to establish whether a square root of an arbitrary complex matrix exists in general.

Theorem 1.9 ([HLII]). *Let $\mathbf{M} \in \mathbb{C}^{d \times d}$ have the Jordan canonical form $\mathbf{Z}\mathbf{\Lambda}\mathbf{Z}^{-1}$, where $\mathbf{\Lambda} = \text{diag}(\mathbf{J}_0, \mathbf{J}_1)$, such that \mathbf{J}_0 collects all Jordan blocks corresponding to the eigenvalue 0, and \mathbf{J}_1 collects the remaining ones. Assume further that*

$$d_i := \dim(\ker \mathbf{M}^i) - \dim(\ker \mathbf{M}^{i-1})$$

- *has the property that for all $i \in \mathbb{N}_{>0}$, no more than one element of the sequence satisfies $d_i = 2i + 1$. Then $\sqrt{\mathbf{M}} = \mathbf{Z}\sqrt{\mathbf{\Lambda}}\mathbf{Z}^{-1}$, where $\sqrt{\mathbf{\Lambda}} = \text{diag}(\sqrt{\mathbf{J}_0}, \sqrt{\mathbf{J}_1})$.*

If \mathbf{M} is a real matrix, a similar theorem holds and there exist various numerical algorithms for calculating real square roots, see for example [Hig87].

1.1.1.2 Roots of Stochastic Matrices

Remember the following two definitions.

- Definition 1.10. *A matrix $\mathbf{M} \in \mathbb{K}^{d \times d}$ is said to be nonnegative if $0 \leq \mathbf{M}_{ij} \forall i, j = 1, \dots, d$.*

Definition 1.11. *A matrix $\mathbf{Q} \in \mathbb{K}^{d \times d}$ is said to be stochastic if it is nonnegative and $\sum_{k=1}^d \mathbf{Q}_{ik} = 1 \forall i =$*

- *$1, \dots, d$.*

In contrast to finding a general root of a matrix, very little is known about the existence of nonnegative roots of nonnegative matrices—or stochastic roots of stochastic matrices—if $d \geq 3$. For stochastic matrices

- and in the case $d = 2$, a complete characterisation can be given explicitly, and for $d \geq 3$, all stochastic roots that are functions of the original matrix are known, as demonstrated in [HGo3]. Further special classes of matrices for which a definite answer exists can be found in [HLII]. But even for $d = 3$, the general case is still an open question—see [LinII, ch. 2.3] for details.

Indeed, a stochastic matrix may have no stochastic root, a primary or nonprimary root—or both. To make things worse, if a matrix has a p^{th} stochastic root, it might or might not have a q^{th} stochastic root if $p \nmid q$ — p is not a divisor of q —, $q > p$ or $q \nmid p$, $q < p$.

A related open problem is the inverse eigenspectrum problem, as described in the extensive overview in [ELNo4]. While the sets $\Omega_n \subset \mathbb{D}$ —denoting all the possible eigenvalues of an n -dimensional nonnegative

matrix—can be given explicitly, and hence also Ω_n^p , almost nothing is known about the sets for the entire eigenspectrum. Any progress in this area might yield necessary conditions for the existence of stochastic roots.

In recent years, some approaches have been developed to approximate stochastic roots numerically, see the comments in [HG03, sec. 4]. Unfortunately, most algorithms are highly unstable and do not necessarily converge to a stochastic root. A direct method using nonlinear optimisation techniques is difficult and depends heavily on the algorithm employed [Lin11].

It remains an open question whether there exists an efficient algorithm that decides whether a stochastic matrix \mathbf{Q} has a stochastic root.

In this chapter, we will prove that this question is NP-hard to answer. •

1.1.1.3 The Choi Isomorphism

For the results on CPTP maps, we will need the following basic definition and results.

Definition 1.12. *Let $\mathbf{A} : \mathcal{H} \rightarrow \mathcal{H}$ be a linear map on $\mathcal{H} = \mathbb{C}^{d \times d}$. We say that \mathbf{A} is positive if for all Hermitian and positive definite $\rho \in \mathcal{H}$, $\mathbf{A}\rho$ is Hermitian and positive definite. It is said to be completely positive if $\mathbf{A} \otimes \mathbb{1}_n$ is positive $\forall n \in \mathbb{N}$.*

A map \mathbf{A} which is completely positive and trace-preserving—i.e. $\text{tr}(\mathbf{A}\rho) = \text{tr} \rho \forall \rho \in \mathcal{H}$ —is called a completely positive trace-preserving map, or short CPTP map.

In contrast to positivity, complete positivity is easily characterised using the well-known Choi-Jamiolkowski isomorphism—see [Cho75, Th. 2].

Remark 1.13. *Let the notation be as in definition 1.12 and pick a basis e_1, \dots, e_d of \mathbb{C}^d . Then \mathbf{A} is completely positive if and only if the Choi matrix*

$$\mathbf{C}_{\mathbf{A}} := (\mathbb{1}_d \otimes \mathbf{A})\Omega\Omega^T = \sum_{i,j=1}^d e_i e_j^T \otimes \mathbf{A}(e_i e_j^T)$$

is positive semidefinite, where $\Omega := \sum_{i=1}^d e_i \otimes e_i$.

The condition of trace-preservation then translates to the following.

Remark 1.14. *A map \mathbf{A} is trace-preserving if and only if $\text{tr}_2(\mathbf{C}_{\mathbf{A}}) = \mathbb{1}_d$, where tr_2 denotes the partial trace over the second pair of indices.*

1.1.2 NP-Toolbox

1.1.2.1 Boolean Satisfiability Problems

Definition 1.15 (1-IN-3SAT).

Instance: n_v Boolean variables m_1, \dots, m_{n_v} and n_c clauses $R(m_{i_1}, m_{i_2}, m_{i_3})$ where $i = 1, \dots, n_c$, usually denoted as a 4-tuple $(n_v, n_c, \{m_i\}_{1 \leq i \leq n_v}, \{m_{ij}\}_{1 \leq i \leq n_c, 1 \leq j \leq 3})$, where each $m_{ij} \in \{m_1, \dots, m_{n_v}\}$

- denotes the three Boolean variables occurring in the i^{th} clause. The Boolean operator R satisfies

$$R(a, b, c) = \begin{cases} \text{TRUE} & \text{if exactly one of } a, b \text{ or } c \text{ is TRUE} \\ \text{FALSE} & \text{otherwise.} \end{cases}$$

Question: Does there exist a truth assignment to the Boolean variables such that every clause contains exactly one true variable?

- Note that we can without loss of generality assume that none of the R terms contains negations. In order to translate a given 1-IN-3SAT instance to this case, assume one is given a term $R(a, b, \neg c)$. By adding an auxiliary variable c' and a term $R(c, c, c')$, c is forced to be negative in order for the clause to be satisfiable.

1.1.2.2 Subset Sum Problems

We start out with the following variant of a well-known NP-complete problem—see for example [GJ79] for a reference.

Definition 1.16 (SUBSET SUM, VARIANT).

- **Instance.** Multiset S of integer or rational numbers, $l \in \mathbb{Q}$.

Question. Does there exist a multiset $T \subsetneq S$ such that $|\sum_{t \in T} t - \sum_{s \in S \setminus T} s| < l$?

From the definition, we immediately observe the following rescaling property.

- Corollary 1.17. Let $a \in \mathbb{Q} \setminus \{0\}$ and (S, l) a SUBSET SUM instance. Then $\text{SUBSET SUM}(S, l) = \text{SUBSET SUM}(aS, |a|l)$.

We define two variants of SUBSET SUM where we demand one of the subsets to contain a certain number of elements.

Definition 1.18 (SUBSET SUM _{m} , $m \in \mathbb{Z}$).

Instance. Multiset S of reals with $|S|$ even, $l \in \mathbb{R}$.

Question. Does there exist a multiset $T \subsetneq S$ with $|T| = m$ and such that $|\sum_{t \in T} t - \sum_{s \in S \setminus T} s| < l$?

Definition 1.19 (SIGNED SUBSET SUM_m).

Instance. Multiset S of positive integers or reals, $x, y \in \mathbb{R} : x \leq y$.

Question. Does there exist a multiset $T \subset S$ with $|T| = m$ and such that $x < \sum_{t \in T} t - \sum_{s \in S \setminus T} s < y$?

We observe the following.

Lemma 1.20. $SUBSET\ SUM_m \leftarrow SUBSET\ SUM$.

Proof. If (S, l) is a SUBSET SUM instance, then

$$SUBSET\ SUM(S, l) = \bigvee_{m=1}^{|S|} SUBSET\ SUM_m(S, l). \quad \square$$

Remark 1.21. It is clear that $SUBSET\ SUM_m(S, l) = FALSE$ for $|S| \leq m \vee 0 \geq m$. Furthermore, $SUBSET\ SUM_m(S, l) = SUBSET\ SUM_{|S|-m}(S, l)$.

Observe that this remark indeed makes sense, as SUBSET SUM₀ should give FALSE, which is the desired outcome for $m = |S|$.

We define yet another variant of SUBSET SUM, which captures the special case when we require both subsets to have the same number of elements. •

Definition 1.22 (EVEN SUBSET SUM).

Instance. Multiset S of reals with $|S|$ even, $l \in \mathbb{R}$.

Question. Does there exist a multiset $T \subsetneq S$ with $|T| = |S|/2$ and such that $|\sum_{t \in T} t - \sum_{s \in S \setminus T} s| < l$?

We further reduce SUBSET SUM to EVEN SUBSET SUM.

Lemma 1.23. $EVEN\ SUBSET\ SUM \leftarrow SUBSET\ SUM$.

Proof. Let (S, l) be an SUBSET SUM instance. Define $S' := S \cup \{0, \dots, 0\} : |S'| = 2|S|$. Then if $EVEN\ SUBSET\ SUM(S', l) = TRUE$, we know that there exists $T' \subset S' : |\sum_{t \in T'} t - \sum_{s \in S' \setminus T'} s| < l$. Let then $T := T'$ without the 0s. It is obvious that then $|\sum_{t \in T} t - \sum_{s \in S \setminus T} s| < l$. The FALSE case reduces analogously, hence the claim follows. □

For EVEN SUBSET SUM, we generalise corollary 1.17 to the following scaling property.

Lemma 1.24. Let $a \in \mathbb{Q} \setminus \{0\}$, $c \in \mathbb{Q}$, and (S, l) an EVEN SUBSET SUM instance. Then $EVEN\ SUBSET\ SUM(S, l) = EVEN\ SUBSET\ SUM(aS + c, |a|l)$, where addition and multiplication is defined element-wise. •

Proof. Straightforward, since we require $|S| = 2|T| = 2|S \setminus T|$. □

For definition 1.19, we finally show

Lemma 1.25. $SIGNED\ SUBSET\ SUM_m \longleftarrow SUBSET\ SUM_m$.

Proof. Immediate from $SIGNED\ SUBSET\ SUM_m(S, -l, l) = SUBSET\ SUM_m(S, l)$. □

1.1.2.3 Partition Problems

Another well-known NP-complete problem which will come into play in the proof of theorem 1.88 is set partitioning.

Definition 1.26 (PARTITION).

- **Instance.** *Multiset A of positive integers or rationals.*

Question. *Does there exist a multiset $T \subsetneq A$ with $\sum_{t \in T} t = \sum_{s \in A \setminus T} s$?*

Lemma 1.27. *For the special case of SUBSET SUM with instance (S, l) , where the bound $l \in \mathbb{R}$ equals the total sum of the instance numbers $l = \sum_{s \in S} s$, we obtain the equivalence $SUBSET\ SUM(\cdot, \Sigma) \longleftrightarrow PARTITION(\cdot)$.*

Proof. Let S be the multiset of a SUBSET SUM instance (S, l) , where we assume without loss of generality that all $S \ni s \geq 0$. Now first assume $\Sigma_S = 0$. In that case the claim follows immediately, since the problems are identical.

Without loss of generality, we can thus assume $\Sigma_S > 0$ and consider the set $S' := S \cup \{-\Sigma_S/2, -\Sigma_S/2\}$, such that $\Sigma_{S'} = 0$.

If now $SUBSET\ SUM(S', 0) = \text{TRUE}$, we know that there exists $T \subsetneq S' : |\sum_{t \in T} t - \sum_{s \in S' \setminus T} s| = 0$. Now assume T contains both copies of $-\Sigma_S/2$. Then clearly

$$|\sum_{t \in T} t - \sum_{s \in S' \setminus T} s| = |-\Sigma_S + \sum_{t \in T \setminus \{-\Sigma_S\}} t - \sum_{s \in S \setminus T} s| > 0,$$

since $|S \setminus T| > 0$. The same argument shows that exactly one $-\Sigma_S/2 \in T, S \setminus T$, and hence $PARTITION(S) = \text{TRUE}$.

On the other hand, if $PARTITION(S) = \text{TRUE}$, then it immediately follows that $SUBSET\ SUM(S, \Sigma_S) = \text{TRUE}$. □

Finally observe the following extension of lemma 1.27.

Lemma 1.28. *Let $\epsilon > 0$, f a polynomial. Then $SUBSET\ SUM(\cdot, \Sigma + f(\epsilon)) \longleftrightarrow PARTITION(\cdot)$.*

Proof. The proof is the same as for lemma 1.27, but we consider $S \cup \{-\Sigma_S/2 - f(\epsilon)/2, -\Sigma_S/2 - f(\epsilon)/2\}$ instead. □

1.1.3 Equivalence of Computational Questions

In the following we denote with S some arbitrary finite index set, not necessarily the same for all problems. We begin by defining the following decision problems.

Definition 1.29 (CPTP DIVISIBILITY).

Instance. CPTP map $\mathbf{B} \in \mathbb{Q}^{d \times d}$.

Question. Does there exist a CPTP map $\mathbf{A} : \mathbf{A}^2 = \mathbf{B}$?

Definition 1.30 (CPTP ROOT).

Instance. Family of matrices $(\mathbf{A}_s)_{s \in S}$ that comprises all the roots of a matrix \mathbf{B} .

Question. Does there exist an $s \in S : \mathbf{A}_s$ is a CPTP map?

Definition 1.31 (STOCHASTIC DIVISIBILITY).

Instance. Stochastic matrix $\mathbf{P} \in \mathbb{Q}^{d \times d}$.

Question. Does there exist a stochastic matrix $\mathbf{Q} : \mathbf{Q}^2 = \mathbf{P}$?

Definition 1.32 (STOCHASTIC ROOT).

Instance. Family of matrices $(\mathbf{Q}_s)_{s \in S}$ comprising all the roots of a matrix \mathbf{P} .

Question. Does there exist an $s \in S : \mathbf{Q}_s$ stochastic?

Definition 1.33 (NONNEGATIVE ROOT).

Instance. Family of matrices $(\mathbf{M}_s)_{s \in S}$ comprising all the roots of a matrix \mathbf{N} , where all \mathbf{M}_s have at least one positive entry.

Question. Does there exist an $s \in S : \mathbf{M}_s$ nonnegative?

We want to point out that the input size for the ROOT problems is at most polynomial in the binary size of the index set S , i.e. $\text{poly } \log |S|$; in particular, the families are *not* given as an explicit list of $|S|$ matrices (which would correspond to an $\Omega(|S|)$ instance size). This is a natural assumption: since the index set is finite but we demand that the family lists *all* roots, we know that no degenerate (continuous families of) roots exist, so we are in the realm where the matrices are non-degenerate (see the discussion below definition 1.8). We reduce the general case to the non-degenerate case in lemma 1.47. This allows us to state the following central theorem. •

Theorem 1.34. *The reductions as shown in fig. 1.1 hold.*

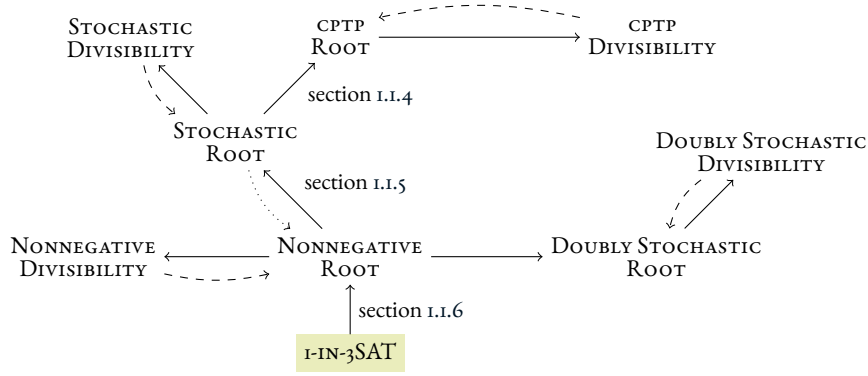


Figure 1.1: Complete chain of reduction for our programs. The dashed line between the DIVISIBILITY and ROOT problems holds for non-derogatory matrices, respectively. The dotted line between STOCHASTIC ROOT and NONNEGATIVE ROOT holds only for irreducible matrices. The doubly stochastic and nonnegative branch are included for completeness but not described in detail here—see corollary 1.41.

Proof outline. We give here a high-level overview over the line of reduction, and refer the reader to the

- corresponding rigorous proofs later in the chapter.

The implication STOCHASTIC DIVISIBILITY \leftarrow STOCHASTIC ROOT is immediate, and we rigorously state it here. If \mathbf{P} is not stochastic, the answer is negative. If it is stochastic, we can apply STOCHASTIC DIVISIBILITY. The opposite direction holds for non-derogatory stochastic \mathbf{P} : in this case we can enumerate all roots of \mathbf{P} as a finite family which forms a valid instance for STOCHASTIC ROOT.

The reduction STOCHASTIC ROOT \leftarrow NONNEGATIVE ROOT can be resolved by lemmas 1.39 and 1.40—we construct a family of matrices $(\mathbf{Q}_s)_{s \in S}$ that contains a stochastic root iff $(\mathbf{M}_s)_{s \in S}$ contains a nonnegative root. The result then follows from applying STOCHASTIC ROOT.

- A special case is when our stochastic matrix \mathbf{P} is irreducible, i.e. when it can be written as

$$\mathbf{S}\mathbf{P}\mathbf{S}^T = \begin{pmatrix} \mathbf{P}_{11} & \mathbf{P}_{12} \\ 0 & \mathbf{P}_{22} \end{pmatrix}$$

under a permutation \mathbf{S} , and such that \mathbf{P}_{11} and \mathbf{P}_{22} are square and nonzero. In this case, any nonnegative root $\mathbf{Q}_{s'} : \mathbf{Q}_{s'}^2 = \mathbf{P}$ is stochastic, and thus STOCHASTIC ROOT \leftarrow NONNEGATIVE ROOT—see [HL11, sec. 3, sec. B.11] for details.

The link CPTP DIVISIBILITY \leftarrow CPTP ROOT again needs the following intermediate step. If \mathbf{A} is not CPTP, the answer is negative. If it is CPTP, then we can apply CPTP DIVISIBILITY. Similarly, if \mathbf{A} is non-derogatory, the reduction works in the opposite direction as well.

The direction CPTP ROOT \leftarrow STOCHASTIC ROOT follows from corollary 1.38. We start out with a family $(\mathbf{Q}_s)_{s \in S}$ comprising all the roots of a stochastic matrix \mathbf{P} . Then we define an embedding map to

define a new family of CPTP maps $(\mathbf{A}_s := \text{emb } \mathbf{Q}_s)_{s \in S}$ (see definition 1.36) —this family then comprises all of the roots of $\mathbf{B} := \mathbf{A}_k^2 \equiv \mathbf{A}_s^2 \forall k, s$. Furthermore, by lemma 1.37, there exists a CPTP \mathbf{A}_s if and only if there exists a stochastic \mathbf{Q}_s , and the reduction follows. •

Finally, we can extend our reduction to the programs DOUBLY STOCHASTIC ROOT and DOUBLY STOCHASTIC DIVISIBILITY as well as NONNEGATIVE DIVISIBILITY, defined analogously, see our comment in corollary 1.41 and the complete reduction tree in fig. 1.1.

The heavy lifting is done in the reduction from 1-IN-3SAT to NONNEGATIVE ROOT, which we address in section 1.1.6. □ •

At this point, we observe the following fact.

Lemma 1.35. *All the above DIVISIBILITY and ROOT problems in definition 1.29 to 1.33 are contained in NP.*

Proof. It is straightforward to come up with a witness and a verifier circuit that satisfies the definition of the decision class NP. For example in the CPTP case, a witness is a matrix root that can be checked to be a CPTP map using remark 1.13 and squared in polynomial time, which is the verifier circuit. Both circuit and witness are clearly poly-sized and hence the claim follows. □

By encoding an instance of 1-IN-3SAT into a family of nonnegative matrices $(\mathbf{M}_s)_{s \in S}$, we show the implication 1-IN-3SAT \rightarrow NONNEGATIVE ROOT and 1-IN-3SAT \rightarrow (DOUBLY) STOCHASTIC/CPTP DIVISIBILITY, accordingly, from which NP-hardness of (DOUBLY) STOCHASTIC/CPTP DIVISIBILITY follows. The entire chain of reduction can be seen in fig. 1.1.

1.1.4 Reduction of STOCHASTIC ROOT to CPTP ROOT

This reduction is based on the following embedding.

Definition 1.36. *Let $\{e_i\}$ be an orthonormal basis of \mathbb{K}^d . The embedding emb is defined as*

$$\text{emb} : \mathbb{K}^{d \times d} \hookrightarrow \mathbb{K}^{d^2 \times d^2},$$

$$\mathbf{A} \mapsto \mathbf{B} := \sum_{i,j=1}^d \mathbf{A}_{ij} (e_i \otimes e_i)(e_j \otimes e_j)^T = \sum_{i,j=1}^d \mathbf{A}_{ij} (e_i e_j^T) \otimes (e_i e_j^T).$$

We observe the following.

Lemma 1.37. *We use the same notation as in remark 1.13. Let $\mathbf{A} \in \mathbb{K}^{d \times d}$ and $\mathbf{B} := \text{emb } \mathbf{A}$. Then \mathbf{A} is positive (nonnegative) if and only if the Choi matrix $\mathbf{C}_\mathbf{B}$ is positive (semi-)definite. Furthermore, the row sums of \mathbf{A} are 1—i.e. $\sum_{j=1}^d \mathbf{A}_{ij} = 1 \forall j = 1, \dots, d$ —if and only if $\text{tr}_2(\mathbf{C}_\mathbf{B}) = \mathbb{1}_d$. In addition, the spectrum of \mathbf{B} satisfies $\sigma(\mathbf{B}) \subseteq \{\mathbf{A}_{ij}\} \cup \{0\}$. •*

- Proof.* The first claim follows directly from the matrix representation of our operators. There, the Choi isomorphism is manifest as the reshuffling operation

$$\cdot^\Gamma : \mathbb{K}^{d^2 \times d^2} \longrightarrow \mathbb{K}^{d^2 \times d^2}, [(e_i e_j^\top) \otimes (e_i e_j^\top)]^\Gamma \longmapsto (e_i e_i^\top) \otimes (e_j e_j^\top). \quad (1.1)$$

For more details, see e.g. [BŽo6].

The second statement follows from

$$\begin{aligned} \text{tr}_2(C_{\mathbf{B}}) &= \text{tr}_2 \left(\sum_{i,j=1}^d \mathbf{A}_{ij} (e_i e_j^\top) \otimes (e_i e_j^\top) \right) \\ &= \sum_{i,j=1}^d \mathbf{A}_{ij} e_i e_i^\top = \text{diag} \left(\sum_{j=1}^d \mathbf{A}_{1j}, \dots, \sum_{j=1}^d \mathbf{A}_{dj} \right). \end{aligned}$$

The final claim is trivial. □

This remark immediately yields the following consequence.

Corollary 1.38. *For a family of stochastic matrices $(\mathbf{Q}_s)_{s \in S}$ parametrised by the index set S , there exists a family of square matrices $(\mathbf{A}_s)_{s \in S} := (\text{emb } \mathbf{Q}_s)_{s \in S}$, such that $(\mathbf{Q}_s)_{s \in S}$ contains a stochastic matrix if and only if $(\mathbf{A}_s)_{s \in S}$ contains a CPTP matrix.*

1.1.5 Reduction of NONNEGATIVE ROOT to STOCHASTIC ROOT

The difference between NONNEGATIVE ROOT and STOCHASTIC ROOT is the extra normalisation condition in the latter, see definition 1.11. The following two lemmas show that this normalisation does not pose an issue, so we can efficiently reduce the problem NONNEGATIVE ROOT to STOCHASTIC ROOT.

Lemma 1.39. *For a family of square matrices $(\mathbf{M}_s)_{s \in S}$ parametrised by the index set S , all of which with at least one positive entry, there exists a family of square matrices $(\mathbf{Q}_s)_{s \in S}$ such that $(\mathbf{M}_s)_{s \in S}$ contains a nonnegative matrix if and only if $(\mathbf{Q}_s)_{s \in S}$ contains a stochastic matrix and such that $\text{rank } \mathbf{Q}_s = \text{rank } \mathbf{M}_s + 2 \forall s \in S$. Furthermore, $(\mathbf{Q}_s)_{s \in S}$ can be constructed efficiently from $(\mathbf{M}_s)_{s \in S}$.*

Proof. We explicitly construct our family $(\mathbf{Q}_s)_{s \in S}$ as follows. Pick an $s \in S$ and denote $\mathbf{M} := \mathbf{M}_s$. Let d be the dimension of \mathbf{M} . We first pick $a \in \mathbb{R}^+$ such that $a \max_{ij} \mathbf{M}_{ij} = 1/2$ (the value of $1/2$ is arbitrary,

and any finite number $\leq 43/81$ works here), and define

$$\begin{aligned}\mathbf{Q}_s &:= \frac{1}{1764d} \begin{pmatrix} 1764a\mathbf{M} + 637 & 735 - 1260a\mathbf{M} & 392 - 504a\mathbf{M} \\ 735 - 1260a\mathbf{M} & 900a\mathbf{M} + 1029 & 360a\mathbf{M} \\ 392 - 504a\mathbf{M} & 360a\mathbf{M} & 144a\mathbf{M} + 1372 \end{pmatrix} \\ &\equiv \frac{a}{d}AA^T \otimes \mathbf{M} + \frac{1}{d}(BB^T + CC^T) \otimes \mathbf{1},\end{aligned}$$

where by sum of matrix \mathbf{M} and scalar x we mean $\mathbf{M} + x\mathbf{1}$. $\mathbf{1} := (1)_{1 \leq i, j \leq d} \in \mathbb{R}^{d \times d}$, and

$$A := \left(1, -\frac{5}{7}, -\frac{2}{7}\right)^T, \quad B := \left(\frac{1}{6}, \frac{1}{2}, -\frac{2}{3}\right)^T, \quad C := -\frac{1}{\sqrt{3}}(1, 1, 1)^T.$$

Observe that $\{A, B, C\}$ form an orthogonal set—if one wishes, normalising and pulling out the constant as eigenvalue to the corresponding eigenprojectors would work equally well.

By construction, \mathbf{Q}_s is nonnegative if and only if \mathbf{M}_s is. Since the row sums of \mathbf{Q}_s are always 1, \mathbf{Q}_s is stochastic if and only if \mathbf{M}_s is nonnegative, and the claim follows. \square

Lemma 1.40. *Let the notation be as in lemma 1.39 and write $\sqrt{\mathbf{N}}$ for the set of roots of \mathbf{N} , see definition 1.8. Assume $(\mathbf{M}_s)_{s \in S} = \sqrt{\mathbf{N}}$ for some $\mathbf{N} \in \mathbb{C}^{d \times d}$. Then there exists a $\mathbf{P} \in \mathbb{C}^{d \times d}$, such that $\mathbf{Q}_s^2 = \mathbf{P} \forall s \in S$ and $(\mathbf{Q}_s)_{s \in S} \subset \sqrt{\mathbf{P}}$. Furthermore, the complement of $(\mathbf{Q}_s)_{s \in S}$ in $\sqrt{\mathbf{P}}$ does not contain any stochastic roots.*

Proof. The first statement is obvious, since for all $s \in S$,

$$\mathbf{Q}_s^2 = \frac{a^2}{d^2} \frac{78}{49} AA^T \otimes \mathbf{M}_s^2 + \frac{1}{d} \left(\frac{13}{18} BB^T + CC^T \right) \otimes \mathbf{1} =: \mathbf{P},$$

and hence clearly $(\mathbf{Q}_s)_{s \in S} \subset \sqrt{\mathbf{P}}$.

The last statement is not quite as straightforward—it is the main reason our carefully crafted matrix \mathbf{Q}_s has its slightly unusual shape. All possible roots of \mathbf{P} are of the form

$$\sqrt{\mathbf{P}} = \frac{a}{d} AA^T \otimes \sqrt{\mathbf{N}} \pm \frac{1}{d} (BB^T \pm CC^T) \otimes \mathbf{1}.$$

It is easy to check that none of the other sign choices yields any stochastic matrix, so the claim follows.¹ \square

Corollary 1.41. *The results of lemma 1.39 and 1.40 also hold for doubly stochastic matrices—observe that our construction of \mathbf{Q}_s is already doubly stochastic.*

¹The reader may try to find a simpler matrix that does the trick.

1.1.6 Reduction of 1-IN-3SAT to NONNEGATIVE ROOT

We now embed an instance of a Boolean satisfiability problem, 1-IN-3SAT—see definition 1.15 for details—into a family of matrices $(\mathbf{M}_s)_{s \in S}$ in a way that there exists an s such that \mathbf{M}_s is nonnegative if and only if the instance of 1-IN-3SAT is satisfiable. The construction is inspired by the one in [CEW12b].

We identify

$$\text{TRUE} \longleftrightarrow 1, \quad \text{FALSE} \longleftrightarrow -1. \quad (1.2)$$

Denote with $(m_{i1}, m_{i2}, m_{i3}) \in \{\pm 1\}^3$ the three Boolean variables occurring in the i^{th} Boolean clause, and let $m_i \in \{\pm 1\}$ stand for the single i^{th} Boolean variable. Then 1-IN-3SAT translates to the inequalities

$$-\frac{3}{2} \leq m_{i1} + m_{i2} + m_{i3} \leq -\frac{1}{2} \quad \forall i = 1, \dots, n_c. \quad (1.3)$$

- **Theorem 1.42.** *Let $(n_v, n_c, \{m_i\}, \{m_{ij}\})$ be a 1-IN-3SAT instance. Then there exists a family of matrices $(\mathbf{M}_s)_{s \in S}$ such that $\exists s : \mathbf{M}_s$ nonnegative iff the instance is satisfiable.*

To prove this, we first need the following technical lemma.

- **Lemma 1.43.** *Let $(n_v, n_c, \{m_i\}, \{m_{ij}\})$ be a 1-IN-3SAT instance. Then there exists a family of matrices $(\mathbf{C}_s)_{s \in S}$ such that $\exists s : \text{the first } n_c \text{ on-diagonal } 4 \times 4 \text{ blocks of } \mathbf{C}_s \text{ are nonnegative iff the instance is satisfiable. In addition, we have } \mathbf{C}_s^2 = \mathbf{C}_t^2 \forall s, t. \text{ Furthermore, } (\mathbf{C}_s)_{s \in S} \subset \sqrt{\mathbf{C}_s^2}, \text{ and the complement contains no nonnegative root.}$*
- **Proof.** We begin with a simple example. Consider the matrix

$$\mathbf{A} := \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \equiv (1+i) \begin{pmatrix} 1 & -i \\ i & 1 \end{pmatrix} + (1-i) \begin{pmatrix} 1 & i \\ -i & 1 \end{pmatrix}.$$

\mathbf{A} is clearly non-degenerate; if we take the square of \mathbf{A} , and then list its four square-roots, we obtain the four matrices

$$a(1+i) \begin{pmatrix} 1 & -i \\ i & 1 \end{pmatrix} + b(1-i) \begin{pmatrix} 1 & i \\ -i & 1 \end{pmatrix} = \begin{pmatrix} (1+i)a + (1-i)b & (1-i)a + (1+i)b \\ (-1+i)a - (1+i)b & (1+i)a + (1-i)b \end{pmatrix},$$

where $a, b \in \{\pm 1\}$. Like this, we can enforce conditions on the choice of signs for a, b : if a and b have opposite sign, for instance, this matrix will be complex-valued; for $a = b = 1$ we recover \mathbf{A} , and for $a = b = -1$ we obtain $-\mathbf{A}$. By combining a list of such matrix blocks with orthogonal projectors, we can encode more complicated conditions on the sign choices, and obtain an overall matrix with a list of roots such that precisely one of the roots is non-negative if and only if the encoded 1-IN-3SAT instance—where

the boolean variables correspond to the sign choices—is satisfiable.

For every Boolean variable m_k , define a vector $v_k \in \mathbb{R}^d$ such that their first n_c elements are defined via

$$(v_k)_i := \begin{cases} 1 & m_k \text{ occurs in } i^{\text{th}} \text{ clause} \\ 0 & \text{otherwise.} \end{cases}$$

We will specify the dimension d later—obviously $d \geq n_c$, and the free entries are used to orthonormalise all vectors in the end. For now, we denote the orthonormalisation region with \vec{o} . We further define the vectors $c_1, c_2 \in \mathbb{R}^d$ to have all 1s in the first n_c entries, i.e. $c_{1,2} = (1, \dots, 1, \vec{o}_{1,2})$. Let then

$$\begin{aligned} \mathbf{C}'_s := & c_1 c_1^T \otimes \begin{pmatrix} 1 & 1 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} + \frac{1}{2} c_2 c_2^T \otimes \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \\ & + \sum_{k=1}^{n_v} p_k v_k v_k^T \otimes \begin{pmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}. \end{aligned} \quad (1.4)$$

The subscript s stands for a specific choice of sign for every eigenvalue in this matrix; in this case, all signs are positive, but taking the square of \mathbf{C}'_s and then the square-root leaves this sign choice free for each eigenvalue.

The variables p_k denote a specific choice of the rescaled Boolean variables m_i , which—in order to avoid degeneracy—have to be distinct, e.g. via

$$p_k = \left(1 - \frac{1}{N} - \frac{k}{N n_v}\right) m_k \quad \forall k = 1, \dots, n_v. \quad (1.5)$$

The p_{ij} are defined accordingly from the m_{ij} and $N \in \mathbb{N}$ is large but fixed.

To exemplify this embedding, consider the k^{th} 4×4 on-diagonal block corresponding to the k^{th} Boolean clause involving the variables p_{k1}, p_{k2} and p_{k3} . All 4×4 matrices in eq. (1.4) are rank 2 and thus provide us with two independent sign choices in front of the two projectors comprising it: we denote them with $c_{11}, c_{12} \in \{\pm\sqrt{2}\}$ for the first projector, and $c_{21}, c_{22} \in \{\pm 1\}$ for the second projector, respectively. For the third projector, we denote the sum of the sign choices with $m \in \{-3, -1, 1, 3\}$ (and for the sake of

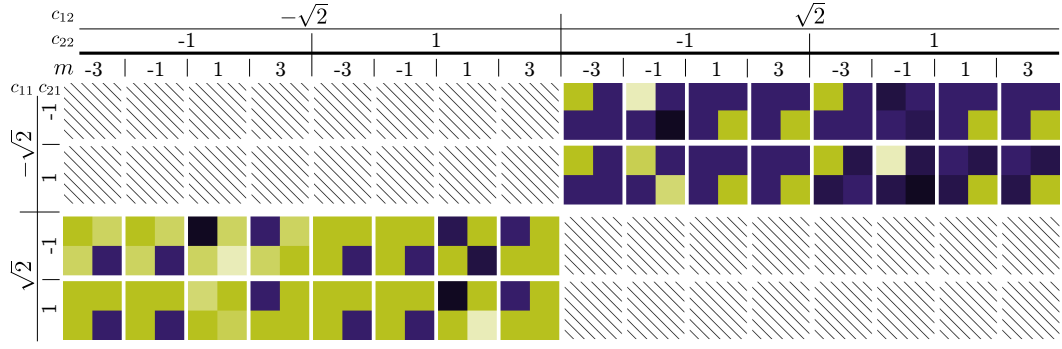


Figure 1.2: \mathbf{C}'_s for various sign choices of the eigenvalues c_{ij} , $i, j = 1, 2$ corresponding to the eigenvectors $c_{1,2}$. Only all positive signs and $m := \sum_j m_{ij} = -1$ yields a nonnegative block (third from right in top row). Hatching signifies complex numbers, the colour scale is the same as in fig. 1.3.

clarity we drop the rescaling from eq. (1.5)). Overall, we obtain the nonzero entries

- $$c_{11} \frac{1+i}{\sqrt{2}} \begin{pmatrix} 1 & -i \\ i & 1 \end{pmatrix} - c_{12} \frac{1-i}{\sqrt{2}} \begin{pmatrix} 1 & i \\ -i & 1 \end{pmatrix} + c_{21} \frac{1}{2} \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} - c_{22} \frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} + \begin{pmatrix} 0 & m \\ -m & 0 \end{pmatrix}.$$

To avoid complex entries, $c_{11} \stackrel{!}{=} -c_{12}$. The two off-diagonal entries encode the 1-IN-3SAT inequalities from eq. (1.3); by demanding nonnegativity, we obtain the inequalities

$$\frac{3}{2} + p_{i1} + p_{i2} + p_{i3} \geq 0 \quad \text{and} \quad -\frac{1}{2} - p_{i1} - p_{i2} - p_{i3} \geq 0.$$

It is straightforward to verify that no other sign choice for the eigenvalues of the first two terms in eq. (1.4) yields nonnegative blocks, which we demonstrate in fig. 1.2. From this, the first claim of the lemma follows.

We can always orthonormalise the vectors $c_{1,2}$ and v_k using the freedom left in \vec{o} , hence we can achieve that $\mathbf{C}_s^2 = \mathbf{C}_t^2 \forall s, t$, from which the last two claims follows. \square

1.1.7 Orthonormalisation and Handling the Unwanted Inequalities

As in [CEW12b], we have unwanted inequalities—the off-diagonal blocks in the first $4n_c$ entries and the blocks involving the orthonormalisation region \vec{o} . We first deal with the off-diagonal blocks in favour of enlarging the orthonormalisation region, creating more—potentially negative—entries in there, and then fix the latter.

For this, we first pad \mathbf{C}'_s

$$\mathbf{C}_s := \begin{pmatrix} \mathbf{C}'_s & 0 \\ 0 & 0 \end{pmatrix} \in \mathbb{C}^{d \times d},$$

where we have used an obvious block notation to pad \mathbf{C}'_s with zeroes. •

1.1.7.1 Off-Diagonal Blocks

We begin with the following lemma.

Lemma 1.44. *Let the family $(\mathbf{C}_s)_{s \in S}$ be defined as in the proof of lemma 1.43, and $(n_v, n_c, \{m_i\}, \{m_{ij}\})$ the corresponding 1-IN-3SAT instance. Then there exists a matrix $\mathbf{E} \in \mathbb{C}^{d \times d}$ such that the top left $4n_c \times 4n_c$ block of $\mathbf{C}_s + \mathbf{E}$ has at least one negative entry $\forall s$ iff the instance is not satisfiable. Furthermore, $\text{im } \mathbf{C}_s \perp \text{im } \mathbf{E} \forall s$, and $\mathbf{C}_s + \mathbf{E}'$ has negative entries $\forall s, \forall \mathbf{E}' \in \sqrt{\mathbf{E}^2} \setminus \{\mathbf{E}\}$.* •

Proof. Define

$$\mathbf{E}_1 := E_1 E_1^T \otimes \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \end{pmatrix} \quad \text{where } E_1 := (1, \dots, 1, \vec{\sigma})^T.$$

Then \mathbf{E}_1 has rank 1.

From this mask, we now erase the first n_c on-diagonal 4×4 -blocks, while leaving all other entries in the upper left $4n_c \times 4n_c$ block positive. Define $b_i := (e_i, \vec{\sigma}) \in \mathbb{C}^d$ for $i = 1, \dots, n_c$ where e_i denotes the i^{th} unit vector, and let

$$\mathbf{E} := \frac{7}{2} \mathbf{E}_1 - \frac{7}{2} \sum_{i=1}^{n_c} t_i b_i b_i^T \otimes \begin{pmatrix} 1 & 1 & -1 & 0 \\ 1 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

The variables t_i are chosen close to 1 but distinct, e.g.

$$t_i := \left(1 - \frac{1}{M} - \frac{i}{M n_c} \right), \tag{1.6}$$

where $M \in \mathbb{N}$ large but fixed. Then \mathbf{E} has rank $n_c + 1$, and adding \mathbf{E} to \mathbf{C}_s trivialises all unwanted inequalities in the upper left $4n_c \times 4n_c$ block. By picking M large enough, the on-diagonal inequalities are left intact.

One can check that all other possible sign choices for the roots of \mathbf{E} create negative entries in parts of the upper left block where \mathbf{C}_s is zero $\forall s$. Furthermore, \mathbf{C}_s and \mathbf{E} have distinct nonzero eigenvalues by construction—the orthogonality condition is again straightforward, hence the last two claims follow. \square

1.1.7.2 Orthonormalisation Region

Lemma 1.45. *Let $4n < d$ and $\delta \gg 1$. There exists a nonnegative rank 2 matrix $\mathbf{D} \in \mathbb{C}^{d \times d}$ such that the top left $4n \times 4n$ block of \mathbf{D} has entries $\mathbf{D}_{ij} \in \mathcal{O}(\delta^{-2})$ if $j \nmid 4$ and the rest of the matrix entries are $\mathcal{O}(\delta^{-1})$. If $\mathbf{D}' \in \sqrt{\mathbf{D}^2}$, either the same holds true for \mathbf{D}' , or $\mathbf{D}'_{ij} < 0 \forall j < 4n + 1, j \nmid 4$.*

Proof. Define

$$E_2 := \left(\underbrace{\frac{1}{\delta}, \dots, \frac{1}{\delta}}_{n \text{ times}}, 1, \dots, 1 \right) \in \mathbb{C}^d$$

and let $\mathbf{E}_2 := E_2 E_2^T \otimes \mathbf{1}_4$, where $\mathbf{1}_4 := (1)_{1 \leq i, j \leq 4}$. Let further

$$\Delta := \left(\underbrace{\frac{1}{\delta}, \dots, \frac{1}{\delta}}_{n \text{ times}}, -\frac{1}{\delta}, \dots, -\frac{1}{\delta}, a \right) \in \mathbb{C}^d,$$

where $0 < a < 1$ is used to orthonormalise Δ and E_2 , which is the case if

$$a = -\frac{n}{\delta^2} + \frac{d - n - 1}{\delta}.$$

By explicitly writing out the rank 2 matrix

$$\mathbf{D} := \mathbf{E}_2 \pm \Delta \Delta^T \otimes \begin{pmatrix} 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 0 \end{pmatrix},$$

it is straightforward to check that \mathbf{D} fulfils all the claims of the lemma—see fig. 1.3 for an example. \square

1.1.8 Lifting Singularities

The reader will have noted by now that even though we have orthonormalised all our eigenspaces, ensuring that the nonzero eigenvalues are all distinct, we have at the same time introduced a high-dimensional kernel in \mathbf{C}_s , \mathbf{E} and \mathbf{D} . The following lemma shows that this does not pose an issue.

Lemma 1.46. *Let $(\mathbf{A}_s)_{s \in S}$ be the family of primary rational roots of some degenerate $\mathbf{B} \in \mathbb{Q}^{d \times d}$. Then there exists a non-degenerate matrix \mathbf{B}' , such that for the family $(\mathbf{A}'_s)_{s \in S}$ of roots of \mathbf{B}' , we have \mathbf{A}_s positive*

iff \mathbf{A}'_s positive. Furthermore, the entries of \mathbf{A}'_s are rational with bit complexity $\tau(\mathbf{A}'_s) = \mathcal{O}(\text{poly}(\tau(\mathbf{A}_s)))$.

Proof. Take a matrix $\mathbf{A} \in (\mathbf{A}_s)_{s \in S}$. We need to distort the zero eigenvalues $\{\lambda_i^{(0)}\}$ slightly away from 0. Using notation from definition 1.50, a conservative estimate for the required smallness without affecting positivity would be

$$\lambda_i^{(0)} \mapsto \lambda_i'^{(0)} \quad : \quad 0 < \lambda_i'^{(0)} \leq \frac{1}{C \cdot d^3 \cdot \max_{i,j} \{|\mathbf{Z}_{ij}|, |\mathbf{Z}_{ij}^{-1}|\}},$$

where we used the Jordan canonical form $\mathbf{A} = \mathbf{Z}\mathbf{\Lambda}\mathbf{Z}^{-1}$ for some invertible \mathbf{Z} and $\mathbf{\Lambda} = \text{diag}(\mathbf{J}_0, \mathbf{J}_1)$, such that \mathbf{J}_0 collects all Jordan blocks corresponding to the eigenvalue 0, and \mathbf{J}_1 collects the remaining ones. \square

This will lift all remaining degeneracies and singularities, without affecting our line of argument above. Observe that all inequalities in our construction were bounded away from 0 with enough head space independent of the problem size, so positivity in the lemma is sufficient.

We thus constructed an embedding of 1-IN-3SAT into non-derogatory and non-degenerate matrices, as desired. It is crucial to note that we do not lose anything by restricting the proof to the study of these matrices, as the following lemma shows.

Lemma 1.47. *There exists a Karp reduction of the DIVISIBILITY problems when defined for all matrices to the case of non-degenerate and non-derogatory matrices.*

Proof. As shown in lemma 1.35, containment in NP for this problem is easy to see, also in the degenerate or derogatory case. Since 1-IN-3SAT is NP-complete, there has to exist a poly-time reduction of the DIVISIBILITY problems—when defined for *all* matrices—to 1-IN-3SAT. Now embed this 1-IN-3SAT-instance with our construction. This yields a poly-time reduction to the non-degenerate non-derogatory case. \square

1.1.9 Complete Embedding

We now finally come to the proof of theorem 1.42.

Theorem 1.42. Construct the family $(\mathbf{C}_s + \mathbf{E})_{s \in S}$ using lemma 1.43 and lemma 1.44, ensuring that all orthonormalising is done, which preliminarily fixes the dimension d . By lemma 1.45, we now construct a mask $\mathbf{D}(\delta)$ of dimension $d + d'$, where $d' > 0$ is picked such that we can also orthonormalise all previous vectors with respect to E_2 and δ . •

By lemmas 1.43, 1.44, 1.45 and 1.46, the perturbed family $(\mathbf{M}'_s)_{s \in S} := (\mathbf{C}_s + \mathbf{E} + N\mathbf{D}(\delta))'_s$ —where N and $\delta \in \mathbb{Q}$ are chosen big enough so that all unwanted inequalities are trivially satisfied—fulfils the claims of the theorem and the proof follows. \square

We finalise the construction as follows. In theorem 1.42, we have embedded a given 1-IN-3SAT instance into a family of matrices $(\mathbf{M}_s)_{s \in S}$, such that the instance is satisfiable if and only if at least one of those matrices is nonnegative.

By rescaling the entire matrix such that $\max_{i,j}(\mathbf{M}_s)_{ij} = 1/2$, we could show that this instance of 1-IN-3SAT is satisfiable if and only if the normalised matrix family $(\mathbf{Q}_s)_{s \in S}$, which we construct explicitly, contains a stochastic matrix.

As shown in section 1.1.3, this can clearly be answered by STOCHASTIC DIVISIBILITY, as the family $(\mathbf{Q}_s)_{s \in S}$ comprises all the roots of a unique matrix \mathbf{P} . If this matrix is *not* stochastic, our instance of 1-IN-3SAT is trivially not satisfiable. If the matrix *is* stochastic, we ask STOCHASTIC DIVISIBILITY for an answer—a positive outcome signifies satisfiability, a negative one non-satisfiability.

1.1.10 Bit Complexity of Embedding

To show that our results holds for only polynomially growing bit complexity, observe the following proposition.

- Proposition 1.48. *The bit complexity $\tau(\mathbf{M}_s)$ of the constructed embedding of a 1-IN-3SAT instance $(n_v, n_c, \{m_i\}, \{m_{ij}\})$ equals $\mathcal{O}(\text{poly}(n_v, n_c))$.*

Proof. We can ignore any construction that multiplies by a constant prefactor, for example lemma 1.39 and lemma 1.40. The renormalisation for lemma 1.39 to $\max_{i,j} \mathbf{M}_{s,ij} = 1/2$ does not affect τ either.

The rescaling in eq. (1.5) and eq. (1.6) yields a complexity of $\mathcal{O}(\log n_v)$, and the same thus holds true for lemma 1.43 and lemma 1.44.

The only other place of concern is the orthonormalisation region. Let us write a_i for all vectors that need orthonormalisation. In the n^{th} step, we need to make up for $\mathcal{O}(n)$ entries with our orthonormalisation, using the same amount of precision to solve the linear equations $(a_i^T a_n = 0)_{1 \leq i < n}$. This has to be done with a variant of the standard Gauss algorithm, e.g. the Bareiss algorithm—see for example [Bar68]—which has nonexponential bit complexity.

Together with the lifting of our singularities, which has polynomial precision, we obtain $\tau(\mathbf{M}_s) = \mathcal{O}(\text{poly}(n_v, n_c))$. Completing the embedding in section 1.1.9 changes the bit complexity by another polynomial factor, at most, and hence the claim follows. \square

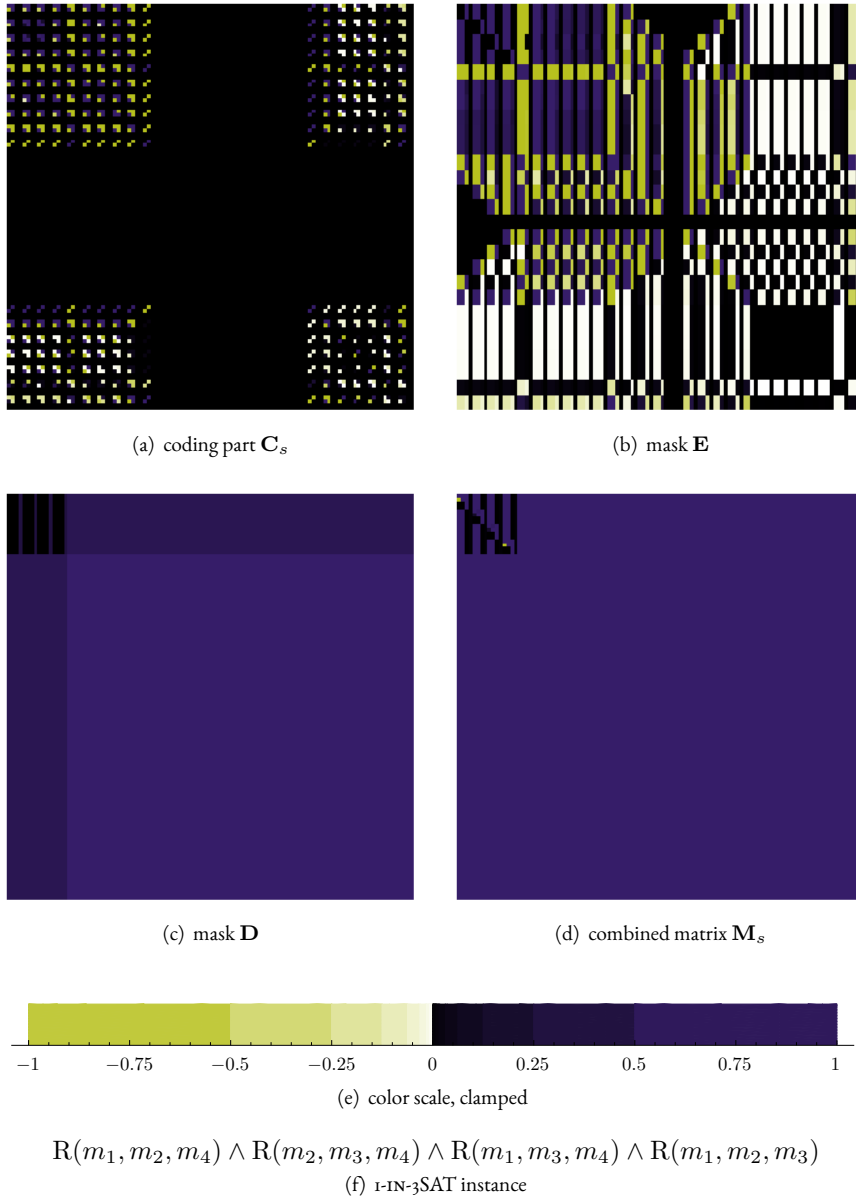


Figure 1.3: One branch of an unsatisfiable instance of I -IN-3SAT encoded into a matrix of total rank 19. The negative entries—two bright dots—in the upper left block in the combined matrix (d) indicate that this branch does not satisfy the given instance. By looking at all other blocks, one sees that none translates to a nonnegative matrix. Observe that in this naïve implementation the orthonormalisation region is sub-optimally large.

1.2 Distribution Divisibility

Underlying stochastic and quantum channel divisibility, and—to some extent—a more fundamental topic, is the question of divisibility and decomposability of probability distributions and random variables. An illustrative example is the distribution of the sum of two rolls of a standard six-sided die, in contrast to the single roll of a twelve-sided die. Whereas in the first case the resulting random variable is obviously the sum of two uniformly distributed random variables on the numbers $\{1, \dots, 6\}$, there is no way to achieve the outcome of the twelve-sided die as any sum of nontrivial “smaller” dice—in fact, there is no way of dividing *any* uniformly distributed discrete random variable into the sum of non-constant random variables. In contrast, a uniform continuous distribution can always be decomposed² into two *different* distributions.

To be more precise, a random variable X is said to be divisible if it can be written as $X = Y + Z$, where Y and Z are non-constant independent random variables that are identically distributed (iid). Analogously, infinite divisibility refers to the case where X can be written as an infinite sum of such iid random variables.

If we relax the condition $Y \stackrel{d}{=} Z$ —i.e. we allow Y and Z to have different distributions—we obtain the much weaker notion of decomposability. This includes using other sources of randomness, not necessarily uniformly distributed.

Both divisibility and decomposability have been studied extensively in various branches of probability theory and statistics. Early examples include Cramer’s theorem [Cra36], proven in 1936, a result stating that a Gaussian random variable can only be decomposed into random variables which are also normally distributed. A related result on χ^2 distributions by Cochran [CW34], dating back to 1934, has important implications for the analysis of covariance.

An early overview over divisibility of distributions is given in [SK79]. Important applications of n -divisibility—the divisibility into n iid terms—is in modelling, for example of bug populations in entomology³ [Kat77], or in financial aspects of various insurance models⁴ [Tho77a]; [Tho77b]. Both examples study the overall distribution and ask if it is compatible with an underlying subdivision into smaller random events. The authors also give various conditions on distributions to be infinitely divisible, and list numerous infinitely divisible distributions.

Important examples for infinite divisibility include the Gaussian, Laplace, Gamma and Cauchy distributions, and in general all normal distributions. It is clear that those distributions are also finitely divisible, and decomposable. Examples of indecomposable distributions are Bernoulli and discrete uniform distributions

²All continuous uniform distributions decompose into the sum of a discrete Bernoulli distribution and another continuous uniform distribution. This decomposition is never unique.

• ³If X is the number of bugs collected over a certain time period and in some area, and the assumption is that $X = \sum_{i=1}^5 0Y_i$ for iid Y_i representing a small part of the area, respectively, the question is whether one can deduce from $X \sim \text{LOG}$ that the $Y_i \sim \text{LOG}$ as well, i.e. whether the bug distribution of a smaller lot could be deduced from the overall distribution.

• ⁴If X is the overall money paid out throughout a year, the question is whether an assumption on X ’s distribution—e.g. a Pareto distribution—would be compatible with an assumption that $X = \sum_{i=1}^5 2Y_i$, i.e. that the money paid over the year can be sub-divided into weekly iid payments.

on $\{1, \dots, p\}$ for p prime.

However, there does not yet exist a straightforward way of checking whether a given discrete distribution is divisible or decomposable. We will show in this work that the question of decomposability is NP-hard, whereas divisibility is in P. In the latter case, we outline a computationally efficient algorithm for solving the divisibility question. We extend our results to weak-membership formulations (where the solution is only required to within an error ϵ in total variation distance), and argue that the continuous case is computationally trivial as the indecomposable distributions form a dense subset.

We start out in section 1.2.1 by introducing general notation and a rigorous formulation of divisibility and decomposability as computational problems. The foundation of all our distribution results is by showing equivalence to polynomial factorisation, proven in section 1.2.2. This will allow us to prove our main divisibility and decomposability results in section 1.2.3 and 1.2.4, respectively.

1.2.1 Preliminaries

1.2.1.1 Discrete Distributions

In our discussion of distribution divisibility and decomposability, we will use the standard notation and language as described in the following definition.

Definition 1.49. *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a discrete probability space, i.e. Ω is at most countably infinite and the probability mass function $\mathbb{P} : \Omega \rightarrow [0, 1]$ —or pmf, for short—fulfils $\sum_{x \in \Omega} \mathbb{P}(x) = 1$. We take the σ -algebra \mathcal{F} to be maximal, i.e. $\mathcal{F} = 2^\Omega$, and without loss of generality assume that the state space $\Omega = \mathbb{N}$. Denote the distribution described by \mathbb{P} with \mathcal{D} . A random variable $X : \Omega \rightarrow B$ is a measurable function from the sample space to some set B , where usually $B = \mathbb{R}$.*

For the sake of completeness, we repeat the following well-known definition of characteristic functions.

Definition 1.50. *Let \mathcal{D} be a discrete probability distribution with pmf \mathbb{P} , and $X \sim \mathcal{D}$. Then*

$$\phi_X(\omega) := \mathbb{E}(e^{i\omega X}) = \int_{\Omega} e^{i\omega x} dF_X(x) = \sum_{x \in \Omega} \mathbb{P}(x) e^{i\omega x}$$

defines the characteristic function of \mathcal{D} .

It is well-known that two random variables with the same characteristic function have the same cumulative density function.

Definition 1.51. *Let the notation be as in definition 1.49. Then the distribution \mathcal{D} is called finite if $\mathbb{P}(k) = 0 \forall k \geq N$ for some $N \in \mathbb{N}$.*

Remark 1.52. Let \mathfrak{D} be a discrete probability distribution with pmf \mathfrak{p} . We will—without loss of generality—assume that $\mathfrak{p}(0) \neq 0$ and $\mathfrak{p}(k) = 0 \forall k < 0$ for the pmf \mathfrak{p} of a finite distribution. It is a straightforward shift of the origin that achieves this.

1.2.1.2 Continuous Distributions

Definition 1.53. Let $(\mathcal{X}, \mathcal{A})$ be a measurable space, where \mathcal{A} is the σ -algebra of \mathcal{X} . The probability distribution of a random variable X on $(\mathcal{X}, \mathcal{A})$ is the Radon-Nikodym derivative f , which is a measurable function with $\mathbb{P}(X \in A) = \int_A f d\mu$, where μ is a reference measure on $(\mathcal{X}, \mathcal{A})$.

Observe that this definition is more general than definition 1.49, where the reference measure is simply the counting measure over the discrete sample space Ω . Since we are only interested in real-valued univariate continuous random variables, observe the following important

Remark 1.54. We restrict ourselves to the case of $\mathcal{X} = \mathbb{R}$ with \mathcal{A} the Borel sets as measurable subsets and the Lebesgue measure μ . In particular, we only consider distributions with a probability density function f —or pdf, for short—i.e. we require the cumulative distribution function $\mathbb{P}(x) := \mathbb{P}(X \leq x) \equiv \int_{y \leq x} f(y) dy$ to be absolutely continuous.

Corollary 1.55. The cumulative distribution function \mathbb{P} of a continuous random variable X is almost everywhere differentiable, and any piecewise continuous function f with $\int_{\mathbb{R}} f(x) dx = 1$ defines a valid continuous distribution.

1.2.1.3 Divisibility and Decomposability of Distributions

To make the terms mentioned in the introduction rigorous, note the two following definitions.

Definition 1.56. Let X be a random variable. It is said to be n -decomposable if $X = Z_1 + \dots + Z_n$ for some $n \in \mathbb{N}$, where Z_1, \dots, Z_n are independent non-constant random variables. X is said to be indecomposable if it is not decomposable.

Definition 1.57. Let X be a random variable. It is said to be n -divisible if it is n -decomposable as $X = \sum_{i=1}^n Z_i$ and $Z_i \stackrel{d}{=} Z_j \forall i, j$. X is said to be infinitely divisible if $X = \sum_{i=1}^{\infty} Z_i$, with $Z_i \sim \mathfrak{D}$ for some nontrivial distribution \mathfrak{D} .

If we are not interested in the exact number of terms, we also simply speak of *decomposable* and *divisible*. We will show in section 1.2.4.7 that—in contrast to divisibility—the question of decomposability into more than two terms is not well-motivated.

Observe the following extension of remark 1.52.

Lemma 1.58. Let \mathfrak{D} be a discrete probability distribution with pmf p . If p obeys remark 1.52, then we can assume that its factors do as well. In the continuous case, we can without loss of generality assume the same.

Proof. Obvious from positivity of convolutions in case of divisibility. For decomposability, we can reach this by shifting the terms symmetrically. \square

1.2.1.4 Markov Chains

To establish notation, we briefly state some well-known properties of Markov chains

Remark 1.59. Take discrete iid random variables $Y_1, \dots, Y_n \sim \mathfrak{D}$ and write $\mathbb{P}(Y_i = k) = p_k := p(k)$ for all $k \in \mathbb{N}$, independent of $i = 1, \dots, n$. Define further

$$X_i := \begin{cases} Y_1 + \dots + Y_i & i > 0 \\ 0 & \text{otherwise.} \end{cases}$$

Then $\{X_n, n \geq 0\}$ defines a discrete-time Markov chain, since

$$\begin{aligned} \mathbb{P}(X_{n+1} = k_{n+1} | X_0 = k_0 \wedge \dots \wedge X_n = k_n) &= \mathbb{P}(Y_{n+1} = k_{n+1} - k_n) \\ &\equiv p_{k_{n+1} - k_n}. \end{aligned}$$

This last property is also called stationary independent increments. •

Remark 1.60. Let the notation be as in remark 1.59. The transition probabilities of the Markov chain are then given by

$$P_{ij} := \begin{cases} p_{j-i} & j \geq i \\ 0 & \text{otherwise.} \end{cases}$$

In matrix form, we write the transition matrix

$$\mathbf{P} := \begin{pmatrix} p_0 & p_1 & p_2 & \cdots \\ & p_0 & p_1 & \cdots \\ & & p_0 & \cdots \\ & & & \ddots \end{pmatrix}.$$

Working with transition matrices is straightforward—if the initial distribution is given by $\pi := (1, 0, \dots)$, then obviously $(\pi \mathbf{P})_i = p_i$. Iterating \mathbf{P} then yields the distributions of X_2, X_3, \dots , respectively—e.g. $(\pi \mathbf{P}^2)_i = \mathbb{P}(X_2 = i) \equiv \mathbb{P}(Y_1 + Y_2 = i)$.

We know that X_2 is divisible—namely into $X_2 = Y_1 + Y_2$, by construction—but what if we ask this question the other way round? We will show in the next section that there exists a relatively straightforward way to calculate if an (infinite) matrix in the shape of \mathbf{P} has a stochastic root—i.e. if \mathcal{D} is divisible. Observe that this is not in contradiction with theorem 1.1, as the theorem does not apply to infinite operators.

In contrast, the more general question of whether we can write a finite discrete random variable as a sum of nontrivial, potentially distinct random variables will be shown to be NP-hard.

1.2.2 Equivalence to Polynomial Factorisation

Starting from our digression in section 1.2.1.4 and using the same notation, we begin with the following definition.

Definition 1.61. Denote with \mathbf{S} the shift matrix $S_{ij} := \delta_{i+1,j}$. Then we can write

$$\mathbf{P} = p_0 \mathbf{1} + p_1 \mathbf{S} + p_2 \mathbf{S}^2 + \dots = \sum_{i=0}^{\infty} p_i \mathbf{S}^i \in \mathbb{R}_{[0,1]}[\mathbf{S}].$$

Since \mathbf{S} just acts as a symbol, we write

$$f_{\mathcal{D}}(x) := \sum_{i=0}^N p_i x^i \in \mathcal{R} \quad \text{where} \quad \mathcal{R} := \mathbb{R}_{\geq 0}[x]/\sim,$$

and $f \sim g :\Leftrightarrow f = cg, c > 0$. We call $f_{\mathcal{D}}$ the characteristic polynomial of \mathcal{D} —not to be confused with the characteristic polynomial of a matrix. The equivalence space \mathcal{R} defines the set of all characteristic polynomials, and can be written as

- $$\mathcal{R} = \bigcup_{i=n}^{\infty} \mathcal{R}_i \quad \text{where} \quad \mathcal{R}_n := \{f \in \mathcal{R} : \deg f = n\}.$$

We mod out the overall scaling in order to keep the normalisation condition $\sum_k p(k) = 1$ implicit—if we write $f_{\mathcal{D}}$, we will always assume $f_{\mathcal{D}}(1) = 1$. An alternative way to define these characteristic polynomials is via characteristic functions, as given in definition 1.50.

Definition 1.62. $f_{\mathcal{D}}(e^{i\omega}) = \phi_X(\omega)$.

The reason for this definition is that it allows us to reduce operations on the transition matrix \mathbf{P} or products of characteristic functions ϕ_X to algebraic operations on $f_{\mathcal{D}}$. This enables us to translate the divisibility problem into a polynomial factorisation problem and use algebraic methods to answer it. Because we will make use of it later, we also observe the following.

Definition 1.63. We define norms on the space of characteristic polynomials of degree $N \in \mathcal{R}_N$ —via $\|f_{\mathfrak{D}}\|_{N,p} := \|(p_i)_{0 \leq i \leq N}\|_{\ell^p}$. If N is not explicitly specified, we usually assume $N = \deg f_{\mathfrak{D}}$. •

First note the following proposition.

Proposition 1.64. There is a 1-to-1 correspondence between finite distributions \mathfrak{D} and characteristic polynomials $f_{\mathfrak{D}}$, as defined in definition 1.61.

Proof. Clear by definition 1.62 and the uniqueness of characteristic functions. □

While this might seem obvious, it is worth clarifying, since this correspondence will allow us to directly translate results on polynomials to distributions.

The following lemma reduces the question of divisibility and decomposability—see definition 1.56 and 1.57—to polynomial factorisation.

Lemma 1.65. A finite discrete distribution \mathfrak{D} is n -divisible iff there exists a polynomial $g \in \mathcal{R}$ such that $g^n = f_{\mathfrak{D}}$. \mathfrak{D} is n -decomposable iff there exist polynomials $g_1, \dots, g_n \in \mathcal{R}$ such that $\prod_{i=1}^n g_i = f_{\mathfrak{D}}$.

Proof. Assume that \mathfrak{D} is n -divisible, i.e. that there exists a distribution \mathfrak{D}' and random variables $Z_1, \dots, Z_n \sim \mathfrak{D}'$ such that $X = \sum_{i=1}^n Z_i$. Denote with \mathbf{Q} the transition matrix of \mathfrak{D}' , as defined in remark 1.60, and write q for its probability mass function. Then

$$\mathbb{P}(X = j) = \mathbb{P}\left(\sum_{i=1}^n Z_i = j\right) = (\mathbf{Q}^n \pi)_j,$$

as before. Write $g_{\mathfrak{D}'}$ for the characteristic polynomial of \mathfrak{D}' . By definition 1.61, $g^n(\mathbf{S}) \equiv f_{\mathfrak{D}}(\mathbf{S})$, and hence $g_{\mathfrak{D}'}^n = f_{\mathfrak{D}}$. Observe that

$$1 = \sum_i p(i) = f_{\mathfrak{D}}(1) \equiv g_{\mathfrak{D}'}^n(1) = \left(\sum_i q(i)\right)^n,$$

and hence $\sum_i q(i) = 1$ is normalised automatically.

The other direction is similar, as well as the case of decomposability, and the claim follows. □

1.2.3 Divisibility

1.2.3.1 Computational Problems

We state an exact variant of the computational formulation of the question according to definition 1.57—i.e. one with an allowed margin of error—as well as a weak membership formulation.

Definition 1.66 (DISTRIBUTION DIVISIBILITY $_n$).

Instance. Finite discrete random variable $X \sim \mathfrak{D}$.

Question. Does there exist a finite discrete distribution $\mathfrak{D}' : X = \sum_{i=1}^n Z_i$ for random variables $Z_i \sim \mathfrak{D}'$?

Observe that this includes the case $n = 2$, which we defined in definition 1.57.

Definition 1.67 (WEAK DISTRIBUTION DIVISIBILITY $_{n,\epsilon}$).

Instance. Finite discrete random variable $X \sim \mathfrak{D}$ with pmf $p_X(k)$.

Question. If there exists a finite discrete random variable Y with pmf $p_Y(k)$, such that $\|p_X - p_Y\|_\infty < \epsilon$ and such that

1. Y is n -divisible—return YES
2. Y is not n -divisible—return NO.

1.2.3.2 Exact Divisibility

Theorem 1.68. DISTRIBUTION DIVISIBILITY $_n \in P$.

Proof. By lemma 1.65 it is enough to show that for a characteristic polynomial $f \in \mathcal{R}_N$, we can find a $g \in \mathcal{R} : g^n = f$ in polynomial time. In order to achieve this, write $(f)^{1/n}$ as a Taylor expansion with rest, i.e.

$$\sqrt[n]{f(x)} = p(x) + R(x) \quad \text{where } p \in \mathcal{R}_{N/n}, R \in \mathcal{R}_N.$$

If $R \equiv 0$, then $g = p$ n -divides f , and then the distribution described by f is n -divisible. Since the series expansion is constructive and can be done efficiently—see [Mül87]—the claim follows.

If the distribution coefficients are rational numbers, another method is to completely factorise the polynomial—e.g. using the LLL algorithm, which is known to be easy in this setting—sort and recombine the linear factors, which is also in $O(\text{poly}(\text{ord } f))$, see for example [HHN11]. Then check if all the polynomial root coefficients are positive. □

We collect some further facts before we move on.

Remark 1.69. Let p be the probability mass function for a finite discrete distribution \mathfrak{D} , and write $\text{supp } p = \{k : p(k) \neq 0\}$. If $\max \text{supp } p - \min \text{supp } p =: w$, then \mathfrak{D} is obviously not n -divisible for $n > w/2$, and furthermore not for any n that do not divide w , $n < w/2$. Indeed, \mathfrak{D} is not n -divisible if the latter condition holds for either $\max \text{supp } p$ or $\min \text{supp } p$.

Remark 1.70. Let $X \sim \mathfrak{D}$ be an n -divisible random variable, i.e. $\exists Z_1, \dots, Z_n \sim \mathfrak{D}' : \sum_{i=1}^n Z_i = X$. Then \mathfrak{D}' is unique.

Proof. This is clear, because $\mathbb{R}[x]$ is a unique factorisation domain. □

1.2.3.3 Divisibility with Variation

As an intermediate step, we need to extend theorem 1.68 to allow for a margin of error ϵ , as captured by the following definition.

Definition 1.71 (DISTRIBUTION DIVISIBILITY $_{n,\epsilon}$).

Instance. Finite discrete random variable $X \sim \mathcal{D}$ with pmf $p_X(k)$.

Question. Do there exist finite random variables $Z_1, \dots, Z_n \sim \mathcal{D}'$ with pmfs $p_Z(k)$, such that $\| \underbrace{p_Z * \dots * p_Z}_{n \text{ times}} - p_X \|_\infty < \epsilon$?

Lemma 1.72. DISTRIBUTION DIVISIBILITY $_{n,\epsilon}$ is in \mathbf{P} .

Proof. Let $f(x) = \sum_{i=0}^N p_i x^i$ be the characteristic polynomial of a finite discrete distribution, and $\epsilon > 0$. By padding the distribution with 0s, we can assume without loss of generality that $N = \deg f$ is a multiple of n . A polynomial root—if it exists—has the form $g(x) = \sum_{i=0}^N a_i x^i$, where $a_i \geq 0 \forall i$. Then

$$\begin{aligned} g(x)^n &= (\dots + a_3 x^3 + a_2 x^2 + a_1 x + a_0)^n \\ &= \dots + ((n-1)a_1^2 + n a_0^{n-2} a_2) x^2 + n a_0^{n-1} a_1 x + a_0^n. \end{aligned}$$

Comparing coefficients in the divisibility condition $f(x) = g(x)^n$, the latter translates to the set of inequalities

$$\begin{aligned} a_0^n &\in (p_0 - \epsilon, p_0 + \epsilon) \\ n a_0^{n-1} a_1 &\in (p_1 - \epsilon, p_1 + \epsilon) \\ (n-1)a_1^2 + n a_0^{n-2} a_2 &\in (p_2 - \epsilon, p_2 + \epsilon) \\ &\vdots \end{aligned}$$

Each term but the first one is of the form $h_i(a_1, \dots, a_{i-1}) + n a_0^{n-i} a_i \in (p_i - \epsilon, p_i + \epsilon)$, where $h_i \geq 0 \forall i$ is monotonic. This can be rewritten as $a_i \in U_{\epsilon/n a_0^{n-i}}((p_i - h_i(a_1, \dots, a_{i-1}))/n a_0^{n-i})$. It is now easy to solve the system iteratively, keeping track of the allowed intervals I_i for the a_i . •

If $I_i = \emptyset$ for some i , we return NO, otherwise YES. We have thus developed an efficient algorithm to answer DISTRIBUTION WEAK DIVISIBILITY $_{n,\epsilon}$, and the claim of lemma 1.72 follows. □

Remark 1.73. Given a random variable X , the algorithm constructed in the proof of lemma 1.72 allows us to calculate the closest n -divisible distribution to X in polynomial time.

Proof. Straightforward, e.g. by using binary search over ϵ . □

1.2.3.4 Weak Divisibility

For the weak membership problem, we reduce WEAK DISTRIBUTION DIVISIBILITY $_{n,\epsilon}$ to DISTRIBUTION DIVISIBILITY $_{n,\epsilon}$.

Theorem 1.74. *WEAK DISTRIBUTION DIVISIBILITY $_{n,\epsilon} \in \mathbf{P}$.*

Proof. Let \mathfrak{D} be a finite discrete distribution. If DISTRIBUTION DIVISIBILITY $_{n,\epsilon}$ answers YES, we know that there exists an n -divisible distribution ϵ -close to \mathfrak{D} . In case of NO, \mathfrak{D} itself is not n -divisible, hence we know that there exists a non- n -divisible distribution close to \mathfrak{D} . \square

1.2.3.5 Continuous Distributions

Let us briefly discuss the case of continuous distributions—continuous meaning a non-discrete state space \mathcal{X} , as specified in section 1.2.1.2. Although divisibility of continuous distributions is well-defined and widely studied, formatting the continuous case as a computational problem is delicate, as the continuous distribution must be specified by a finite amount of data for the question to be computationally meaningful. The most natural formulation is the continuous analogue of definition 1.57 as a weak-membership problem. However, we can show that this problem is computationally trivial.

First observe the following intermediate result.

Lemma 1.75. *Take $f \in C_{c,b}^+$ with $\text{supp } f \subset A \cup B$, where $A := [0, M]$, $B := [2M, 3M]$, $M \in \mathbb{R}_{>0}$. We claim that if f is divisible, then both $f|_A$ and $f|_B$ are divisible.*

Proof. Due to symmetry, it is enough to show divisibility for $f|_A$. Assuming f is divisible, we can write $f = r * r$, i.e. $f(x) = \int_{\mathbb{R}} r(x-y)r(y)dy$. It is straightforward to show that $r(x) = 0 \forall x < 0$. Define

$$\bar{r}(x) = \begin{cases} r(x) & x \in A/2 \\ 0 & \text{otherwise,} \end{cases} \quad (1.7)$$

where $A/2 := \{a/2 : a \in A\}$. Then

$$\begin{aligned} (\bar{r} * \bar{r})(x) &= \int_{\mathbb{R}} \bar{r}(x-y)\bar{r}(y)dy \\ &= \int_{\mathbb{R}} dy \begin{cases} r(x-y) & x-y \in A/2 \\ 0 & \text{otherwise} \end{cases} \cdot \begin{cases} r(y) & y \in A/2 \\ 0 & \text{otherwise.} \end{cases} \end{aligned}$$

We see that $(\bar{r} * \bar{r})(x) = 0$ for $x \notin A$. For $x \in A$, the support of the integrand is contained in $\{y : y \in x - A/2 \wedge y \in A/2\} = x - A/2 \cap A/2 := S_x$, and hence we can write $(\bar{r} * \bar{r})(x) = \int_{S_x} r(x-y)r(y)dy$.

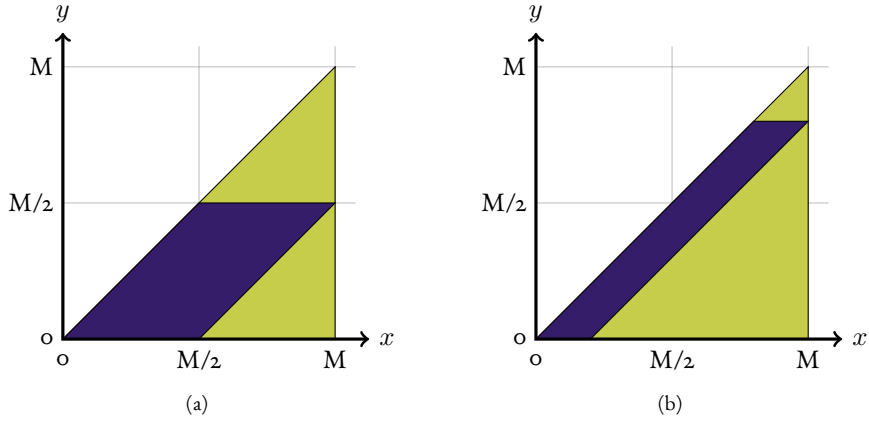


Figure 1.4: (a) Integration domains in lemma 1.75 for $\bar{r} * \bar{r}$ (purple) and $f|_A$ (light green), respectively. (b) Example for integration domains in proposition 1.96 for $\bar{r} * \bar{s}$ (purple) and $f|_A$ (light green), respectively.

It hence remains to show that $f|_A(x) = \int_{S_x} r(x-y)r(y)dy \forall x \in A$. The integrand $r(x-y)r(y) = 0 \forall y < 0 \vee y > x$. The difference in the integration domains can be seen in fig. 1.4. We get two cases.

Let $x \in A$. Assume $\exists y' \in (M/2, M)$ such that $r(x-y')r(y') > 0$. Let $x' := 2y'$. We then have $r(y')^2 = r(x'-y')r(y') > 0$, and due to continuity $f(x') > 0$, contradiction, because $x' \in (M, 2M)$.

Analogously fix $x' \in (M/2, M)$. Assume $\exists y' \in (0, x' - M/2)$ such that $r(x'-y')r(y') > 0$, and thus $r(x'-y') > 0$, where $a := x' - y' > M/2, 2a \in (M, 2M)$. Then $r(a)^2 = r(2a-a)r(a) > 0$, due to continuity $f(2a) > 0$, again contradiction. \square

Proposition 1.76. Let $\mathcal{C}_{c,b}^+$ denote the set of piecewise continuous nonnegative functions of bounded support. Then the set of nondivisible functions, $\mathcal{I} := \{f : \nexists r \in \mathcal{C}_{c,b} : f = r * r\}$ is dense in $\mathcal{C}_{c,b}$.

Proof. It is enough to show the claim for functions $f \in \mathcal{C}_{c,b}^+$ with $\inf \text{supp } f \geq 0$. Let $\epsilon > 0$, and $M := \sup \text{supp } f$. Take $j \in \mathcal{C}_{c,b}$ to be nondivisible with $\text{supp } j \subset (2M, 3M)$, and define

$$g(x) := \begin{cases} f(x) & x < M \\ \epsilon j(x) / \|j\|_\infty & x \in (2M, 3M) \\ 0 & \text{otherwise.} \end{cases}$$

By construction, $\|f - g\|_\infty < \epsilon$, but $g|_{(2M, 3M)} \equiv j$ is not divisible, hence by lemma 1.75 g is not divisible, and the claim follows. \square

Corollary 1.77. Let $\epsilon > 0$. Let X be a continuous random variable with pdf $p_X(k)$. Then there exists a nondivisible random variable Y with pdf $p_Y(k)$, such that $\|p_X - p_Y\| < \epsilon$.

Proof. Let $\epsilon > 0$ small. Since $\mathcal{C}_{e,b} \subset \{f \text{ integrable}\} =: L$, we can pick $f_M \in L : \text{supp } f_M \in (-M, M)$, $\|p_X - f_M\| < \epsilon/3$ and $\|f_M\| = 1 + \delta$ with $|\delta| \leq \epsilon/3$. Then

$$\left\| p_X - \frac{f_M}{\|f_M\|} \right\| = \left\| p_X - \frac{f_M}{1 + \delta} \right\| \leq \|p_X - f_M\| + \frac{\epsilon}{2} \|f_M\| \leq \epsilon.$$

and proposition 1.76 finishes the claim. □

Corollary 1.78. *Any weak membership formulation of divisibility in the continuous setting is trivial to answer, as for all $\epsilon > 0$, there always exists a nondivisible distribution ϵ close to the one at hand. Similar considerations apply to other formulations of the continuous divisibility problem.*

1.2.3.6 Infinite Divisibility

Let us finally and briefly discuss the case of infinite divisibility. While interesting from a mathematical point of view, the question of infinite divisibility is ill-posed computationally. Trivially, discrete distributions cannot be infinitely divisible, as follows directly from theorem 1.68. A similar argument shows that neither the ϵ , nor the weak variant of the discrete problem is a useful question to ask, as can be seen from lemma 1.72 and 1.74.

By the same arguments as in section 1.2.3.5, the weak membership version is easy to answer and thus trivially in P.

1.2.4 Decomposability

1.2.4.1 Computational Problems

We define the decomposability analogue of definition 1.66 and 1.67 as follows.

Definition 1.79 (DISTRIBUTION DECOMPOSABILITY).

Instance. Finite discrete random variable $X \sim \mathcal{D}$.

Question. Do there exist finite discrete distributions $\mathcal{D}', \mathcal{D}'' : X = Z_1 + Z_2$ for random variables $Z_1 \sim \mathcal{D}', Z_2 \sim \mathcal{D}''$?

Definition 1.80 (WEAK DISTRIBUTION DECOMPOSABILITY $_{\epsilon}$).

Instance. Finite discrete random variable $X \sim \mathcal{D}$ with pmf $p_X(k)$.

Question. If there exists a finite discrete random variable Y with pmf $p_Y(k)$, such that $\|p_X - p_Y\|_{\infty} < \epsilon$ and such that

1. Y is decomposable—return YES

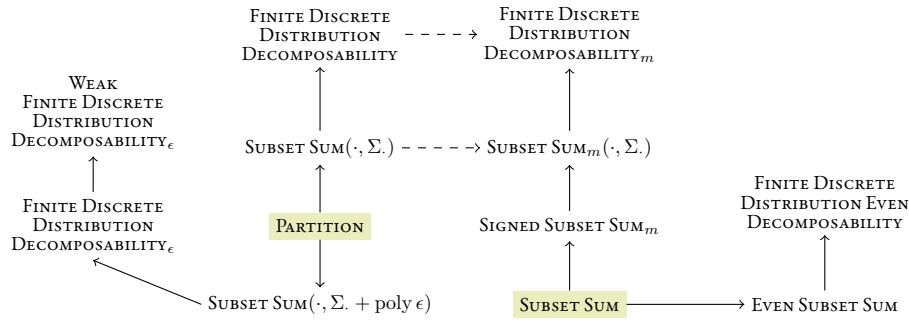


Figure 1.5: Complete chain of reduction for our discrete programs. The dashed lines are obvious and not mentioned explicitly.

2. Y is indecomposable—return No.

In this section, we will show that DISTRIBUTION DECOMPOSABILITY is NP-hard, for which we will need a series of intermediate results. Requiring the support of the first random variable Z_1 to have a certain size, i.e. $|\text{supp}(p_{\mathcal{D}'})| = m$, yields the following program.

Definition 1.81 (DISTRIBUTION DECOMPOSABILITY $_m$, $m \geq 2$).

Instance. Finite discrete random variable $X \sim \mathcal{D}$ with $|\text{supp}(p_{\mathcal{D}'})| > m$.

Question. Do there exist finite discrete distributions $\mathcal{D}', \mathcal{D}'' : X = Z_1 + Z_2$ for random variables $Z_1 \sim \mathcal{D}', Z_2 \sim \mathcal{D}''$ and such that $|\text{supp}(p_{\mathcal{D}'})| = m$?

We then define DISTRIBUTION EVEN DECOMPOSABILITY to be the case where the two factors have equal support.

The full reduction tree can be seen in fig. 1.5.

Analogous to lemma 1.35, we state the following observation.

Lemma 1.82. All the above DECOMPOSABILITY problems in definition 1.79 to 1.80 are contained in NP.

Proof. It is straightforward to construct a witness and a verifier that satisfies the definition of the decision class NP. For example in definition 1.89, a witness is given by two tables of numbers which are easily checked to form finite discrete distributions. Convolving these lists and comparing the result to the given distribution can clearly be done in polynomial time. Both verification and witness are thus poly-sized, and the claim follows. \square

1.2.4.2 Even Decomposability

We continue by proving that DISTRIBUTION EVEN DECOMPOSABILITY is NP-hard. We will make use of a variant of the well-known SUBSET SUM problem, which is NP-hard—see lemma 1.23 for a proof. The interested reader will find a rigorous digression in section 1.1.2.

This immediately leads us to the following intermediate result.

Lemma 1.83. *DISTRIBUTION EVEN DECOMPOSABILITY is NP-hard.*

Proof. Let (S, l) be an instance of EVEN SUBSET SUM. We will show that there exists a polynomial $f \in \mathcal{R}$ of degree $2|S|$ such that f is divisible into $f = g \cdot h$ with $\deg g = \deg h$ iff (S, l) is a YES instance. We will explicitly construct the polynomial $f \in \mathcal{R}$. As a first step, we transform the EVEN SUBSET SUM instance (S, l) , making it suited for embedding into f .

Let $N := |S|$ and denote the elements in S with s_1, \dots, s_N . We perform a linear transformation on the elements s_i via

$$\bullet \quad b_i := a \left(s_i - \frac{1}{|S|} \sum_{s \in S'} s \right) + \frac{al}{2|S|} \quad \text{for } i = 1, \dots, N, \quad (1.8)$$

where $a \in \mathbb{R}_{>0}$ is a free scaling parameter chosen later such that $|b_i| < \delta \in \mathbb{R}_+$ small. Let $B := \{b_1, \dots, b_N\}$. By lemma 1.24, we see that $\text{EVEN SUBSET SUM}(S, l) = \text{EVEN SUBSET SUM}(B, al)$. Since further $\sum_i b_i = al/2 > 0$, we know that (B, al) is a YES instance if and only if there exist two non-empty

- disjoint subsets $B_1 \cup B_2 = B$ with $|B_1| = |B_2|$ such that both

$$\sum_{i \in B_1} b_i > 0 \quad \text{and} \quad \sum_{i \in B_2} b_i > 0. \quad (1.9)$$

The next step is to construct the polynomial f and prove that it is divisible into two polynomial factors $f = g \cdot h$ if and only if (B, al) is a YES instance. We first define quadratic polynomials $g(b_i, x) := x^2 + b_i x + 1$ for $i = 1, \dots, N$, and set $f_T(x) := \prod_{b \in T} g(b, x)$ for $T \subset B$. Observe that for suitably small δ , the $g(b_i, x)$ are irreducible over $\mathbb{R}[x]$. With this notation, we claim that $f_B(x)$ has the required properties.

In order to prove this claim, we first show that for sufficiently small scaling parameter a , a generic subset

- $T \subset B$ with $n := |T|$ and $f_T(x) =: \sum_{i=0}^{2|T|} c_i x^i$, the coefficients c_i satisfy

$$c_0 = 1, \quad (1.10)$$

$$\text{sgn}(c_1) = \text{sgn}(\Sigma), \quad (1.11)$$

$$c_{2j} > 0 \quad \text{for } j = 1, \dots, |T|, \quad (1.12)$$

$$\text{sgn}(c_{2j+1}) \geq \text{sgn}(\Sigma) \quad \text{for } j = 1, \dots, |T| - 1, \quad (1.13)$$

where $\Sigma := \sum_{t \in T} t$. Indeed, if then $f_B = g \cdot h$, where $g, h \in \mathcal{R}$, then $g = f_{B_1}$ and $h = f_{B_2}$ for aforementioned subsets $B_1, B_2 \subsetneq B$, and conversely if (B, al) is a YES instance, then $f_B = f_{B_1} \cdot f_{B_2}$ —remember that $\mathbb{R}[x]$ is a unique factorisation domain, so all polynomials of the shape f_T necessarily

decompose into quadratic factors.

By construction, $c_0 = 1$ and $c_1 = n\Sigma$, so the first two assertions follow immediately. To address eq. (I.12) and I.13, we further split up the even and odd coefficients into

$$c_j =: \begin{cases} c_{j,0} + c_{j,2} + \dots + c_{j,j} & \text{if } j \text{ even} \\ c_{j,1} + c_{j,3} + \dots + c_{j,j} & \text{if } j \text{ odd,} \end{cases} \quad (\text{I.14})$$

where $c_{j,k}$ is the coefficient of $x^j b_{i_1} \dots b_{i_k}$. We thus have $c_{j,k} = \mathcal{O}(\delta^k)$ in the limit $\delta \rightarrow 0$ —we will implicitly assume the limit in this proof and drop it for brevity. Our goal is to show that the scaling in δ suppresses the combinatorial factors, i.e. that c_j is dominated by its first terms $c_{j,0}$ and $c_{j,1}$, respectively.

In order to achieve this, we need some more machinery. First consider $g(\delta, x) = x^2 + \delta x + 1$. It is imminent that for an expansion

$$g(\delta, x)^n =: \sum_{j=0}^{2n} x^j \sum_{k=0}^n d_{j,k} \delta^k,$$

we get coefficient-wise inequalities

$$|c_{j,k}| \leq d_{j,k} \quad \forall j = 0, \dots, 2n, k = 0, \dots, n. \quad (\text{I.15})$$

We will calculate the coefficients $d_{j,k}$ of $g(\delta, x)^n$ explicitly and use them to bound the coefficients $c_{j,k}$ of $f_T(x)$.

Using a standard Cauchy summation and the uniqueness of polynomial functions, we obtain

$$\begin{aligned}
g(\delta, x)^n &= \sum_{j=0}^n \frac{1}{j!} (1+x^2)^{n-j} x^j (n)_j \delta^j \\
&= \sum_{j=0}^n \frac{\delta^j}{j!} (n)_j x^j \sum_{k=0}^{n-j} \binom{n-j}{k} x^{2k} \\
&\equiv \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} \frac{\delta^j}{j!} (n)_j \binom{n-j}{k} x^{j+2k} \\
&= \sum_{j=0}^{\infty} \sum_{l=0}^j \frac{\delta^l}{l!} (n)_l \binom{n-l}{j-l} x^{2j-l} \\
&\equiv \sum_{j=0}^n \sum_{l=0}^j \frac{\delta^l}{l!} (n)_l \binom{n-l}{j-l} x^{2j-l} \\
&= \sum_{j=0}^n \sum_{l=j}^{2j} \frac{\delta^{2j-l}}{(2j-l)!} (n)_{2j-l} \binom{n-2j+l}{l-j} x^l.
\end{aligned}$$

With $(n)_l$ we denote the falling factorial, i.e. $(n)_l = n(n-1)(n-2)\cdots(n-l+1)$. By convention, $(n)_0 = 1$.

Addressing even and odd powers of x separately, we can thus deduce that

$$\begin{aligned}
g(\delta, x)^n &= \sum_{j=0}^{2n} x^j \begin{cases} \sum_{k=0}^{\lfloor \frac{j}{2} \rfloor} \frac{\delta^{2k+1}}{(2k+1)!} \frac{(n)_{\lceil \frac{j}{2} \rceil + k}}{(\lfloor \frac{j}{2} \rfloor - k)!} & \text{if } j \text{ odd} \\ \sum_{k=0}^{\frac{j}{2}} \frac{\delta^{2k}}{(2k)!} \frac{(n)_{\frac{j}{2} + k}}{(\frac{j}{2} - k)!} & \text{if } j \text{ even} \end{cases} \\
&= \sum_{j=0}^{2n} x^j \begin{cases} (n)_{\lceil \frac{j}{2} \rceil} \sum_{k=0}^{\lfloor \frac{j}{2} \rfloor} \frac{\delta^{2k+1}}{(2k+1)!} \frac{(n - \lceil \frac{j}{2} \rceil)_k}{(\lfloor \frac{j}{2} \rfloor - k)!} & \text{if } j \text{ odd} \\ (n)_{\frac{j}{2}} \sum_{k=0}^{\frac{j}{2}} \frac{\delta^{2k}}{(2k)!} \frac{(n - \frac{j}{2})_k}{(\frac{j}{2} - k)!} & \text{if } j \text{ even.} \end{cases}
\end{aligned}$$

A straightforward estimate shows that for the even and odd case, we obtain the coefficient scaling

$$g(\delta, x)^n = \sum_{j=0}^{2n} x^j \begin{cases} (n)_{\lceil \frac{j}{2} \rceil} \sum_{k=0}^{\lfloor \frac{j}{2} \rfloor} \delta^{2k+1} \mathcal{O}(n^k) & \text{if } j \text{ odd} \\ (n)_{\frac{j}{2}} \sum_{k=0}^{\frac{j}{2}} \delta^{2k} \mathcal{O}(n^k) & \text{if } j \text{ even,} \end{cases}$$

which means that e.g. picking $\delta = \mathcal{O}(1/n^2)$ is enough to exponentially suppress the higher order combinatorial factors.

We will now separately address the even and odd case—eq. (1.12) and 1.13.

Even Case. As the constant coefficients $c_{j,0} = \mathcal{O}(1)$ in δ , it is the same as for $g(\delta, x)^n$ and by eq. (1.15), we immediately get

$$\frac{|c_{j,2} + \dots + c_{j,j}|}{c_{j,0}} = \mathcal{O}(\delta).$$

Odd Case. Note that if $\Sigma < 0$, we are done, so assume $\Sigma > 0$ in the following. A simple combinatorial argument gives

$$c_{j,1} = \binom{n-1}{(j-1)/2} \Sigma,$$

so it remains to show that $c_{j,1} > -c_{j,3} - \dots - c_{j,j}$. Analogously to the even case, by eq. (1.15), we conclude

$$\frac{|c_{j,3} + \dots + c_{j,j}|}{c_{j,1}} = \mathcal{O}(\delta),$$

which finalises our proof. □

1.2.4.3 m-Support Decomposability

In the next two sections we will generalise the last result to $\text{DISTRIBUTION DECOMPOSABILITY}_m$. As a first observation, we note the following.

Lemma 1.84. *Let $f(n)$ be such that $(f(n)\beta(f(n), n+1-f(n)))^{-1} = \mathcal{O}(\text{poly}(n))$. Then $\text{DISTRIBUTION DECOMPOSABILITY}_{f(\cdot|\cdot)} \in \mathbf{P}$.*

Proof. See proof of theorem 1.68, and an easy scaling argument for $\binom{n}{f(n)}$ completes the proof. As in remark 1.21, this symmetrically extends to $\text{DISTRIBUTION DECOMPOSABILITY}_{|\cdot|-f(\cdot|\cdot)} \in \mathbf{P}$. □

Observe that $f(n) = n/2$ yields exponential growth, hence the remark is consistent with the findings in section 1.2.4.2.

We now address the general case. As in the last section, we need variants of the SUBSET SUM problem, SUBSET SUM_m , and $\text{SIGNED SUBSET SUM}_m$, see definitions 1.18 and 1.19. Both are shown to be \mathbf{NP} -hard in lemmas 1.20 and 1.25, or by the following observation. In order to avoid having to take absolute values in the definition of SUBSET SUM_m , we reduce it to multiple instances of $\text{SIGNED SUBSET SUM}_m$, by using the following interval partition of the entire range $(-l, l)$.

Remark 1.85. For every $a > 0, l > 0$, there exists a partition of the interval $(-l - 2a, l + 2a) = \bigcup_{i=0}^{N-1} (x_i, x_{i+1})$ with suitable $N \in \mathbb{N}$ such that $x_{i+1} - x_i = 2a$ and

$$(-l, l) = \left(\bigcup_{i=1}^{N-2} (x_i, x_{i+1}) \right) \setminus ((x_0, x_1) \cup (x_{N-1}, x_N)).$$

This finally leads us to the following result.

Lemma 1.86. *DISTRIBUTION DECOMPOSABILITY_m is NP-hard.*

Proof. We will show the reduction $\text{DISTRIBUTION DECOMPOSABILITY}_m \leftarrow \text{SUBSET SUM}_m$. Let m be fixed. Let (S, l) be an SUBSET SUM instance. For brevity, we write $\Sigma_S := \sum_{s \in S} s$. Without loss of generality, by corollary 1.17, we again assume $\Sigma_S \geq 0$. Now define $a := 2(|S|l + 2m\Sigma_S - |S|\Sigma_S)/(2m - |S|)$. Using remark 1.85, pick a suitable subdivision of the interval $(-l - 2a, l + 2a)$, such that

$$\begin{aligned} \text{SUBSET SUM}_m(S, l) &= \left(\bigvee_{i=1}^{N-2} \text{SIGNED SUBSET SUM}_m(S, x_i, x_{i+1}) \right) \\ &\quad \wedge \neg \text{SIGNED SUBSET SUM}_m(S, x_0, x_1) \\ &\quad \wedge \neg \text{SIGNED SUBSET SUM}_m(S, x_{N-1}, x_N). \end{aligned}$$

One can verify that

$$\begin{aligned} &\text{SIGNED SUBSET SUM}_m(S, x_i - a, x_i + a) \\ &= \text{SIGNED SUBSET SUM}_m(S + c(m, i), -\Sigma_{S+c(m, i)}, \Sigma_{S+c(m, i)}) \\ &= \text{SUBSET SUM}_m(S + c(m, i), \Sigma_{S+c(m, i)}), \end{aligned}$$

where we chose $c(m, i) = x_i/(2m - |S|)$. The latter program we can answer using the same argument as for the proof of lemma 1.83, and the claim follows. \square

As a side remark, this also confirms the following well-known fact.

Corollary 1.87. *Let $f(n)$ be as in lemma 1.84. Then $\text{SUBSET SUM}_{f(\cdot)} \in \text{P}$.*

1.2.4.4 General Decomposability

We have already invented all the necessary machinery to answer the general case.

Theorem 1.88. *DISTRIBUTION DECOMPOSABILITY is NP-hard.*

Proof. Follows immediately from lemma 1.83, where we consider the special set of SUBSET SUM instances for which (S, l) is such that $l = \sum_{s \in S} s$. We show in lemma 1.27 that SUBSET SUM(\cdot, Σ) is still NP-hard, thus the claim follows. \square

1.2.4.5 Decomposability with Variation

As a further intermediate result—and analogously to definition 1.71—we need to allow for a margin of error ϵ .

Definition 1.89 (DISTRIBUTION DECOMPOSABILITY $_{\epsilon}$).

Instance. Finite discrete random variable $X \sim \mathcal{D}$ with pmf $p_X(k)$.

Question. Do there exist finite discrete random variables $Z_1 \sim \mathcal{D}'$, $Z_2 \sim \mathcal{D}''$ with pmfs $p_{Z_1}(k)$, $p_{Z_2}(k)$, such that $\|p_{Z_1} * p_{Z_2} - p_X\|_{\infty} < \epsilon$?

This definition leads us to the following result.

Lemma 1.90. DISTRIBUTION DECOMPOSABILITY $_{\epsilon}$ is NP-hard.

Proof. First observe that we can restate this problem in the following equivalent form. Given a finite discrete distribution \mathcal{D} with characteristic polynomial $f_{\mathcal{D}}$, do there exist two finite discrete distributions \mathcal{D}' , \mathcal{D}'' with characteristic polynomials $f_{\mathcal{D}'}$, $f_{\mathcal{D}''}$ such that $\|f_{\mathcal{D}} - f_{\mathcal{D}'} f_{\mathcal{D}''}\|_d < \epsilon$? Here, we are using the maximum norm from definition 1.63, and assume without loss of generality that $\deg f_{\mathcal{D}} = \deg f_{\mathcal{D}'}$, $\deg f_{\mathcal{D}''}$.

As $f_{\mathcal{D}}$ is a polynomial, we can consider its Viète map $v : \mathbb{C}^n \rightarrow \mathbb{C}^n$, where $n = \deg f_{\mathcal{D}}$, which continuously maps the polynomial roots to its coefficients. It is a well-known fact—see [Whi72] for a standard reference—that v induces an isomorphism of algebraic varieties $w : \mathbb{A}_k^n / S_n \xrightarrow{\sim} \mathbb{A}_k^n$, where S_n is the n^{th} symmetric group. This shows that w^{-1} is polynomial, and hence the roots of $f_{\mathcal{D}'} f_{\mathcal{D}''}$ lie in an $\mathcal{O}(\epsilon)$ -ball around those of $f_{\mathcal{D}}$. By a standard uniqueness argument we thus know that if $f_{\mathcal{D}} = \prod_i f_i$ with $f_i = x^2 + b_i x + c_i$ as in the proof of lemma 1.83, then $f_{\mathcal{D}'} = \prod_i g_i$ with $g_i = a_i x^2 + b'_i x + c_i$, where $a_i = c_i = 1 + \mathcal{O}(\epsilon)$, $b'_i = b_i + \mathcal{O}(\epsilon)$ —we again implicitly assume the limit $\epsilon \rightarrow 0$.

We continue by proving the reduction DISTRIBUTION DIVISIBILITY $_{\epsilon} \leftarrow$ SUBSET SUM(\cdot, Σ + poly ϵ), which is NP-hard as shown in lemma 1.28. Let $S = \{s_i\}_{i=1}^N$ be a SUBSET SUM multiset. We claim that it is satisfiable if and only if the generated characteristic function $f_S(x)$ —where we used the notation of the proof of lemma 1.83—defines a finite discrete probability distribution and the corresponding random variable X is a YES instance for DISTRIBUTION DIVISIBILITY $_{\epsilon}$.

First assume f_S is such a YES instance. Then $\sum_{s \in S} s \geq 0$, and there exist two characteristic polynomials $g = \prod_i g_i$ and $h = \prod_i h_i$ as above and such that $\|f_S - gh\|_d < \epsilon$. We also know that if $g_i = a_i x^2 + b_i x + c_i$, then $\exists T \subsetneq S$ such that $\{b_i\}_i \in B_{\epsilon}(T) \subseteq \mathbb{R}^{|T|}$, where $T \subsetneq S$ and $B_{\epsilon}(T)$ denotes an ϵ ball

around the set T , and analogously for $h_i = a'_i x^2 + b'_i x + c'_i$, with $\{b'_i\}_i \in B_\epsilon(S \setminus T) \subseteq \mathbb{R}^{|S|-|T|}$. For the linear coefficients, we thus have

$$\begin{aligned} \left| \sum_{s \in S} s - \sum_{t \in T} t - \sum_{s \in S \setminus T} s \right| &= \left| \sum_{s \in S} s - \sum_{i=1}^{|T|} b_i - \sum_{i=1}^{|S \setminus T|} b'_i + \mathcal{O}(\epsilon) \right| \\ &\leq \mathcal{O}(\epsilon) \leq \sum_{s \in S} s + \mathcal{O}(\epsilon). \end{aligned} \tag{1.16}$$

Now the case if f_S is a No instance. Assume there exists a nontrivial multiset $T \subsetneq S$ satisfying

$$\left| \sum_{t \in T} t - \sum_{s \in S \setminus T} s \right| < \sum_{s \in S} s + \mathcal{O}(\epsilon).$$

Then by construction $\sum_{t \in T} t, \sum_{s \in S \setminus T} s \geq -\mathcal{O}(\epsilon)$ and $f_T \cdot f_{S \setminus T} = f_S$, contradiction, and the claim follows. \square

1.2.4.6 Weak Decomposability

Analogously to section 1.2.3.4, we now address the weak membership problem of decomposability.

Theorem 1.91. *WEAK DISTRIBUTION DECOMPOSABILITY $_\epsilon$ is NP-hard.*

Proof. In order to show the claim, we prove the reduction $\text{WEAK DISTRIBUTION DECOMPOSABILITY}_\epsilon \leftarrow \text{DISTRIBUTION DECOMPOSABILITY}_{g(\epsilon)}$, where the function $g = \mathcal{O}(\epsilon)$. It is clear that the polynomial factor leaves the NP-hardness of the latter program intact.

We use the same notation as in the proof of lemma 1.90. Let f_S be a YES instance of $\text{DISTRIBUTION DECOMPOSABILITY}_\epsilon$, and define $S' := \{s + \mathcal{O}(\epsilon) : s \in S\}$. From eq. (1.16) it immediately follows that then $f_{S'}$ is a YES instance of $\text{DISTRIBUTION DECOMPOSABILITY}_{g(\epsilon)}$, where we allow $g = \mathcal{O}(\epsilon)$. We have hence shown that there exists an $\mathcal{O}(\epsilon)$ ball around each YES instance that *solely* contains YES instances.

A similar argument holds for the No instances. It is clear that these cases can be answered using $\text{WEAK DISTRIBUTION DECOMPOSABILITY}_\epsilon$, and the claim follows. \square

1.2.4.7 Complete Decomposability

Another interesting question to ask is for the complete decomposition of a finite distribution \mathfrak{D} into a sum of indecomposable distributions. We argue that this decomposition is not unique.

Proposition 1.92. *There exists a family of finite distributions $(\mathfrak{D}_n)_{n \in \mathbb{N}}$ with probability mass functions $p_n(k) : \max \text{supp } p_n(k) = 4n$ and such that, for each \mathfrak{D}_n , there are at least $n!$ distinct decompositions into indecomposable distributions.*

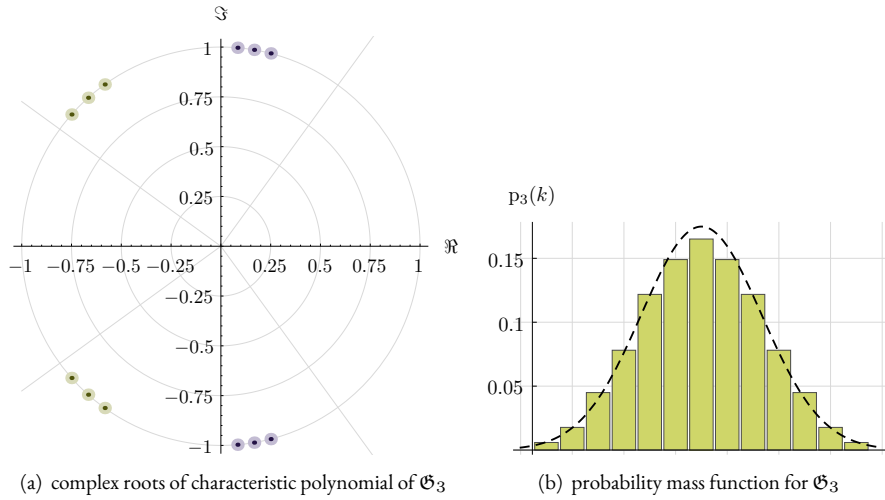


Figure 1.6: Counterexample construction of proposition 1.92. The dashed line shows a normal distribution for comparison.

Proof. We explicitly construct the family $(\mathfrak{D}_n)_{n \in \mathbb{N}}$. Let $n \in \mathbb{N}$. We will define a set of irreducible quadratic polynomials $\{p_k, n_k$ for $k = 1, \dots, n\}$ such that n_k are *not* positive, but $p_k n_l$ are positive quartics $\forall k, l$ —and thus define valid probability distributions. Since $\mathbb{R}[x]$ is a unique factorisation domain the claim then follows.

Following the findings in the proof of lemma 1.83, it is in fact enough to construct a set $\{a_k, b_k : 0 < |a_k| < 2, -2 < b_k < 0$ for $k = 1, \dots, n\} \subset \mathbb{R}^{2n}$ and such that $a_k + b_l > 0 \forall k, l$ —then let $p_k := 1 + a_k x + x^2, n_k := 1 + b_k x + x^2$. It is straightforward to verify that e.g.

$$a_k := 1 + \frac{k}{2n} \quad \text{and} \quad b_k := -\frac{k}{2n}$$

fulfil these properties. □

Remark 1.93. Observe that for $b_k := -k/2n^2$, the construction in proposition 1.92 allows decompositions into m indecomposable terms, where $m = n, \dots, 2n$.

Corollary 1.94. \mathcal{R} is not a unique factorisation domain.

Proposition 1.92 and remark 1.93 show that an exponential number of complete decompositions—all of which have different distributions—do not give any further insight into the distribution of interest—indeed, as the number of positive indecomposable factors is not even unique, asking for a non-maximal decomposition into indecomposable terms does not answer more than whether the distribution is decomposable at all.

Indeed, the question whether one *can* decompose a distribution into indecomposable parts can be trivially answered with YES, but if we include the condition that the factors have to be non-trivial, or for decomposability into a certain number of terms—say $N \geq 2$ or the maximum number of terms—the problem is also obviously NP-hard by the previous results.

In short, by theorem 1.88, we immediately obtain the following result.

Corollary 1.95. *Let \mathcal{D} be a finite discrete distribution. Deciding whether one can write \mathcal{D} as any nontrivial sum of irreducible distributions is NP-hard.*

1.2.4.8 Continuous Distributions

Analogous to our discussion in section 1.2.3.5, the exact and ϵ variants of the decomposability question are computationally ill-posed. We again point out that answering the weak membership version is trivial, since the set of indecomposable distributions is dense, as the following proposition shows.

Proposition 1.96. *Let $\mathcal{C}_{c,b}^+$ denote the piecewise linear nonnegative functions of bounded support. Then the set of indecomposable functions, $\mathcal{J} := \{f : \nexists r, s \in \mathcal{C}_{c,b} : f = r * s\}$ is dense in $\mathcal{C}_{c,b}$.*

Proof. We first extend lemma 1.75, and again take $f \in \mathcal{C}_{c,b}^+ : \text{supp } f \subset A \cup B$. While not automatically true that $r(x), s(x) = 0 \forall x < 0$, we can assume this by shifting r and s symmetrically. We also assume $\inf \text{supp } f = 0$, and hence $\inf \text{supp } r = \inf \text{supp } s = 0$ —see lemma 1.58 for details.

Since $f(x) = 0 \forall x \in (M, 2M)$, we immediately get $r(x) = s(x) = 0 \forall x \in (M, 2M)$. Furthermore, $\exists m \in (0, M) : r(x) = s(y) = 0 \forall x \in (m, M], y \in (M - m, M]$. Analogously to eq. (1.7), we define

$$\bar{r}(x) = \begin{cases} r(x) & x \in [0, m] \\ 0 & \text{otherwise} \end{cases} \quad \text{and} \quad \bar{s}(x) = \begin{cases} r(x) & x \in [0, M - m] \\ 0 & \text{otherwise.} \end{cases} \quad (1.17)$$

The integration domain difference is derived analogously, and can be seen in an example in fig. 1.4. We again consider the two cases separately.

Let $x \in A$. Assume $\exists y' \in (M - m, M)$ such that $r(x - y')s(y') > 0$. Then $s(y') > 0$, contradiction. Now fix $x' \in (m, M)$, and assume $\exists y' \in (0, x' - m) : r(x' - y')s(y') > 0$. Since $x' - y' > x' - x' + m = m$, $r(x' - y') > 0$ yields another contradiction.

The rest of the proof goes through analogously. □

Corollary 1.97. *Let $\epsilon > 0$. Let X be a continuous random variable with pmf $p_X(k)$. Then there exists a indecomposable random variable Y with pmf $p_Y(k)$, such that $\|p_X - p_Y\| < \epsilon$.*

Proof. See corollary 1.77. □

1.3 Chapter Summary

In section 1.1, we have shown that the question of existence of a stochastic root for a given stochastic matrix is in general at least as hard as answering 1-IN-3SAT, i.e. it is NP-hard. By corollary 1.41, this NP-hardness result also extends to NONNEGATIVE and DOUBLY STOCHASTIC DIVISIBILITY, which proves theorem 1.1. A similar reduction goes through for CPTP DIVISIBILITY in corollary 1.38, proving NP-hardness of the question of existence of a CPTP root for a given CPTP map.

In section 1.2, we have shown that—in contrast to CPTP and stochastic matrix divisibility—distribution divisibility is in P, proving theorem 1.4. On the other hand, if we relax divisibility to the more general decomposability problem, it becomes NP-hard as shown in theorem 1.6. We have also extended these results to weak membership formulations in theorem 1.5 and 1.7—i.e. where we only require a solution to within ϵ in the appropriate metric—showing that all the complexity results are robust to perturbation.

Finally, in section 1.2.3.5 and 1.2.4.8, we point out that for continuous distributions—where the only computationally the only meaningful formulations are the weak membership problems or closely related variants—questions of divisibility and decomposability become computationally trivial, as the nondivisible and indecomposable distributions independently form dense sets.

As containment in NP for all of the NP-hard problems is easy to show (lemma 1.35 and 1.82), these problems are also NP-complete. Thus our results imply that, apart for the distribution divisibility problem which is efficiently solvable, all other divisibility problems for maps and distributions are equivalent to the famous $P = NP$ conjecture, in the following precise sense: A polynomial-time algorithm for answering any one of these questions—(DOUBLY) STOCHASTIC, NONNEGATIVE or CPTP DIVISIBILITY, or either of the DECOMPOSABILITY variants—would prove $P = NP$. Conversely, solving $P = NP$ would imply that there exists a polynomial-time algorithm to solve all of these DIVISIBILITY problems.

In the next chapter, we will be discussing another static problem of quantum many-body systems—the question of approximating ground state energies of local Hamiltonians to within a certain precision. Interestingly, as we will see, this will firmly move us into the realm of *quantum* complexity classes like QMA and BQP, and away from NP and P. However, since the quantum world is intrinsically probabilistic, this should not come as a surprise.

2 Hamiltonian Complexity: Turing’s Wheelbarrow

If I seem to wander, if I seem to stray,
remember that true stories
seldom take the straightest way.
—Patrick Rothfuss, *the Name of the Wind*

Complex physical behaviour can emerge from even very simple rules. Yet if the system is *too* simple, one can often rule out the possibility of any exotic behaviour. Just how simple can a system be to nonetheless feature complex properties? Much of the progress in Hamiltonian complexity and related areas over the last decade can be viewed as improving our understanding of where this boundary between simple and complex lies.

For example, consider 1D spin chains with translationally-invariant nearest neighbour interactions. Hastings proved that if the Hamiltonian describing the system is gapped, the ground state entanglement has to follow an area law [Has07]. In 1D, the area law means that the entanglement entropy between any contiguous region and its complement is upper-bounded by a constant, independent of the size of the region. It was believed that even for non-gapped Hamiltonians, area-law violations would contribute at most log corrections in the system size. Such long-range correlations in a spin chain’s ground state which scale with the system’s size are a common indicator of criticality, i.e. they show that the system is close to a quantum phase transition. The entanglement entropy is then expected to scale logarithmically with the number of spins, since critical spin chains can often be related to a conformal field theory.

However, using Hamiltonian complexity techniques, Irani [Ira07] constructed an example of a spin chain in 1D that exhibits violation of the area-law *beyond* logarithmic corrections, indicating that one cannot describe such behaviour by a conformal field theory. Irani’s construction breaks translational-invariance, so it cannot directly be compared to systems satisfying area laws. A later construction [GI13] can give a similar area-law violation whilst preserving translational-invariance. However, the required local dimension, $O(10^6)$, is vast. It is therefore at best questionable whether this area-law violation could ever be observed in practice. Does this mean that such violations only occur for some peculiar theoretical models with non-translationally-invariant couplings, or unrealistically large Hilbert space dimensions?

- We now know that the answer to this question is negative. First, it was shown by Bravyi2012a that, even for frustration-free spin-1 chains (i.e. local dimension 3), one can construct interactions that yield highly entangled ground states, indicating critical behaviour. In fact, this result delineates a strict dimension
- threshold for the presence of ground-state entanglement in frustration-free systems. For frustration-free spin-1/2 chains (i.e. local dimension 2) with translationally-invariant nearest neighbour interactions, it was already known that ground states are unentangled [Che+11]. Building on this, Movassagh et al. [MS14] constructed models which give power-law violation of the area-law for translationally-invariant spin-5/2 chains (i.e. local dimension¹ 6), significantly improving on the bound on the local dimension threshold for power-law area-law violation from Gottesman and Irani’s result.

Similar dimension-related physicality questions also surround Cubitt et al.’s result which proves that deciding whether a system is gapped or gapless in the thermodynamic limit is an undecidable problem, even for 2D spin lattices with translationally-invariant local interactions [CPW15a]. Again, the local Hilbert space dimension in the model they describe is vast. Bravyi and Gosset recently derived necessary and sufficient conditions for a gapped or gapless phase for frustration-free spin-1/2 chains [BG15]. So at the other end of the local dimension scale, the spectral gap problem is decidable in some cases. However, there is evidence that an astronomical local dimension may not be a fundamental ingredient in the emergent behaviour that gives rise to undecidability of the spectral gap. The abrupt change in the spectrum at very large system sizes that is behind the undecidability, can also occur on 2D lattices of far lower-dimensional spins [Bau+16]. Again, this poses an immediate question of whether there is some local dimension threshold above which undecidability can occur, but below which it cannot.

- The original and most widely-studied question in Hamiltonian complexity theory, however, is that of estimating the ground state energy of a local Hamiltonian. Kitaev showed that this problem is QMA-hard [KSV02] (i.e. at least as hard as every other problem in the complexity class QMA—the quantum generalisation of NP). Similar to a spin glass, when cooled down these QMA-hard systems are predicted to get stuck in one of their many meta-stable configurations, and will take exponentially long (in the system size) to find their global minimum-energy configuration. QMA-hardness-inspired constructions lie behind all the results mentioned above. Yet even though the parameters describing QMA hard ground state Hamiltonians have been improved successively [KKR06]; [OT05]; [Aha+09b]; [HNN13]; [GI13], a lower local dimension threshold below which systems cannot feature spin-glass-like frustration is not known; for non-translationally-invariant systems we know that this bound can be at most 8. For the more physically relevant case of spin chains *with* translational symmetry, however, the best-known bound is $O(10^6)$, due to Gottesman and Irani [GI13], which is unphysically large. From a physical perspective it makes a dramatic
-

¹[MS14] in fact prove their result for local dimension 5 but breaking strict translational invariance by adding boundary terms at the ends of the chain. Using a trick due to [GI13], the boundary terms can be removed at the cost of increasing the local dimension by 1.

difference if the complexity threshold is e.g. 7, or 1000.

In this work, we improve the best-known upper bound on the local Hilbert space dimension required for QMA-hardness in translationally-invariant spin chains by several orders of magnitude, showing that the question of estimating the ground state energy of a local translationally-invariant Hamiltonian with nearest-neighbour interactions remains hard, even for spins on a chain with local dimension ≈ 40 .

2.1 Extended Introduction and Overview of Results

2.1.1 Historical Context

Hamiltonians are the one-stop shop for describing physical properties of multi-body quantum systems, and are of paramount interest for an array of disciplines ranging from experimental condensed matter physics to theoretical computer science [OT05]; [KKR06]; [Aha+09b]; [GI13]; [BH12]; [CM14]; [PM15]; [CPW15a]; [WL15]. While computer scientists are interested in the computational power of different models, for physicists it is important to calculate the structure of the low-energy spectrum of quantum systems, in particular to approximate the minimum energy of the system, i.e. the ground state energy.

The decision problem of determining whether such a local Hamiltonian operator has lowest energy—or eigenvalue—below some α or above some β , with $\beta > \alpha$, can be thought of as the quantum analogue of the maximum satisfiability problem MAX-SAT. Similar to the well-known 3-SAT, this asks for the maximum number of clauses of a Boolean formula in conjunctive normal form that can be satisfied simultaneously. In the quantum case, each local term \mathbf{h} of \mathbf{H} is analogous to a clause while a global state $|\psi\rangle$ is analogous to a global variable assignment, and the smaller $\langle\psi|\mathbf{h}|\psi\rangle$ is, the closer $|\psi\rangle$ is to satisfying the corresponding clause \mathbf{h} . The LOCAL HAMILTONIAN problem formalises the notion of maximizing the number of local terms of \mathbf{H} which can be simultaneously minimised by some global state $|\psi\rangle$, in the sense that $\langle\psi|\mathbf{H}|\psi\rangle$ is small. Physically, this minimum is equal to the lowest energy of the system.

Formally, we can state the LOCAL HAMILTONIAN problem as the following promise problem.

Definition 2.1 (k -LOCAL HAMILTONIAN).

Input. An integer n and a k -local Hamiltonian \mathbf{H} on a multipartite Hilbert space $(\mathbb{C}^d)^{\otimes n}$, and two real numbers $\beta > \alpha$ such that $\beta - \alpha \geq 1/p(n)$, for some fixed polynomial $p(n)$. The smallest eigenvalue λ_{\min} of \mathbf{H} is promised to be either smaller than α or greater than β .

Question. Is $\lambda_{\min} < \alpha$? •

The ground state energy of a many-body quantum system plays a crucial role in physics, and the question of providing an estimate for it for a system at hand—as captured by this definition of the k -LOCAL HAMILTONIAN problem—is an active field of research. But how hard is it as a computational problem?

	locality	local dimension	geometry and symmetries
Kitaev (1999)	5	2	arbitrary
Kempe, Kitaev, Regev [KKR06]	2	2	arbitrary
Oliveira, Terhal [OT05]	2	2	2D, planar, nearest-neighbour interactions
Aharonov, Gottesman, Irani, Kempe [Aha+09b]	2	12	line, nearest-neighbour
Hallgren, Nagaj, Narayanaswami [HNN13]	2	8	line, nearest-neighbour
Gottesman, Irani [GI13]	2	huge ($\approx 10^6$)	line, nearest-neighbour, translationally-invariant

Table 2.1: Brief historic overview of QMA (QMA_{EXP} for [GI13]) completeness results in Hamiltonian complexity.

- The complexity of k -LOCAL HAMILTONIAN has a track record of long-standing interest (cf. table 2.1). The foundations were laid with Feynman’s paper [Fey86] on encoding quantum circuits into the ground state of a Hamiltonian, which motivated a whole series of interesting and increasingly sophisticated results showing that variants of this problem are QMA- or QMA_{EXP}-complete.²

On the other hand, just as in classical computer science 2-SAT is solvable in polynomial time, its quantum analogue—the QUANTUM 2-SAT, a special case of the 2-LOCAL HAMILTONIAN problem³—can also be solved deterministically in polynomial time: [Bra11] proved an $O(n^4)$ runtime bound, and later a linear-time algorithm was discovered independently by [Ara+15] and [BG16]. Yet the resemblance with classical results goes further: QUANTUM 4-SAT and later QUANTUM 3-SAT were shown to be QMA₁-complete [Bra11]; [GN13]. In the same spirit, a recent result shows that in case of one-dimensional *gapped* local Hamiltonians, there exists an efficient randomised algorithm for approximating the ground state as a matrix product state [LVV15] (this result is independent of the local dimension).

- However, the LOCAL HAMILTONIAN problem, as defined in definition 2.1, allows the Hamiltonian to be frustrated (e.g. by going beyond local projectors), and encompasses Hamiltonians whose gap closes inverse-polynomially in the system’s size. It is thus a natural question to ask whether this more general LOCAL HAMILTONIAN problem remains computationally hard, even under restrictions motivated on *physical* grounds (e.g. for translationally-invariant interactions and for qubits), or whether there is a fundamental local dimension threshold below which it becomes tractable.

²QMA_{EXP} is to QMA what NEXP is to NP. This is a necessary technicality whenever the input has to be specified in unary. The energy gap still scales inverse-polynomially with system size n , and the physical implications are exactly the same as for QMA-completeness. We define these complexity classes rigorously in section 2.2.1.2, and explain their difference in detail in section 2.2.1.4.

³More specifically, QUANTUM 2-SAT asks whether a sum of 2-local terms, where each term is a 2-qubit projector that acts on any pair of qubits, is frustration-free, i.e. has a 0-energy eigenstate or, equivalently, a state that simultaneously satisfies all local constraints.

To motivate this further, it is crucial to note that Hamiltonian constructions in the spirit of [Fey86] are a proof-of-concept and may not necessarily be natural, in the sense that we would not encounter them in nature describing an actual physical system. There are three fundamental criteria for judging the “physicality” of a Hamiltonian: the interactions should be geometrically local, the dimension of the interacting subsystems should be small, and the interactions should exhibit translational invariance. These properties apply to physical systems we typically encounter in nature. For example, translational invariance means that if the Hamiltonian is specified on a lattice, the interactions are the same independently of the location within the lattice.

Starting with Kitaev’s original proof of QMA-completeness of 5-LOCAL HAMILTONIAN [KSV02], the locality and local dimension of the constructions were improved successively [KKR06]; [OT05]; [Aha+09b], see table 2.1. For spins of local dimension 8 coupled by nearest-neighbour interactions on a chain, QMA-hardness was proven by Hallgren et al. [HNN13]. All of these results make heavy use of the non-translationally-invariant nature of interactions, which vastly simplify the encoding of the problem instance and verifier circuit into the local structure of the Hamiltonian. The QMA_{EXP}-hardness result by Gottesman and Irani [GI13], which features a 2-local Hamiltonian on a line with translationally-invariant nearest-neighbour interactions, shows that having translational symmetry does not change the complexity class of the LOCAL HAMILTONIAN problem. But one caveat remains: the local dimension is unphysically large, on the order of 10^6 .

2.1.2 Main Result

Our goal is to significantly improve on this best-known upper bound on the local dimension. We develop a set of new methods to prove that the complexity threshold above which the LOCAL HAMILTONIAN problem is computationally hard is at most 42, even under the strict physicality constraints outlined above. More precisely, we prove the following main theorem.

Theorem 2.2. The LOCAL HAMILTONIAN problem with translationally-invariant interactions between neighbouring spins on a chain with local dimension 42 is QMA_{EXP}-complete. This holds true even for Hamiltonians with local terms of the form $\mathbf{h} + p(n)\mathbf{b}$, where \mathbf{h} and \mathbf{b} are fixed 2-local interactions and $p(n)$ is a fixed polynomial in the chain length n .

Following the notation in [GI13], we label this class of problems 2-TILH, for translationally-invariant 2-LOCAL HAMILTONIAN. Analogous to all past hardness constructions, we prove our result by explicitly defining a family of QMA_{EXP}-hard instances of 2-TILH. More precisely, the instances we construct are so-called *history state* Hamiltonians: by choosing the local constraints in \mathbf{H} suitably, one can create a Hermitian operator with a ground state spanned by states that are a uniform superposition over the history of a computation, such that the state at step t is *entangled* with a corresponding state $|t\rangle$ in a separate time

register (i.e. $\sum_t |t\rangle \otimes |\psi_t\rangle$). Measuring the time register at time t then yields the state of the computation at this step. This “program counter”, as Feynman describes it, can be thought of as a clock or a finite automaton driving the application of quantum gates. Originally, only linearly-evolving clock constructions were used, since analysing the spectrum of a Hamiltonian with branching computational paths is more difficult. More recently, QMA-hardness constructions in 1D and 2D have used limited branching and cycles [HNN13]; [BT14a]. These have also been exploited in the slightly different context of adiabatic and Hamiltonians quantum computation [NW08]; [Nag12]; [GTV15].

Whereas recent results [PM15] make use of perturbation gadgets—approximating higher-order interactions in the low-energy subspace of the system by an effective high-energy theory—it is known that this does not work in one-dimensional systems [Aha+09b]. The improvements in [HNN13] over [Aha+09b] are possible however by approximating 4-local interactions by a sum of 2-local interactions, effectively introducing illegal transitions that have to be penalised. Perturbation gadgets and locality reduction both depend on introducing a large energy scale to project out illegal subspaces. Our results, on the other hand, do *not* use perturbation theory⁴.

Our findings are based on the following three main technical contributions:

1. All previous constructions encode one of the standard models of quantum computation (almost always the circuit model, with the exception of [GI13] which encodes a quantum Turing machine), which are not optimised for this task. We design a new universal model of quantum computation—a *quantum ring machine* (QRM)—which we prove to be quantum Turing-complete. The periodicity of the QRM’s computational steps make it particularly well-suited for local Hamiltonian constructions.
2. We next introduce *unitary labelled graphs* and their associated Hamiltonians, which can accommodate a non-deterministic clock construction to drive quantum computation. This vastly generalises Feynman’s original clock construction [Fey86], which corresponds to a path graph in our setup. Mirroring Kitaev’s analysis [KSV02], our Hamiltonian is also equivalent to a Laplacian of the corresponding graph, which allows us to analyse its spectrum using a combination of spectral graph theory and matrix analysis techniques. These techniques let us analyse ground states of much more complicated Hamiltonians than previously possible.
3. We define yet another computational model—a *quantum Thue system*, or a quantum string rewriting system—that on the one hand is particularly well-suited for embedding a computational model into local interactions of a Hamiltonian; and on the other hand, under simple *local* constraints on the

⁴The polynomial in theorem 2.2 is an artefact of the construction. A standard trick from [GI13] can reduce the Hamiltonian to fixed $O(1)$ interactions by slightly increasing the local dimension, see remark 2.73.

rewriting rules, necessarily produces Hamiltonians that correspond to unitary labelled graphs. Quantum Thue systems can in a sense be thought of as an assembly language for compiling computational models into local translationally-invariant Hamiltonians, which could also be used for adiabatic quantum computation, or Hamiltonian quantum computers (cf. [NW08]; [WL15]).

In light of our result being rather involved and technical, we want to give a poor man’s overview of our findings, which—without any proofs—outline the technical contributions in this chapter. We want to emphasise that we made an effort to keep each section largely self-contained; in particular the section on spectral analysis of graphs with unitary edge labels, quantum ring machines, and quantum Thue systems can be read independently of each other. The QMA_{EXP} -hardness proof in section 2.2.5 of course utilises all of our developed machinery, but in such a way that the proof of existence of QMA_{EXP} -hard instances themselves are given a separate section.

Since the latter part is somewhat technical and specific, we want to point out that one does not need to understand the construction itself to follow the idea behind the hardness proof, which hopefully facilitates an understanding of the result.⁵

2.1.3 Proof Ideas and Techniques

2.1.3.1 Spectral Analysis for Hamiltonians Encoding Non-Deterministic Computation

As briefly explained in the introduction, the fundamental idea behind encoding quantum computation into the ground state of a Hamiltonian is based on the concept of *history states*, introduced by Feynman in 1986. For some quantum circuit represented by local gates $\mathbf{U}_1, \dots, \mathbf{U}_T$ on a Hilbert space \mathcal{H} , we define a Hamiltonian on the product space $\mathbb{C}^T \otimes \mathcal{H}$ as

$$\mathbf{H} := \sum_{t=1}^{T-1} (|t\rangle\langle t| \otimes \mathbb{1} + |t+1\rangle\langle t+1| \otimes \mathbb{1} - |t+1\rangle\langle t| \otimes \mathbf{U}_t - |t\rangle\langle t+1| \otimes \mathbf{U}_t^\dagger). \quad (2.1)$$

The ground state of this Hermitian operator is spanned by states of the form $\sum_t |t\rangle \otimes |\psi_t\rangle$, where $|\psi_t\rangle = \mathbf{U}_t \cdots \mathbf{U}_1 |\phi\rangle$ for some $|\phi\rangle \in \mathcal{H}$. For any $|\phi\rangle$, $\ker \mathbf{H}$ thus encodes the uniform superposition over the history of the quantum circuit acting on $|\phi\rangle$. An intuitive way of thinking about these ground states is that they represent quantum computation driven by a clock, i.e. for each increment of the clock register, the corresponding quantum gate is applied to the computational register.⁶

⁵The reason behind unhitching the explicit construction of QMA_{EXP} -hard instances in this way is to allow for further optimisation of the local dimension to go through without having to re-prove all of the claims; in fact, we encourage the interested reader to have a stab at finding a quantum Thue systems following the four properties given in lemma 2.59, but with an alphabet that is smaller than ours.

⁶The notion of time in this context is meaningless, but simplifies an intuitive understanding on how computation is embedded into the ground state of \mathbf{H} .

Essentially all past results employ such history state Hamiltonians with a linear clock, i.e. for every computational step, there exists precisely one unique forward and backward transition. For local Hamiltonian constructions—i.e. where \mathbf{H} is a sum of local terms—this implies that each *local* rule has to know the exact location within the overall computation.

To be more specific, consider a spin chain of length n as the Hilbert space $\mathcal{H}_{\text{loc}}^{\otimes n}$. The interactions on this chain then take the form of a set of local *rewriting rules* acting on neighbouring sections of spins: for $|\psi_i\rangle, |\phi_i\rangle \in \mathcal{H}_{\text{loc}}^{\otimes k}$ for some constant $k < n$, we encode the evolution $|\psi_i\rangle \mapsto |\phi_i\rangle$ by a local Hamiltonian term $\mathbf{h}_i = (|\psi_i\rangle - |\phi_i\rangle)(\langle\psi_i| - \langle\phi_i|)$. The overall Hamiltonian is then a sum of these local interactions over all spins, i.e.

$$\mathbf{H} = \sum_j \mathbb{1}_{1, \dots, j-1} \otimes \left(\sum_i \mathbf{h}_i \right)_{j, \dots, j+k-1} \otimes \mathbb{1}_{j+k, \dots, n}. \quad (2.2)$$

If the global evolution defined by the terms \mathbf{h}_i is unique, this implies that it is always possible to locally determine the global state of the computation.⁷ This means that locally, we have to store this state in one way or another: under this requirement it is difficult to push the limits of local Hilbert space dimension down, and much could be gained if we could e.g. allow the local computational state to be ambiguous to some extent (but such that if the wrong transition is applied, the computation does not proceed to tamper with the actual outcome of the embedded circuit).

In our work, we go beyond linear clock constructions, and prove a series of spectral graph-theoretic results which allow us to analyse more complicated history state Hamiltonians. We outline these novel techniques below.

If all $|\psi_i\rangle$ and $|\phi_i\rangle$ in eq. (2.2) are standard basis vectors, then each rule corresponds to an edge in a graph G with vertices labelled by the canonical basis of the spin chain. \mathbf{H} thus equals the Laplacian of the graph G (whose spectrum is accessible) and the ground state of \mathbf{H} is given by the uniform superposition over connected graph components of G . We call ground states of such Hamiltonians as in eq. (2.2) *history states*, since they encode the closure of states reachable under the given rewriting rules.

To analyse the spectrum of more general non-basis transitions $|\psi_i\rangle \mapsto |\phi_i\rangle$, one needs to prove that this choice still allows \mathbf{H} to be at least unitarily equivalent to a graph Laplacian Δ , e.g. by explicitly constructing a unitary similarity transform \mathbf{W} such that $\mathbf{W}^\dagger \mathbf{H} \mathbf{W} = \Delta \otimes \mathbb{1}$. Most if not all QMA-hard constructions since Kitaev's go along this route; however, in the language of graphs, the unitary equivalence could only be proven if Δ is the Laplacian of a path graph.⁸ Just as in eq. (2.2), this graph essentially corresponds to the

⁷Hamiltonians such as in eq. (2.2) are combined with a series of local projectors which single out a computationally valid ground state, so strictly speaking the local rules will only have to discriminate the current computational state locally within this valid subspace—cf. [HNN13], where this is exploited to break down 4-local interactions to 2-local ones.

⁸A more complex construction with a local clock was considered in [BT14a], where the authors consider a 2D surface and allow executing transitions in parallel, as long as the execution front behaves in a time-like fashion. To analyse the spectrum of the resulting Hamiltonian, they relate the propagation terms to the diffusion of a string on a torus, corresponding to a ferromagnetic Heisenberg model with partially twisted periodic boundary conditions. Their analysis, while elegant, is specific to their string diffusion-type execution order of quantum gates. These Hamiltonians cannot generally be translationally-invariant, as the circuit

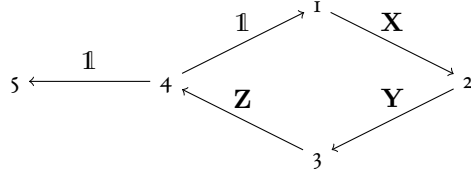


Figure 2.1: Example of a unitary labelled graph (ULG) with vertices $\{1, 2, 3, 4, 5\}$ and three non-trivial unitaries $\mathbf{X}, \mathbf{Y}, \mathbf{Z} \in U(\mathcal{H})$. The associated Hamiltonian for this ULG, as defined in eq. (2.3), is unitarily equivalent to the Laplacian of the underlying graph if $\mathbf{Z}\mathbf{Y}\mathbf{X} = \mathbb{1}$. We provide an explicit description of this change-of-basis unitary.

finite state automaton “driving” the computation; if it is a path graph, the computational path is limited to a sequential application of transition rules $|\psi_i\rangle \mapsto |\phi_i\rangle$ or gate applications encoded therein.

We extend this notion to allow much more complicated branching in the computational path to occur. In particular, we prove a series of results which guarantee the existence of the partially diagonalising unitary \mathbf{W} solely based on properties of the rewriting rules, without the need to explicitly analyse the overall evolution of the system. This has two major benefits: it allows more powerful state transitions which are not necessarily unique for every step, and it drastically simplifies the spectral analysis of \mathbf{H} for whichever construction we choose to work with, as we do not need to construct the equivalence between \mathbf{H} and Δ explicitly. As an important example, our model is the first to allow multiple threads of computation to run in parallel, which then join at some common state.

In a bottom-up approach, we formalise the notion of a Hamiltonian associated with a graph. Starting from a simple directed graph $G = (V, E)$ we associate a Hilbert space \mathcal{H} to each vertex $v \in V$, and a unitary $\mathbf{U}_{(a,b)} : \mathcal{H} \rightarrow \mathcal{H}$ for every directed edge $(a, b) \in E$. We call such a graph with Hilbert space and family of unitaries a *unitary labelled graph*, or ULG for short. As an example, consider fig. 2.1.

The *associated Hamiltonian* for the ULG is then defined as

$$\mathbf{H}(G) := \sum_{(a,b) \in E} \sum_i (|a\rangle \otimes |i\rangle - |b\rangle \otimes \mathbf{U}_{(a,b)} |i\rangle) (\text{ herm. conj. }), \quad (2.3)$$

where the $|i\rangle$ label a basis of \mathcal{H} . Observe that this construction is more general than a local Hamiltonian on a spin chain as in eq. (2.2): $\mathbf{H}(G)$ is simply a Hermitian operator on the overall Hilbert space $\mathbb{C}^V \otimes \mathcal{H}$ where the vertex labels are completely arbitrary, and not necessarily make $\mathbf{H}(G)$ local in any sense.

The associated Hamiltonian $\mathbf{H}(G)$ bears some structural resemblance with a graph Laplacian, as already mentioned. We prove the following theorem.

Theorem 2.3. *If the product of unitaries along any loop in the graph G is $\mathbb{1}$, a property we call simple, then $\mathbf{H}(G)$ is unitarily equivalent to $\Delta \otimes \mathbb{1}_n$, where $n = \dim \mathcal{H}$ and Δ is the Laplacian of G .*

must be laid out on the 2D surface.

Fig. 2.1 satisfies this theorem if and only if the product of unitaries in the loop are $\mathbf{ZYX} = \mathbb{1}$. We provide an explicit expression for this diagonalising unitary, which can be constructed in poly time using a breadth-first search algorithm along a spanning tree of G .

2.1.3.2 Quantum String Rewriting

In order to reintroduce locality to our Hamiltonian construction, we further develop a notation which facilitates embedding transition rules as in eqs. (2.2) and (2.3) into the ground state of a local Hamiltonian. This notation is heavily motivated by string rewriting models, and we extend this notion to introduce a new quantum Turing-complete model based on transitions able to perform quantum gates on part of the string's alphabet.

As mentioned, past hardness constructions (summarised in table 2.1) encode computation in *local* transition rules that act on spins connected by some underlying graph of interactions. While some of these transitions are classical—i.e. basis-preserving—others act on the spin states with a non-diagonal unitary operator, performing the actual quantum computation. Inspired by classical string rewriting systems, we interpret these quantum interactions as local quantum rewriting rules, and introduce a new abstract rewriting system called *quantum Thue system*. This extends an already-existing model of string rewriting—semi-Thue systems⁹—which are well-studied classically [Thoro].

A (classical) semi-Thue system consists of a finite alphabet Σ and length-preserving replacement rules for strings over this alphabet. Similar to the word problem, computation can be encoded in the question whether there exists a connecting path between some input and output strings s_i and s_f . It is straightforward to simulate universal classical Turing machines with a Thue system, which shows that the latter is a Turing-complete model for classical computation. But what about quantum computation?

For quantum Thue systems, we require that the alphabet splits into a classical and a quantum part, i.e. $\Sigma = \Sigma_{cl} \sqcup \Sigma_q$. Transition rules can be purely classical—between elements of Σ_{cl}^* , quantum—between elements of Σ_q^* , or a mixture thereof, in which case we require that the rule preserves the number of quantum symbols $|s|_q$ of a string $s \in \Sigma^*$. In addition, every rule r acting on at least some quantum symbols has a unitary $\mathbf{U}_r \in U(\mathcal{H}^{\otimes |s|_q})$ attached, where \mathcal{H} is some fixed, finite-dimensional Hilbert space.

Starting on some string s and a state vector $|v\rangle \in \mathcal{H}^{\otimes |s|_q}$, we apply any matching string rewriting rule $s \xrightarrow{r} s'$ in turn. For every replacement, we *also* apply the corresponding unitary to the state vector, i.e. $\mathbf{U}_r |v\rangle = |v'\rangle$. In this fashion, we can model quantum computation, driven by a finite automaton: if we make the underlying classical Thue system implement a Turing machine that writes out a quantum circuit description on the string, and then perform this quantum circuit on a separate set of qubits attached to some quantum symbols, the final state vector will contain the output of a quantum computation.

⁹Named after the Norwegian mathematician Axel Thue. We require all rule sets for quantum string rewriting to be symmetric; a symmetric semi-Thue is simply called Thue system, explaining the name quantum Thue system.

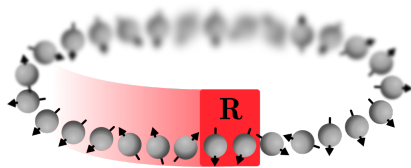


Figure 2.2: Schematic of a quantum ring machine. A fixed unitary \mathbf{R} is cyclically applied to a ring of qudits until one of the qudits indicates a halting configuration.

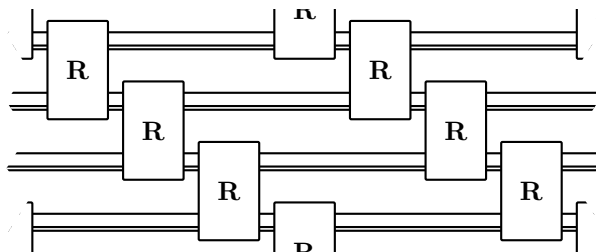


Figure 2.3: Ring machine's evolution implementing a uniform quantum circuit. Double lines carry classical while single lines carry quantum information. Classical wires encode where the next quantum gate from a small universal set will be applied.

One can then show that a quantum Thue system is itself a special case of a unitary labelled graph, which allows us to translate it into a Hamiltonian. We show that the locality of the resulting Hamiltonian only depends on the range of the largest replacement rule, e.g. if one at most replaces a 3-character string, the resulting Hamiltonian will also be 3-local and translationally-invariant.

As replacement rules are not necessarily unique, the computation will have potential ambiguities. As such, we consider all strings connected to the initial starting string s_i via some arbitrary combination of rules, and the size of this set corresponds to the number of basis states that the corresponding *history state* (the ground state of the associated Hamiltonian of the unitary labelled graph defined by the quantum Thue system) is comprised of.

2.1.3.3 A Simpler Computational Model

The complexity class QMA_{EXP} is usually defined in terms of the circuit model, i.e. as a uniform family of verifier circuits: a promise problem $\Pi = (\Pi_{\text{YES}}, \Pi_{\text{NO}})$ is in QMA_{EXP} if there exists a classical Turing machine, such that the verifier circuit for a problem instance $l \in \Pi_{\text{YES}} \cup \Pi_{\text{NO}}$ can be written out by the Turing machine in $O(\exp |l|)$ steps where $|l|$ is the instance size. Being used as an all-purpose computational model, Turing machines have significant downsides: they have complicated transition functions, need a lot of internal states (which translates to an enormous local dimension when encoded in a Hamiltonian) and are rarely written out explicitly (so it is hard to get tight bounds on the required dimension). On the other hand, in past constructions, embedding a circuit directly required the use of non-local clock states marking the position within the circuit, or non-translationally-invariant terms that encode the circuit unambiguously.

We introduce a new computational model which allows us to circumvent the direct use of complicated Turing machines or quantum circuits. The so-called *quantum ring machine* consists of a cyclic ring of *qudits* (i.e. d -dimensional quantum systems) and a unitary \mathbf{R} describing a head that acts on two qudits at a time.

At each time-step, the head moves in the same direction along the ring and cyclically acts on adjacent cells. We give the following definition (see section 2.2.2 for more details).

Definition 2.4 (Quantum ring machine). *A quantum ring machine consists of a ring of n qudits, each of dimension d , and a unitary operator \mathbf{R} acting on a pair of qudits. The n -qudit ring is initialised in state $|\psi_{\text{in}}\rangle$ and the machine proceeds by applying \mathbf{R} cyclically to pairs of adjacent qudits along the ring—see figure 2.3—until one of the qudits indicates halting: its reduced density matrix has support completely inside a certain halting subspace $\mathcal{H}_{\text{halt}}$, while the reduced states of all qudits up to this point were orthogonal to $\mathcal{H}_{\text{halt}}$.*

To show that a quantum ring machine is computationally equivalent to a uniform family of quantum circuits, we encode a classical Turing machine’s transition function into \mathbf{R} , where the internal states, including the Turing machine’s halting flag, are stored as a classical information on the ring. Such ring machine can be used to write out and execute a quantum circuit “on-the-go”: it is universal for whichever uniform circuit class is encompassed by its allowed runtime. Quantum ring machines thus bridge the gap between circuits, which are particularly simple to specify locally but have a complex global structure, and Turing machines, which are difficult to specify locally due to a possibly large number of internal states, but have a straightforward global evolution as the tape only changes in at most one location at each step. A schematic of the ring machine can be found in figs. 2.2 and 2.3.

The ring machine’s simple mechanism allows its evolution to be described by a set of local quantum rewriting rules. These rules operate at a *physical* level while the ring machine operates at a *logical* level—each application of ring machine’s head \mathbf{R} on a pair of logical qubits is implemented by a sequence of physical operations acting on a much larger number of qubits. At any given time the ring machine’s head is positioned on a specific pair of logical qubits, and after each application of \mathbf{R} this location is updated in a similar fashion as Turing machine’s head—it is shifted either up or down along the ring by one position. Overall, \mathbf{R} is a large controlled unitary that acts at a given logical location only if the ring machine’s internal state—stored as a classical bit on the physical tape—is in an active configuration.

2.1.3.4 QMA_{EXP} hardness of 2-TILH

The final proof of theorem 2.2 is based on the following lemma.

Lemma 2.5. *There exists a BQEXP-universal quantum Thue system with 39 symbols, 3 of which are quantum, with attached Hilbert space \mathbb{C}^2 and 2-local rules.*

We prove this by writing out a quantum Thue system which executes a BQEXP-universal quantum ring machine. The quantum Thue system makes heavy use of the new possibilities of ambiguous replacement rules, which allow the history state path to branch. For the QMA_{EXP} hardness proof itself we combine this

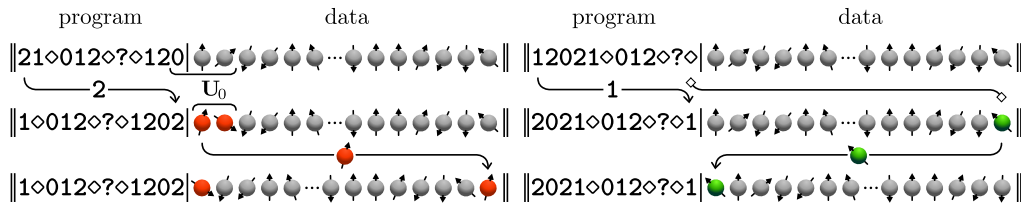


Figure 2.4: Illustration of *Turing's wheelbarrow* construction (see section 2.2.6 for more details). It consists of a tape that stores a program string on the left- and data qubits on the right-hand side. Two types of actions are supported: application of a quantum gate (left figure) and rewinding of the tape (right figure). The rightmost program bit always indicates the next action. For example, 0 indicates that a unitary gate U_0 should be applied to the two leftmost data qubits, and the ring of qubits should then be cyclically rotated one position to the left. On the right, the action of the special symbol \diamond is depicted: it signals the rightmost qubit to move back to the left end of the tape. After each action the program string is cyclically rotated one position to the right.

This system with a series of local penalty terms, which allow us to single out the history state as lowest-energy ground state for any encoded YES instance.

Furthermore, we prove that the quantum This system has a simple history state in the sense of theorem 2.3, which allows us to analyse the spectrum of the resulting Hamiltonian. More specifically, we prove a variant of *Kitaev's geometrical lemma* (see lemmas 2.30 and 2.44) which facilitates the spectral analysis of Hamiltonians that are sums of a unitary labelled graph Hamiltonian and local projectors. This finally allows us to prove our main result, theorem 2.2, that 2-LOCAL-HAMILTONIAN is QMA_{EXP} -hard, even for translationally-invariant nearest-neighbour interactions between spins of local dimension 42.

For completeness, we also want to give a brief overview over the family of hard QTS instances that we construct, but—as mentioned before—the QMA_{EXP} hardness proof does not depend on the precise workings of it; assuming that lemma 2.5 can be proven, theorem 2.2 stands independently.

Treating the Hilbert space of the 2-TLH problem as a physical tape of length n —some symbols quantum, some classical—we write a set of transition rules to perform the following steps that simulate the quantum ring machine.

1. As in the construction by Gottesman and Irani [GI13], we use a counter to translate the chain length n into a *program* string of length $O(\log n)$ on the left hand side of the chain, while on the right hand side we store the physical data qubits, i.e. the ring of qubits our ring machine is executed on.
2. The program on the left hand side contains a physical-level description of a quantum circuit (over a small, finite, universal gate set) for implementing one step of the quantum ring machine, i.e. one application of the ring machine's head \mathbf{R} . The program's rightmost bit always indicates the next gate in the circuit, and this gate is always applied to the two leftmost data qubits on the physical ring (see fig. 2.4).

3. Using the two types of basic commands—“apply gate” and “rewind tape”—shown in fig. 2.4, the quantum circuit implementing \mathbf{R} can be executed cyclically on the physical data qubits, some of which are initialised to ancillary $|0\rangle$'s to be used in the computation.
4. The computation runs until a certain internal classical counter (stored on the ring) terminates. In our construction, we explicitly encode transitions for the gates SWAP , TOFFOLI and a controlled quantum-universal unitary; since SWAP and TOFFOLI are also universal for classical computation, the classical control machinery in the ring machine's head \mathbf{R} (i.e. the Turing machine used to write out the quantum circuit) can be executed exactly (without error). This means that the computation will halt deterministically (as otherwise there could be some overlap with a non-halting state). The transition rules for applying a gate as in fig. 2.4 then have another control gate which only proceeds if the data bit to its right is in a specific configuration, terminating the machine's execution otherwise.
5. The length of the chain is chosen so that the program encodes a quantum ring machine equivalent to a BQEXP verifier circuit. It discriminates between YES and NO instances of the corresponding QMA_{EXP} language depending on whether the ring machine accepts or rejects, and a special symbol in the program description allows us to locally penalise a wrong initialisation of ancillas and a NO output of the computation.

Our construction is universal in the sense that it can be used to implement an arbitrary quantum computation without the need to increase the local dimension (in the same spirit as a universal Turing machine can implement any computation without the need to increase the number of internal states). Since we leave parts of the input unconstrained, we conclude from BQEXP -completeness of these instances that they can be used as a QMA_{EXP} verifier, finalising our claims.

2.1.4 Structure of the Chapter

We summarise several standard definitions in section 2.2.1. In section 2.2.2, we define the aforementioned quantum ring machine and show that it is indeed Turing-complete for quantum computation. Section 2.2.4 formalises the notion of quantum replacement rules and introduces the model of quantum Thue systems. Section 2.2.6 contains a constructive proof of a universal quantum Thue system, and section 2.2.5 combines everything into our main hardness result.

2.2 Turing's Wheelbarrow

2.2.1 Preliminaries

2.2.1.1 Reversible Turing Machines

We give the following standard definition of a (non-deterministic) Turing machine (for more background on Turing machines, see chapter 8 of [ED79]).

Definition 2.6 (Turing machine). *A Turing machine—or TM for short—is a triple (Q, Σ, δ) , where Q is a finite set of internal states containing a distinct initial and halting state q_0 and q_f , respectively, and Σ is a finite set of tape symbols containing a designated blank symbol 0. Let $D := \{\text{left}, \text{right}\}$ be the two possible movement directions of the TM's head. Then each element of the transition set $\delta \subseteq Q \times \Sigma \times \Sigma \times D \times Q$ is a quintuple of the form (q, s, s', d, q') , which means that if the Turing machine reads a symbol s under its head while in state q , it overwrites the symbol by s' , moves the head in direction $d \in D$ and transitions to state q' . At the beginning of the computation, the TM's initial state is q_0 and the tape is initialised to all 0s, except for a finite block of consecutive cells containing the input. The machine halts once its internal state is q_f , for which there is no forward transition.*

As we aim to implement TMs using quantum mechanics, we need them to be deterministic and reversible. The following is based on definition 10 from [Moro8].

Definition 2.7 (Deterministic and reversible Turing machine). *Consider a Turing machine (Q, Σ, δ) , and let $(q_1, s_1, s'_1, d_1, q'_1)$ and $(q_2, s_2, s'_2, d_2, q'_2)$ be any two distinct quintuples in δ . This TM is*

- deterministic if $(q_1 = q_2) \implies (s_1 \neq s_2)$,
- reversible if $(q'_1 = q'_2) \implies (s'_1 \neq s'_2) \wedge (d_1 = d_2)$.

The first condition of definition 2.7 rules out the possibility that $q_1 = q_2$ and $s_1 = s_2$, meaning that the current TM's state and tape symbol should unambiguously determine the rest of the transition. Similarly, the second condition rules out the possibility that $q'_1 = q'_2$ and $s'_1 = s'_2$, as well as the possibility that $q'_1 = q'_2$ and $d_1 \neq d_2$, meaning that the reverse transition also is uniquely determined by the current state and tape symbol, and that the direction of the TM's head movement in reverse is uniquely determined by the current state.

For a deterministic TM, one can regard δ as a partial function, namely $\delta : Q \times \Sigma \rightarrow \Sigma \times D \times Q$, since all combinations of internal state q and tape symbol s have at most one forward transition. For a reversible TM, δ is injective since all combinations of internal state q' and tape symbol s' have at most one backwards transition (whenever such transition exists, it uniquely determines the head movement direction d backwards). In fact, according to definition 2.7, each state of a reversible TM can be entered only from

one direction (this property is referred to as *unidirectionality* in [BV97b]). In other words, it is sufficient to know only the TM’s current state (as opposed to both the state and the tape symbol) to answer the question “From which direction did the TM’s head arrive?”.

Due to unidirectionality, it is often natural to restrict the range of δ to $Q \times \Sigma$. In fact, the transition function δ of a deterministic reversible Turing machine can be replaced by a permutation matrix on $Q \times \Sigma$ without affecting the TM’s behaviour. For our convenience, we state this observation more formally (see also cor. B.2 and thm. 4.2 in [BV97b]).

Lemma 2.8. For any deterministic reversible Turing machine (Q, Σ, δ) , the partial transition function δ can be replaced by a pair (\mathbf{T}_δ, d) , where \mathbf{T}_δ is a permutation matrix on $Q \times \Sigma$ and $d : Q \rightarrow D$ is a function that determines, for each internal state $q \in Q$, the direction from which the TM’s head arrived in q . If we update the TM’s internal state and the current tape symbol according to \mathbf{T}_δ , and then move the TM’s head in the direction opposite to $d(q')$, where q' is the updated state, the behaviour is identical to the original transition function δ .

Proof. The function d is readily obtained because of unidirectionality. A blueprint of \mathbf{T}_δ is obtained by restricting the range of δ to $Q \times \Sigma$ and describing δ ’s action on the elements of this set by a binary matrix. Since the TM is deterministic and reversible, this matrix contains at most one entry 1 in each row and column, so it can be easily extended to a permutation matrix. \square

From now on we will consider only deterministic Turing machines and implicitly assume that they are reversible—this is justified by the following result due to Bennett [Ben73] (see [SP13]; [Moro8] for more background on reversible computation).

Theorem 2.9 (Bennett [Ben73]). Any deterministic TM can be made reversible with at most polynomial overhead in terms of space and time.

2.2.1.2 Quantum Complexity Classes

In this section, we formally define the quantum complexity classes BQP, BQEXP, QMA and QMA_{EXP} in terms of the circuit model, and refer reader to [BV97b]; [Wat12b]; [VW16] for more details on quantum computational complexity.

In what follows, we fix some finite universal set of 2-qubit quantum gates, such as {HADAMARD, CNOT, $R(\pi/4)$ }—see [NC10, ch. 4.5]. We first define a uniform family of quantum circuits over this gate set.

Definition 2.10 (Uniform family of quantum circuits). Let $f : \mathbb{N} \rightarrow \mathbb{N}$ be a function and $(C_n)_{n \in \mathbb{N}}$ be a family of quantum circuits where each C_n

- acts on n qubits and has a distinct output qubit,
- requires at most $f(n)$ additional ancilla qubits initialised in $|0\rangle$,
- contains at most $f(n)$ gates from our universal set.

We say that $(C_n)_{n \in \mathbb{N}}$ is $f(n)$ -uniform if there exists a TM that on input 1^n produces an explicit description of C_n in less than $f(n)$ steps.

Let Σ be a finite set (alphabet), and let Σ^n and $\Sigma^* := \bigcup_{n \geq 0} \Sigma^n$ denote the sets of all length- n and all finite-length strings over Σ , respectively. A *promise problem* over alphabet Σ is a pair $\Pi = (\Pi_{\text{YES}}, \Pi_{\text{NO}})$ such that $\Pi_{\text{YES}} \cap \Pi_{\text{NO}} = \emptyset$, where $\Pi_{\text{YES}}, \Pi_{\text{NO}} \subseteq \Sigma^*$ are the sets of input strings corresponding to YES and NO instances, respectively. We will sometimes write $l \in \Pi$ meaning that $l \in \Pi_{\text{YES}} \cup \Pi_{\text{NO}}$.

Definition 2.11 (Complexity class $\text{BQ}(f)$). A *promise problem* $\Pi = (\Pi_{\text{YES}}, \Pi_{\text{NO}})$ is in $\text{BQ}(f)$, bounded-error quantum $f(n)$ -time, if there exists an $f(n)$ -uniform family of quantum circuits $(C_n)_{n \in \mathbb{N}}$ such that

$$\Pr(C_n(s) = \text{YES}) \geq \frac{2}{3} \quad \text{for } s \in \Pi_{\text{YES}} \quad \text{and} \quad \Pr(C_n(s) = \text{YES}) \leq \frac{1}{3} \quad \text{for } s \in \Pi_{\text{NO}},$$

where $C_n(s)$ denotes the random variable obtained by executing C_n on input $s \in \Pi$ of size $|s| = n$ and measuring the output qubit (the encoding of s as well as the measurement are performed in the computational basis).

We introduced the notation $\text{BQ}(f)$ to emphasise the fact that the definitions of classes BQP and BQEXP are essentially the same up to the bounding function:

$$\text{BQP} := \bigcup_{k \in \mathbb{N}} \text{BQ}(n^k) \quad \text{and} \quad \text{BQEXP} := \bigcup_{k \in \mathbb{N}} \text{BQ}(\exp(n^k)).$$

Trivially, $\text{BQP} \subseteq \text{BQEXP}$ since a longer runtime can only help.

It is well-known (see [Wat12b, Prop. 3]) that for BQP the probabilities of $2/3$ and $1/3$ in definition 2.11 can be exponentially amplified while still remaining in the same complexity class. The same argument works for BQEXP as well, since we only need a polynomial number of repetitions to achieve the desired amplification.

Fact 2.12 (Error-reduction for BQP and BQEXP). For any polynomial p , we can assume that $\Pr(C_n(s) = \text{YES}) \geq 1 - 2^{-p(n)}$ for $s \in \Pi_{\text{YES}}$ and $\Pr(C_n(s) = \text{YES}) \leq 2^{-p(n)}$ for $s \in \Pi_{\text{NO}}$ in the definitions of BQP and BQEXP.

Intuitively, $\text{QMA}(f)$ is the class of promise problems for which the YES/NO answers can be verified by a $\text{BQ}(f)$ verifier.

Definition 2.13 (Complexity class $\text{QMA}(f)$). A promise problem $\Pi = (\Pi_{\text{YES}}, \Pi_{\text{NO}})$ is in $\text{QMA}(f)$, $f(n)$ -time quantum Merlin-Arthur, if there exists an $f(n)$ -uniform family of verifier quantum circuits⁴⁰ $(C_n)_{n \in \mathbb{N}}$ such that

- if $s \in \Pi_{\text{YES}}$, \exists a witness state ρ on at most $f(n)$ qubits such that $C_n(s, \rho) = \text{YES}$ with probability at least $2/3$. This condition is known as completeness.
- if $s \in \Pi_{\text{NO}}$, \forall witness states ρ on at most $f(n)$ qubits $C_n(s, \rho) = \text{YES}$ with probability at most $1/3$. This condition is called soundness.

Observe that the witness size is implicitly constrained by the size of the quantum circuit family, cf. definition 2.10, e.g. for BQP verifiers the witness is poly-sized while for BQEXP verifiers it can be exp-sized. As before, we define

$$\text{QMA} := \bigcup_{k \in \mathbb{N}} \text{QMA}(n^k) \quad \text{and} \quad \text{QMA}_{\text{EXP}} := \bigcup_{k \in \mathbb{N}} \text{QMA}(\exp(n^k)).$$

In particular, note that $\text{QMA} \subseteq \text{QMA}_{\text{EXP}}$ since a QMA verifier can be easily promoted to a QMA_{EXP} verifier. Indeed, while a QMA_{EXP} verifier gets an exponential-size witness and can run for an exponential amount of time, it does not have to (it can instead discard all witness qubits, except for a polynomial number, and verify them in polynomial time).

Soundness and completeness probabilities for QMA and QMA_{EXP} can also be amplified: see theorem 10 in [Wat12b], section 3.2 of [VW16], or lemma 14.1 in [KSV02] (these techniques were originally devised for QMA , but they can be easily adapted also for QMA_{EXP}).

2.2.1.3 Geometrically k -Local Hamiltonians

In this section we introduce basic notions relating to local Hamiltonians and formally state the TLH problem that will play central role. For more background on Hamiltonian complexity, see [Gha+14]; [YSN02]; [Wat12b].

Definition 2.14. An n -qudit Hamiltonian is a Hermitian operator $\mathbf{H} = \mathbf{H}^\dagger$ acting on a multipartite Hilbert space $(\mathbb{C}^d)^{\otimes n}$ consisting of n systems (qudits), each of local dimension d .

We will label the individual systems by elements of $S := \{1, \dots, n\}$. Whenever we talk of a subset of systems $A \subseteq S$, we mean an ordered tuple of distinct elements of S . If \mathbf{h} is a k -qudit Hamiltonian for some $k \leq n$ and $A \subseteq S$ is a subset of $|A| = k$ systems, we write \mathbf{h}_A to denote the n -qudit Hamiltonian that acts as \mathbf{h} on qudits A and trivially (i.e. as $\mathbb{1}$) on the remaining qudits $S \setminus A$. We also write $A + i \subseteq S$ to denote A shifted by $i \in \mathbb{N}$ positions.

⁴⁰ Here we use a slight variation of definition 2.10: we also allow for at most $f(n)$ extra input qubits to store the witness state ρ (this is in addition to the n original input qubits and $f(n)$ ancillary qubits that are initialised in $|0\rangle$).

Definition 2.15. Let \mathbf{H} be an n -qudit Hamiltonian. Then

- \mathbf{H} is k -local if $\mathbf{H} = \sum_i \mathbf{h}(i)_{A_i}$ with $|A_i| \leq k \forall i$;
- \mathbf{H} is k -local and 1D if each $A_i \subseteq \{1, \dots, k\} + t_i$ for some shift t_i ;
- \mathbf{H} is translationally-invariant if $\mathbf{H} = \sum_i \mathbf{h}_{A+i}$ for some $A \subseteq S$ where \mathbf{h} is fixed.

In particular, \mathbf{H} is a 1D translationally-invariant k -local Hamiltonian if $\mathbf{H} = \sum_i \mathbf{h}_{\{1, \dots, k\}+i}$ for some fixed k -qudit Hamiltonian \mathbf{h} .

Our central problem of interest is deciding the ground energy of 1D translationally-invariant k -local Hamiltonians of local dimension d . For brevity, we will refer to this as the TILH problem, following the convention by [GI13].

Definition 2.16 ((k, d) -TILH). Let $\mathbf{H} = \sum_i \mathbf{h}_{\{1, \dots, k\}+i}$ be a 1D translationally-invariant k -local Hamiltonian on a qudit chain of length n , where each qudit has local dimension d and \mathbf{h} is some fixed k -qudit Hamiltonian.

Input. The chain length n and the matrix entries of \mathbf{h} , as well as two real numbers α and β , all up to $\log n$ bits of precision. •

Promise. The operator norm of each local term is bounded, $\|\mathbf{h}\| \leq 1$, and either $\lambda_{\min}(\mathbf{H}) \leq \alpha$ or $\lambda_{\min}(\mathbf{H}) \geq \beta$, where $\lambda_{\min}(\mathbf{H})$ denotes the smallest eigenvalue of \mathbf{H} and $\beta - \alpha \geq 1/p(n)$ for some fixed polynomial $p(n)$.

Output. YES if $\lambda_{\min}(\mathbf{H}) \leq \alpha$, else NO.

We emphasise that the input in definition 2.16 is just the description of the k -local term \mathbf{h} and the chain length n , not the entire (exponentially-sized) Hamiltonian \mathbf{H} . An equivalent variant of the definition relaxes the norm bound to $\|\mathbf{h}\| \leq \text{poly } n$ and gives a promise that either $\lambda_{\min}(\mathbf{H}) \leq \alpha$ or $\lambda_{\min}(\mathbf{H}) \geq \beta$ for some fixed constants $\beta > \alpha$. We can always rescale the overall Hamiltonian by a polynomial factor to switch between the two definitions.

Theorem 2.17 (Kitaev [YSN02]). (k, d) -TILH is in QMA_{EXP} .

Proof. This does not trivially follow from the inclusion $\text{QMA} \subseteq \text{QMA}_{\text{EXP}}$ since the input size for TILH is just $\text{poly } \log n$. However, Kitaev's QMA verifier for the standard LOCAL HAMILTONIAN problem runs in time $\text{poly } n$, which is not polynomial in the input size for TILH. However, the exponential-time verifier of QMA_{EXP} offsets the logarithmically small input size, so the same random sampling argument as presented for QMA in e.g. [YSN02, prop 14.2] goes through. □

2.2.1.4 QMA versus QMA_{EXP}

In this section, we clarify why QMA_{EXP} is the natural class when considering the LOCAL HAMILTONIAN problem with translationally-invariant interactions on a system of size n . When specifying a k -local Hamiltonian $\mathbf{H} = \sum_{i \in I} \mathbf{h}_i$, for some set of interactions I with $|I| = \text{poly } n$, we have to specify each term \mathbf{h}_i individually. Since the locality k and the local dimension d are constant, the total input size in definition 2.1 is thus $l = \text{poly } n$ bits. In contrast, specifying a *translationally invariant* Hamiltonian \mathbf{H} requires only a logarithmic number of bits: since all local terms \mathbf{h}_i are identical and do not vary with the system size n , the only part of the input that varies with n and can thus be used to encode different instances of the problem is the system size n itself.

A fact which we will discuss in great detail in section 2.2.4 is that the gap of a Hamiltonian encoding computation as a superposition of basis states—a so-called *history state construction*—scales inversely polynomially in the runtime, i.e. $1/\text{poly}(f(l))$ for an input of size l and an $f(l)$ -time computation. Contrasting this with the $1/\text{poly } n$ gap required by definitions 2.1 and 2.16 independently—inverse polynomially in the system size, *not* the input size—we conclude the following core differences between QMA and QMA_{EXP} in the context of the LOCAL HAMILTONIAN problem (recall that n denotes the length of the spin chain and l denotes the total size of the input).

QMA. A BQP verifier has $\text{poly } l$ runtime on an input of size l , so the gap of the Hamiltonian that encodes the verifier scales as $1/\text{poly } l$. This agrees with $1/\text{poly } n$ in definition 2.1 since l and n are poly -related. QMA is thus the natural class for the LOCAL HAMILTONIAN problem.

QMA_{EXP}. The BQEXP verifier can run for $\text{exp poly } l$ steps in the input size l . The gap therefore scales as $1/\text{exp poly } l$, which agrees with $1/\text{poly } n$ in definition 2.16 since $l = \text{poly log } n$. QMA_{EXP} is thus the natural class for TILH.

One fact we have glossed over is that even though each instance of TILH is translationally invariant, we could still vary the local interaction for each system size n . As an example, assume that the Hamiltonian \mathbf{H} is specified by a single local term,

$$\mathbf{H} := \sum_{i=1}^{N-1} \mathbf{h}_{i,i+1} \quad \text{where} \quad \mathbf{h} = \begin{pmatrix} 1 & 0 \\ 0 & \alpha(m) \end{pmatrix} \quad \text{with} \quad \alpha(m) = \underbrace{3.1415926 \dots 42}_{m \text{ digits of } \pi}.$$

Then the bit complexity of this input is $O(m)$, and the overall input size—i.e. the possible information specifiable using the two parameters, the system size n and a varying parameter m , is thus $O(m + \log n)$. In order not to overspecify a LOCAL HAMILTONIAN or TILH problem, in each case we have to require

both bit precision and size of the input parameter to be of the same order (within polynomial factors). We conclude with the following remark.

Remark 2.18. It is natural to allow poly n precision of the entries in the local terms of the Hamiltonian when working with QMA, whereas for QMA_{EXP} local terms need to be precision-limited by poly log n .

However, we want to emphasise that *we will only make use of uniformly scaling local interaction terms*, as in [GI13]: this in particular allows us to use coupling constants that scale polynomially in n . We also want to note that the polynomially-closing promise gap of history state constructions might not be the end of the story; at this point in time it is not known whether quantum computation can be encoded into the ground state of a local Hamiltonian for which the promise gap e.g. scales in a sub-linear fashion in the number of computational steps (cf. chapter 5 for an extended discussion).

In [CPW15a], the authors use a phase-estimation algorithm to extract $O(n)$ bits of information from a fixed Hamiltonian term. However, in their construction, the speed at which the gap closes is irrelevant, as long as it remains nonzero in the gapped phase.

With a poly(n)-bounded computation and a $1/\text{poly}(n)$ gap, however, it is not clear how to do this computation in a translationally-invariant manner. For phase estimation of m bits, one requires gates of precision $O(\exp(-m))$, cf. [NC10]—the algorithm depends on being able to perform a unitary \mathbf{U} an exponential number of times, i.e. $\mathbf{U}, \mathbf{U}^2, \mathbf{U}^4, \dots, \mathbf{U}^{2^{m-1}}$. Without having direct access to all powers of this gate—which we do not, if we require bounded local dimension and locality—we need to approximate them in some way: using the Solovay-Kitaev theorem with the required exponential precision $O(\exp(-m))$ results in a circuit of size $\text{poly log } 1/(\exp(-m)) = \text{poly } m$, which limits the amount of information we can extract to $m = O(\log n)$.

It is clear that this is a problem of bootstrapping. For TILH, we only have $O(\log n)$ information available to start the computation with, and again it is not known whether there exists a more direct way of extracting a phase without having to go through the Solovay-Kitaev theorem, which only gives a sufficient upper bound to approximate the phase estimation algorithm.

2.2.1.5 Laplacian Matrix and Algebraic Connectivity of Graphs

In this section we revise general notation and basic results from graph theory. For more background, consult the standard references [Tru13]; [Dier0] and [GR01] on graph theory and algebraic graph theory, respectively.

Definition 2.19. An undirected simple graph $G = (V, E)$ consists of a set of vertices V and a set of edges E , each edge being an unordered pair of distinct elements of V (in particular, there are no self-loops and no multiple edges). If the number of vertices is $n = |V|$ and we label them as $V = \{v_1, \dots, v_n\}$, then the

adjacency matrix of G is $\mathbf{A}(G) := (a_{ij})_{1 \leq i, j \leq n}$ where

$$a_{ij} := \begin{cases} 1 & \{v_i, v_j\} \in E, \\ 0 & \text{otherwise.} \end{cases}$$

We further define the degree matrix $\mathbf{D}(G) := \text{diag}((\deg v_i)_{1 \leq i \leq n})$ where $\deg v_i := \sum_{j=1}^n a_{ij}$.

We will usually omit the qualifiers “undirected” and “simple” in the rest of this thesis. We proceed to introduce basic notions and facts from algebraic graph theory [GR01].

Definition 2.20 (Laplacian matrix). *The Laplacian matrix of a graph G is defined as $\Delta(G) := \mathbf{D}(G) - \mathbf{A}(G)$.*

Since $\mathbf{A}(G)$ and $\Delta(G)$ are linear operators on \mathbb{C}^n where $n = |V|$, it will often be convenient to label the basis vectors of this space by $|v\rangle$ where $v \in V$ and denote the space itself by \mathbb{C}^V .

Definition 2.21. *We write $\lambda_{\min}(\mathbf{M})$ to denote the smallest eigenvalue of Hermitian operator \mathbf{M} . If $\mathbf{M} \geq 0$ then $\lambda_{\min}(\mathbf{M}|_{\text{supp } \mathbf{M}})$ denotes the smallest non-zero eigenvalue of \mathbf{M} .*

Claim 2.22. *For any graph G , $\Delta(G)$ is real symmetric. In fact, $\Delta(G)$ is positive semi-definite with smallest eigenvalue $\lambda_{\min}(\Delta(G)) = 0$ and corresponding eigenvector $(1, \dots, 1)$.*

Proof. By construction, $\mathbf{A}(G)$ and $\mathbf{D}(G)$ are real symmetric and so is $\Delta(G)$. The second claim follows by observing that $\Delta(G)$ is symmetric and diagonally dominant. Alternatively, $\Delta(G)$ can be expressed as a sum of positive semi-definite matrices:

$$\Delta(G) = \sum_{\{a,b\} \in E} (|a\rangle - |b\rangle)(\langle a| - \langle b|), \quad (2.4)$$

where each term is a principal submatrix of the form

$$\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$$

and encodes the Laplacian of a single edge. The last statement follows from the fact that the row sums of $\Delta(G)$ are zero. \square

Definition 2.23 (Algebraic connectivity). *The second smallest eigenvalue of the Laplacian $\Delta(G)$ is denoted with $\alpha(G)$ and called the algebraic connectivity of graph G . The corresponding eigenvector is known as the Fiedler vector.*

Claim 2.24 (Fiedler [Fie73]). For any graph G , $a(G) > 0$ if and only if G is connected.

Lemma 2.25. If $G = G_1 \sqcup \dots \sqcup G_m$ is a disjoint union of connected components G_i then $\Delta(G)$ has eigenvalue 0 with multiplicity exactly m and the next smallest eigenvalue is $\lambda_{\min}(\Delta(G)|_{\text{supp } \Delta(G)}) = \min_i a(G_i)$. Furthermore, $\{|\Phi_1\rangle, \dots, |\Phi_m\rangle\}$ with

$$|\Phi_i\rangle := \frac{1}{\sqrt{|V_i|}} \sum_{v \in V_i} |v\rangle$$

is an orthonormal basis of the 0-eigenspace (ground space) of Δ .

Proof. Note that $\Delta = \Delta_1 \oplus \dots \oplus \Delta_m$ where Δ_i is the Laplacian of G_i . Recall from claim 2.22 that $\Delta_i \geq 0$ and $\lambda_{\min}(\Delta_i) = 0$, hence the m smallest eigenvalues of Δ are equal to 0. Since each G_i is connected, $a(G_i) > 0$ for every i by claim 2.24. Hence the multiplicity of eigenvalue 0 must be m and the $(m+1)$ -st smallest eigenvalue of Δ is positive and equal to $a(G_i)$ for some i . Finally, recall from claim 2.22 that the uniform superposition over all vertices V_i of G_i is a 0-eigenvector of Δ_i , thus $\Delta |\Phi_i\rangle = 0$ for each $i \in \{1, \dots, m\}$. There are no further vectors in the ground space of Δ since Δ has eigenvalue 0 with multiplicity m . \square

Corollary 2.26. If Δ is the Laplacian of graph $G = (V, E)$ and $U \subseteq V$ is some connected component of G , then $|\Phi_U\rangle := \sum_{v \in U} |v\rangle / \sqrt{|U|}$ is a 0-eigenvector of Δ . In fact, any 0-eigenvector of Δ is a linear combination of such vectors.

Claim 2.27 (Fiedler [Fie73]). Let G_L be the path graph on L vertices:

$$G_L := \begin{array}{ccccccc} & 1 & 2 & 3 & \dots & L-1 & L \\ & \circ & \circ & \circ & \text{---} & \circ & \circ \\ & & \text{---} & & & \text{---} & \end{array}$$

Then $a(G_L) = 2(1 - \cos(\pi/L)) \sim \pi^2/L^2$. In particular, $a(G_L) = \Theta(1/L^2)$.

Corollary 2.28. Let G be a connected graph with L vertices. Then $a(G) = \Omega(1/L^2)$.

Proof. The algebraic connectivity is non-decreasing under adding edges [Fie73, corollary 3.2], so for any connected graph on L vertices it is lower-bounded by that of a path graph on L vertices, which is given by claim 2.27. \square

2.2.1.6 Kitaev's Geometrical Lemma for Graphs

We will need Kitaev's geometrical lemma (see Lemma 14.4 in [YSNo2]) whose proof is reproduced below.

Lemma 2.29 ([YSNo2], p. 147). Using notation from definition 2.21, assume $\mathbf{A}, \mathbf{B} \geq 0$ are such that $\lambda_{\min}(\mathbf{A}|_{\text{supp } \mathbf{A}}) \geq \mu$ and $\lambda_{\min}(\mathbf{B}|_{\text{supp } \mathbf{B}}) \geq \mu$, and the null spaces of \mathbf{A} and \mathbf{B} have no vector in

common other than 0, i.e. $\ker \mathbf{A} \cap \ker \mathbf{B} = \{0\}$. Then $\lambda_{\min}(\mathbf{A} + \mathbf{B}) \geq 2\mu \sin^2 \frac{\theta}{2}$, where θ is the angle between subspaces $\ker \mathbf{A}$ and $\ker \mathbf{B}$, i.e.

$$\cos \theta := \max_{\substack{|\alpha\rangle \in \ker \mathbf{A} \\ |\beta\rangle \in \ker \mathbf{B}}} |\langle \alpha | \beta \rangle|$$

where $|\alpha\rangle$ and $|\beta\rangle$ are unit vectors.

Proof. We define $\Pi_{\mathbf{A}}$ to be the projector onto $\ker \mathbf{A}$, and analogously for $\Pi_{\mathbf{B}}$. It follows from $\lambda_{\min}(\mathbf{A}|_{\text{supp } \mathbf{A}}) \geq \mu$ that $\mathbf{A} \geq \mu(\mathbb{1} - \Pi_{\mathbf{A}})$ and similarly for \mathbf{B} . It is hence enough to show that $(\mathbb{1} - \Pi_{\mathbf{A}}) + (\mathbb{1} - \Pi_{\mathbf{B}}) \geq (2 \sin^2 \frac{\theta}{2})\mathbb{1}$, which is equivalent to $(1 + \cos \theta)\mathbb{1} \geq \Pi_{\mathbf{A}} + \Pi_{\mathbf{B}}$. In other words, we want to show that every eigenvalue λ of $\Pi_{\mathbf{A}} + \Pi_{\mathbf{B}}$ satisfies

$$1 + \cos \theta \geq \lambda. \quad (2.5)$$

Let $|\psi\rangle$ be a normalised eigenvector of $\Pi_{\mathbf{A}} + \Pi_{\mathbf{B}}$ with eigenvalue $\lambda \geq 0$. Since eq. (2.5) holds trivially for $\lambda = 0$, we can assume $\lambda > 0$. Since $\Pi_{\mathbf{A}}$ projects onto $\ker \mathbf{A}$, we can find a unit vector $|\psi_{\mathbf{A}}\rangle \in \ker \mathbf{A}$ such that $\Pi_{\mathbf{A}} |\psi\rangle = a |\psi_{\mathbf{A}}\rangle$ for some $a \in \mathbb{C}$; we can adjust the global phase of $|\psi_{\mathbf{A}}\rangle$ to guarantee that $a \geq 0$. Similarly, $\Pi_{\mathbf{B}} |\psi\rangle = b |\psi_{\mathbf{B}}\rangle$ for some unit vector $|\psi_{\mathbf{B}}\rangle \in \ker \mathbf{B}$ and $b \geq 0$. Since $\Pi_{\mathbf{A}}$ and $\Pi_{\mathbf{B}}$ are projectors, $\langle \psi | \Pi_{\mathbf{A}} | \psi \rangle = \langle \psi | \Pi_{\mathbf{A}}^\dagger \Pi_{\mathbf{A}} | \psi \rangle = a^2 \langle \psi_{\mathbf{A}} | \psi_{\mathbf{A}} \rangle = a^2$ and $\langle \psi | \Pi_{\mathbf{B}} | \psi \rangle = b^2$. From $\lambda |\psi\rangle = (\Pi_{\mathbf{A}} + \Pi_{\mathbf{B}}) |\psi\rangle$ we get by linearity that

$$\lambda = \langle \psi | (\Pi_{\mathbf{A}} + \Pi_{\mathbf{B}}) | \psi \rangle = a^2 + b^2.$$

Furthermore,

$$\begin{aligned} \lambda^2 &= \langle \psi | (\Pi_{\mathbf{A}} + \Pi_{\mathbf{B}})^2 | \psi \rangle = a^2 + b^2 + 2ab \operatorname{Re} \langle \psi_{\mathbf{A}} | \psi_{\mathbf{B}} \rangle \\ &\leq \lambda + 2ab |\langle \psi_{\mathbf{A}} | \psi_{\mathbf{B}} \rangle| \leq \lambda + (a^2 + b^2) |\langle \psi_{\mathbf{A}} | \psi_{\mathbf{B}} \rangle| = \lambda(1 + |\langle \psi_{\mathbf{A}} | \psi_{\mathbf{B}} \rangle|) \\ &\leq \lambda \left(1 + \max_{\substack{|\alpha\rangle \in \ker \mathbf{A} \\ |\beta\rangle \in \ker \mathbf{B}}} |\langle \alpha | \beta \rangle| \right) = \lambda(1 + \cos \theta), \end{aligned}$$

and hence $\lambda \leq 1 + \cos \theta$, which proves eq. (2.5). \square

We want to use Kitaev's geometrical lemma to lower bound the smallest eigenvalue of a graph Laplacian when certain vertices are penalised. To be more specific, for a graph $G = (V, E)$ and a set of vertices $P \subsetneq V$, we write a *penalising matrix*

$$\mathbf{P}(G, P) := \sum_{v \in P} |v\rangle\langle v|.$$

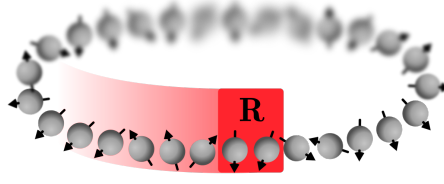


Figure 2.5: Quantum ring machine (QRM). Starting from a ring of qudits $\mathcal{H} = (\mathbb{C}^d)^{\otimes n}$ in an initial configuration $|\psi_{\text{in}}\rangle \in \mathcal{H}$, a unitary $\mathbf{R} \in \text{U}(\mathbb{C}^d \otimes \mathbb{C}^d)$ is applied to pairs of adjacent qudits until one of them is completely in some halting subspace $\mathcal{H}_{\text{halt}} \subseteq \mathbb{C}^d$.

A priori, it is not clear at all what the spectrum of the *penalised Laplacian* $\Delta + \mathbf{P}$ is, however we can obtain a lower bound on the smallest eigenvalue.

Lemma 2.30 (Kitaev’s geometrical lemma for graphs). *Let $G = (V, E)$ be a connected graph. Pick a non-empty subset of penalised vertices $P \subsetneq V$ and write the penalised Laplacian as $\Delta_P(G) := \Delta(G) + \mathbf{P}(G, P)$. Then $\lambda_{\min}(\Delta_P) = \Omega(1/|V|^3)$.*

Proof. Let us first verify that Δ and \mathbf{P} satisfy the prerequisites of lemma 2.29. Since G is connected, $\lambda_{\min}(\Delta|_{\text{supp } \Delta}) = a(G) > 0$ by claim 2.24. Moreover, $\ker \Delta$ is spanned by the all-ones vector $|\Phi_V\rangle := \sum_{v \in V} |v\rangle / \sqrt{|V|}$ according to corollary 2.26. Clearly, $\lambda_{\min}(\mathbf{P}|_{\text{supp } \mathbf{P}}) = 1$ and $|\Phi_V\rangle \notin \ker \mathbf{P}$ since $P \neq \emptyset$, so $\ker \Delta \cap \ker \mathbf{P} = \{0\}$. We can take the constant in lemma 2.29 to be $\mu := \min\{a(G), 1\} = \Omega(1/|V|^2)$, where we used the lower bound $a(G) = \Omega(1/|V|^2)$ from corollary 2.28 on the algebraic connectivity of G . It remains to compute the angle θ between $\ker \Delta = \text{span}\{|\Phi_V\rangle\}$ and $\ker \mathbf{P} = \text{span}\{|v\rangle : v \notin P\}$. We have:

$$\cos \theta = \langle \Phi_V | \left(\frac{1}{\sqrt{|V| - |P|}} \sum_{v \notin P} |v\rangle \right) = \frac{|V| - |P|}{\sqrt{|V|(|V| - |P|)}} = \sqrt{1 - \frac{|P|}{|V|}}$$

and hence

$$2 \sin^2 \frac{\theta}{2} = 1 - \cos \theta = 1 - \sqrt{1 - \frac{|P|}{|V|}} \geq \frac{1}{2} \frac{|P|}{|V|} \geq \frac{1}{2|V|}.$$

We conclude by lemma 2.29 that $\lambda_{\min}(\Delta_P) \geq 2\mu \sin^2 \frac{\theta}{2} = \Omega(1/|V|^3)$. □

2.2.2 Quantum Ring Machine

2.2.2.1 Definition

We define a new computational model, a *quantum ring machine* (QRM), and show that it is poly-time equivalent to a uniform class of quantum circuits. Recall that any uniform class of circuits—such as poly-time circuits or exponential-time circuits—inherits its uniformity condition from the corresponding class of

classical Turing machines producing these circuit families. To prove that QRMs are quantum-universal, we will encode the given Turing machine into a specific instance of a QRM whose inner workings correspond to those of the original Turing machine, but with an additional quantum tape. In other words, the local Hilbert space of our QRM will be partitioned into two parts: a classical part, storing individual cells of the TM's tape and the internal state of the TM, and a quantum part, storing one qubit per cell. However, since a general QRM does not need to have this specific internal structure, we first give an abstract definition.

Definition 2.31. *A quantum ring machine (QRM) is a tuple $(\mathbf{R}, n, |\psi_{\text{in}}\rangle, \mathcal{H}_{\text{halt}})$, where*

- $\mathbf{R} \in \text{U}(\mathbb{C}^d \otimes \mathbb{C}^d)$ is a unitary operator on a pair of qudits, each of dimension d ,
- $n \in \mathbb{N}$ is the total number of qudits on the ring,
- $|\psi_{\text{in}}\rangle \in \mathcal{H}$ is the initial state where $\mathcal{H} := (\mathbb{C}^d)^{\otimes n}$ denotes the joint Hilbert space,
- $\mathcal{H}_{\text{halt}} \subseteq \mathbb{C}^d$ is the halting subspace of each qudit.

Starting from a ring of n qudits initialised in $|\psi_{\text{in}}\rangle$, the operation \mathbf{R} is applied cyclically to adjacent pairs of qudits—see fig. 2.5—until some qudit indicates halting: its reduced density matrix has support completely within the halting subspace $\mathcal{H}_{\text{halt}}$; up until that point, the probability of finding any qudit within $\mathcal{H}_{\text{halt}}$ is zero¹¹.

Fig. 2.6 visualises a QRM as a quantum circuit. Because the ring is cyclic, we can arbitrarily mark a starting position on the ring. Starting at this position, part of the initial state $|\psi_{\text{in}}\rangle$ contains the input while the rest will be used as a workspace. The *input size* is thus upper bounded by the ring size.

In the following definition, we consider a slight extension of QRMs from definition 2.31 where $|\psi_{\text{in}}\rangle$ is replaced by a family of input states $\{|\psi_{\text{in}}(x)\rangle\}_{x \in I}$ for some index set I .

Definition 2.32. *A QRM terminates on $\{|\psi_{\text{in}}(x)\rangle\}_{x \in I}$ if it halts, in finitely many steps, on any initial state $|\psi_{\text{in}}(x)\rangle$ for $x \in I$. Let $(M_n)_n$ be a family of QRMs where M_n has a ring of size n . This family is poly-time terminating if there exists a polynomial p such that M_n terminates in $p(n)$ steps on all states $|\psi_{\text{in}}(x)\rangle$; similarly, it is exponential-time terminating if there exists an exponential function $f(n) = O(\exp(cn))$, for some $c > 0$, such that M_n terminates in $f(n)$ steps on all states $|\psi_{\text{in}}(x)\rangle$.*

¹¹In particular this means that if, after every application of \mathbf{R} , the corresponding two qudits are measured, each with respect to $\mathcal{H}_{\text{halt}}$ and its orthogonal complement $\mathcal{H}_{\text{halt}}^\perp$, then the probability of finding the reduced state in $\mathcal{H}_{\text{halt}}$ should always be either zero or one, with the latter case indicating halting.

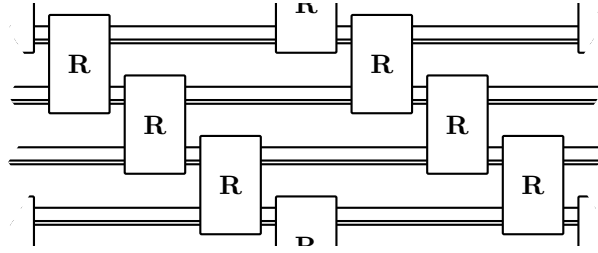


Figure 2.6: Circuit diagram of a QRM with a ring of size 4. The double lines indicate classical wires that are used to store the TM's internal states and tape, as well as a flag indicating either the TM's halting or the direction of its next head movement (see the proof of lemma 2.33 for more details). The internal details of the QRM's unitary operation \mathbf{R} are shown in fig. 2.7.

2.2.2.2 Universality

Lemma 2.33. *Let $\Pi = (\Pi_{\text{YES}}, \Pi_{\text{NO}})$ be a promise problem in BQP. Then there exists a polynomial p and a poly-time terminating family of QRMs $(M_n)_n$,*

$$M_n = (\mathbf{R}, n, \{|\psi_{\text{in}}(x)\rangle\}_{x \in I_n}, \mathcal{H}_{\text{halt}}),$$

with the following properties:

1. All M_n share the same unitary \mathbf{R} and the same terminating subspace $\mathcal{H}_{\text{halt}}$. The ring size of M_n is n .
2. The input states $\{|\psi_{\text{in}}(x)\rangle\}_{x \in I_n}$ of each M_n consist of trivial¹² encodings of instances $I_n := \{x \in \Pi : p(|x|) = n\}$, so that the whole computation fits on a ring of size n .
3. If $x \in \Pi_{\text{YES}}$, the reduced density matrix of the cell that signals halting satisfies an extra constraint: if measured, it collapses to an accepting subspace $\mathcal{H}_{\text{acc}} \subseteq \mathcal{H}_{\text{halt}}$ with probability $\geq 2/3$.
4. If $x \in \Pi_{\text{NO}}$, it collapses to $\mathcal{H}_{\text{acc}} \subseteq \mathcal{H}_{\text{halt}}$ with probability $\leq 1/3$.

Proof. Our goal is to construct a QRM for simulating a classical Turing machine (TM) that produces a description of a uniform quantum circuit. In addition to computing the circuit's description, the QRM also executes it one gate at a time. More formally, the QRM simulates a deterministic and reversible TM (see definition 2.7) augmented with the following quantum features:

- in addition to the classical data, each cell of the TM's tape stores one qubit,

¹²One must be able to produce $|\psi_{\text{in}}(x)\rangle$ from $x \in \Pi$ with a constant-depth quantum circuit (in particular, one cannot cheat by allowing the input to contain the answer to the problem), e.g. see eq. (2.6). This is similar to the types of input encodings one would allow for a poly-time classical TM.

- a special subset of the TM's states is associated with a universal set of two-qubit quantum gates; whenever the TM enters one of these states, the corresponding gate is applied on the two adjacent qubits that are stored in the pair of cells between which the TM's head just moved.

It is straightforward to verify that such quantum-enhanced TM is equivalent to a uniform family of quantum circuits.

Let us now describe the simulation procedure more formally. We write the complex linear span of a finite set S as $\mathbb{C}^S := \text{span}\{|s\rangle \in \mathbb{C}^{|S|} : s \in S\}$ and refer to $(\mathbb{C}^S)^{\otimes n}$ as a *ring* of size $n \in \mathbb{N}$, where each copy of \mathbb{C}^S represents one *cell* of the ring. Each cell further consists of three registers: a *quantum bit* (labelled by $\{0, 1\}$), a *classical data* register (labelled by elements of some finite set Γ), and a *flag* register (labelled by another set F). The standard basis of each ring cell is thus labelled by triples of the form

$$S := \{0, 1\} \times \Gamma \times F.$$

Using the notation from definition 2.7, let (Q, Σ, δ) be the deterministic TM we want to simulate (it is reversible without loss of generality, see theorem 2.9). The first register of the QRM stores the quantum state obtained by executing the quantum circuit produced by the TM. The second register Γ stores the TM's internal state and tape, so $\Gamma := Q \times \Sigma$ where Q is the set of internal states and Σ is the TM's alphabet. The flag symbols F in the third register are used to mark the location of the TM's head.

The flag register's alphabet is given by

$$F := \{\leftarrow, \rightarrow, -, h\}$$

and is used as follows. At any time, exactly one cell on the ring contains an *active flag* (either " \leftarrow ", " \rightarrow ", or " h ") while the rest are padded with " $-$ ". The TM's internal state is always stored in the active cell. Unless the TM has halted (indicated by flag " h "), the active flag shows in which direction (" \leftarrow " for *left* and " \rightarrow " for *right*) the TM's head must be moved before the simulation of the next step can begin. Every time the TM's head moves or its internal state changes, the QRM updates the flag registers and the description of the TM's internal state accordingly. Whenever the TM enters one of the special "quantum" states, the QRM applies the corresponding two-qubit unitary.

Recall from definition 2.31 that QRM operates by cyclically applying a fixed unitary \mathbf{R} on pairs of consecutive cells along the ring (see fig. 2.6). Most of the time \mathbf{R} acts trivially, since a non-trivial action is triggered only when either of the two active flags " \leftarrow " or " \rightarrow " is encountered. Note that \mathbf{R} acts on *two* adjacent cells,

one of them marked by the active flag and the other indicated by the direction of the flag's arrow:



It is crucial that \mathbf{R} is two-local for the following two reasons. First, updating the active location requires changing two symbols (e.g. when the TM's head moves left, we need to replace “ $- \leftarrow$ ” by “ $\leftarrow -$ ” or “ $\rightarrow -$ ”, depending on the direction the head will move next). For applying a two-qubit gate, we clearly also need a two-local interaction (we use the same convention as above to determine on which two qubits the gate is applied).

Recall from lemma 2.8 that, instead of quintuples δ , we can work with a permutation matrix \mathbf{T}_δ on Γ and a function $d : Q \rightarrow \{\text{left, right}\}$ telling us where the TM's head came from. For convenience, we include a special *dummy* state “ \perp ” in Q and a designated *blank* symbol “ $_$ ” in Σ : the dummy state is stored in all cells (except the active cell which stores the actual state of the TM) while the blank symbol is used to initialise the TM's tape. We accordingly extend \mathbf{T}_δ so that it acts trivially on $|\perp, \sigma\rangle$ for any $\sigma \in \Sigma$, and we define $d(\perp) := -$ so that dummy states do not trigger any action in our simulation.

We take the ring size to be $n = p(|x|)$ for an instance $x \in \Pi$, since the TM can access at most that many tape cells. We require that the ring starts out in a well-formed state, i.e. for some binary representation $x = x_1 x_2 \cdots x_l$ and $l = |x|$, a state of the form

$$|\psi_{\text{in}}(x)\rangle := \bigotimes_{j=1}^{n-l-1} (|0\rangle \otimes |\perp, _ \rangle \otimes |-\rangle) \otimes \bigotimes_{i=1}^l (|x_i\rangle \otimes |\perp, _ \rangle \otimes |-\rangle) \otimes (|0\rangle \otimes |q_0, _ \rangle \otimes |\rightarrow\rangle), \quad (2.6)$$

i.e. where all cells but the last are initialised as follows: the TM is in the dummy state “ \perp ”, the TM's tape is initialised to a designated blank symbol “ $_$ ”, and the flag is set to “ $-$ ”. The last cell contains the TM's initial state q_0 and the “ \rightarrow ” flag. The input $x \in \Pi$ is written on the qubit part of the tape, i.e. the first register of each cell.

We can now describe in more detail the steps involved in our simulation, and how to perform them reversibly (see figs. 2.7 and 2.8 for more details):

1. If the active cell has the halting flag “ h ”, the TM has halted so nothing happens.
2. If the active cell has one of the other two flags “ \leftarrow ” or “ \rightarrow ”:
 - a) The Q part of the Γ registers of the active cell and its neighbour—indicated by the flag—are exchanged, thus simulating the movement of the TM's head.
 - b) The flag register of the active cell is uncomputed using the function d .
 - c) Description of the TM's internal state and the current tape symbol is updated using \mathbf{T}_δ .

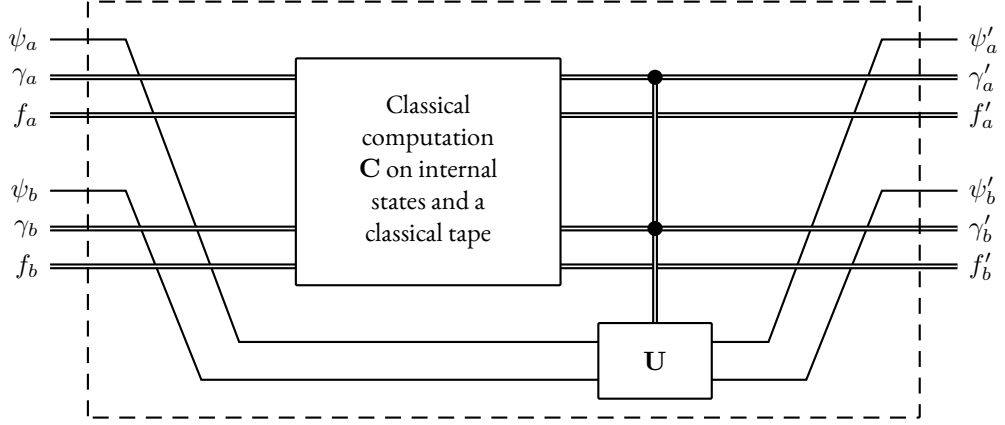


Figure 2.7: Circuit diagram for implementing the QRM's head unitary \mathbf{R} (double wires are classical while single wires are quantum). All computation is classical except for a single classically-controlled quantum gate \mathbf{U} that can be triggered by either of the two Γ registers. A classical circuit for implementing \mathbf{C} is shown in fig. 2.8.

- d) Based on the updated internal state, a new flag register is computed using d (it belongs to the same cell where the TM's new state is stored, and it indicates in which direction the TM's head will move before the next iteration begins).
- e) If the TM is in one of the special states indicating a quantum gate, the corresponding unitary is applied on the two data registers.

We now describe the unitary operator \mathbf{R} that acts on two adjacent QRM's cells:

- For $a = (\psi_a, \gamma_a, f_a) \in S$, write the corresponding basis state as $|a\rangle := |\psi_a, \gamma_a, f_a\rangle \in \mathbb{C}^S$ where $\psi_a \in \{0, 1\}$, $\gamma_a \in \Gamma$, $f_a \in F$, and analogously for $|b\rangle$. Then $|a\rangle \otimes |b\rangle \in \mathbb{C}^{S \times S}$ is also a basis state and we require, up to reordering the registers (see fig. 2.7), that

$$\mathbf{R}(|a\rangle \otimes |b\rangle) = |\Psi\rangle \otimes |\gamma'_a, f'_a\rangle \otimes |\gamma'_b, f'_b\rangle$$

for some $\gamma'_a, \gamma'_b \in \Gamma$ and $f'_a, f'_b \in F$, i.e. \mathbf{R} acts classically on each register except for the quantum data registers $|\psi_a\rangle$ and $|\psi_b\rangle$ (in particular, we allow $|\Psi\rangle \in \mathbb{C}^2 \otimes \mathbb{C}^2$ to be entangled).

- Using the same notation, if $f_a \neq \rightarrow$ and $f_b \neq \leftarrow$, we further demand $|\gamma'_a, f'_a\rangle = |\gamma_a, f_a\rangle$, $|\gamma'_b, f'_b\rangle = |\gamma_b, f_b\rangle$, and $|\Psi\rangle = |\psi_a\rangle \otimes |\psi_b\rangle$, i.e. if neither f_a nor f_b signal “apply head here”, \mathbf{R} acts as the identity operator on all registers.
- The active flag always moves in the direction indicated by the arrow. If $f_a = \rightarrow$ then $f'_a = -$ and $f'_b \neq -$, meaning that the head has moved right. Similarly, if $f_b = \leftarrow$ then $f'_b = -$ and $f'_a \neq -$,

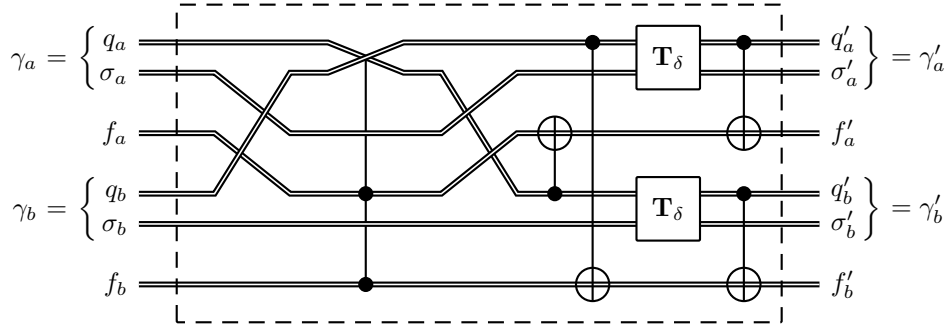


Figure 2.8: Circuit diagram for implementing the classical permutation \mathbf{C} in fig. 2.7 (all wires are classical and all gates are reversible). Conditioned on the flag registers being either “ $\rightarrow -$ ” or “ $- \leftarrow$ ”, the controlled-controlled-SWAP gate exchanges the internal state registers of the two cells. The CNOT gates in the first layer are conditioned on the value of $d(q)$, for state q , and they uncompute the flag register of the opposite cell (the cell where the TM’s head came from). The permutation \mathbf{T}_δ acts on Γ registers of both cells to update the TM’s internal state and the current tape symbol. Recall that \mathbf{T}_δ acts trivially if the state is dummy (at most one of the cells is in a non-dummy state). The final layer of CNOT gates again condition on $d(q')$, where q' is the new state, and update the flag registers to indicate where the TM’s head will move next. These flags will be uncomputed by the next iteration.

meaning that the head has moved left. In each case there are three possible transitions—they indicate whether the TM has halted or in which direction its head has to move next:

$$\begin{array}{cc}
 f_a f_b & f_a f_b \\
 \rightarrow - & - \leftarrow \\
 \Downarrow & \Downarrow \\
 f'_a f'_b & f'_a f'_b \\
 - \rightarrow & \rightarrow - \\
 - \leftarrow & \leftarrow - \\
 - h & h -
 \end{array}$$

Fig. 2.7 shows how \mathbf{R} acts on two adjacent cells. For each cell, the halting subspace $\mathcal{H}_{\text{halt}}$ is spanned by all standard basis vectors with the last register in the halting state $|h\rangle$:

$$\mathcal{H}_{\text{halt}} := \mathbb{C}^2 \otimes \mathbb{C}^\Gamma \otimes |h\rangle.$$

Fig. 2.8 provides details on how to implement \mathbf{C} reversibly.

We construct the desired family of QRMs $(M_n)_n$ in the special form described above. It is straightforward to verify that this ring machine executes the circuit written out by the TM, and the runtime overhead

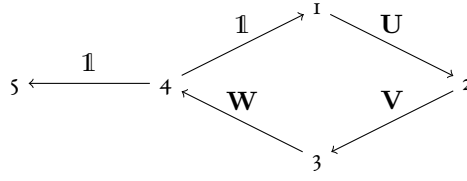


Figure 2.9: Unitary labelled graph from example 2.36. Observe that we mark the direction for the unitaries with an arrow, despite working with undirected graphs.

of M_n as compared to the circuit is at most quadratic. □

Corollary 2.34. *Using an exponential-time terminating family of QRMs, lemma 2.33 holds for BQEXP as well.*

2.2.3 Unitary Labelled Graphs

2.2.3.1 Definitions

The following definition introduces graphs whose vertices are labelled by Hilbert spaces and whose edges are labelled by unitaries between these spaces.

Definition 2.35. *Given an undirected graph $G = (S, E)$ without self-loops, a unitary labelled graph (ULG) is a triple $(G, (\mathcal{H}_v)_{v \in S}, g)$ where*

- $(\mathcal{H}_v)_{v \in S}$ is a family of Hilbert spaces, one space \mathcal{H}_v for each vertex $v \in S$,
- g is a function that assigns to each directed¹³ edge $ab \in E$ some unitary operator $g(ab) \in \mathbf{U}(\mathcal{H}_a)$ so that $g(ab) = g(ba)^\dagger$ (this requires that $\mathcal{H}_a \cong \mathcal{H}_b$ whenever $ab \in E$).

To facilitate notation, we will write an edge and its associated unitary jointly as $(a \leftrightarrow b, \mathbf{U})$ and call it a *rule*. By definition, the rule $(a \leftrightarrow b, \mathbf{U})$ is equivalent to the rule $(b \leftrightarrow a, \mathbf{U}^\dagger)$. With this notation, it is convenient to specify a unitary labelled graph by $G = (S, R)$ where $R := \{(a \leftrightarrow b, g(ab)) : ab \in E\}$ is the corresponding set of rules.

Example 2.36. *Let $S := \{1, 2, 3, 4, 5\}$, $\mathcal{H} = \mathbb{C}^2$ for all vertices, and consider the following set of rules:*

$$R := \{(1 \leftrightarrow 2, \mathbf{U}), (2 \leftrightarrow 3, \mathbf{V}), (3 \leftrightarrow 4, \mathbf{W}), (4 \leftrightarrow 1, \mathbf{1}), (4 \leftrightarrow 5, \mathbf{1})\}.$$

The underlying graph for this example is shown in fig. 2.9.

¹³While G is an undirected graph, we need to arbitrarily direct its edges so that we can discern between labels \mathbf{U} and \mathbf{U}^\dagger assigned to edges ab and ba , respectively.

Definition 2.37. Let $G = (S, R)$ be a ULG. If the product of unitaries along any directed path connecting a and b is equal, and this property holds for all $a, b \in S$, we call the ULG simple. Equivalently, for a ULG to be simple, the product of unitaries along any directed cycle should be $\mathbb{1}$.

The ULG in example 2.36 is simple if and only if $\mathbf{WVU} = \mathbb{1}$.

The following definition assigns a Hamiltonian to each ULG. This Hamiltonian extends the notion of a graph Laplacian, see definition 2.20, to ULGs (while this might not be immediately obvious from the definition, it will be made more clear in lemma 2.41 below).

Definition 2.38. Let $G = (S, R)$ be a connected ULG, \mathcal{H} denote the Hilbert space attached to each of its vertices, $n := \dim \mathcal{H}$ be the dimension of \mathcal{H} , and let $\{|e_i\rangle\}_{i=1}^n$ be some orthonormal basis of \mathcal{H} . The Hamiltonian associated to G is the following Hermitian operator on $\mathbb{C}^S \otimes \mathcal{H}$:

$$\mathbf{H}(G) := \sum_{(a \leftrightarrow b, \mathbf{U}) \in R} \sum_{i=1}^n (|a\rangle \otimes |e_i\rangle - |b\rangle \otimes \mathbf{U} |e_i\rangle)(\langle a| \otimes \langle e_i| - \langle b| \otimes \langle e_i| \mathbf{U}^\dagger). \quad (2.7)$$

This is reminiscent of eq. (2.4) for $\Delta(G)$, the Laplacian of graph G . Furthermore, it also explains why we excluded self-loops in definition 2.35: just as they have no effect on the graph Laplacian, they also impose no changes in the associated Hamiltonian of a simple ULG—the only possible self-loop unitary for such ULG is $\mathbb{1}$, making the corresponding term in eq. (2.7) vanish.

Proposition 2.39. The Hamiltonian $\mathbf{H} = \mathbf{H}(G)$ of a connected UGL $G = (S, R)$, see definition 2.38, is invariant under replacing any rule $(a \leftrightarrow b, \mathbf{U}) \in R$ with the corresponding inverse rule $(b \leftrightarrow a, \mathbf{U}^\dagger)$. Moreover, the matrix entries of \mathbf{H} do not depend on the choice of the basis $\{|e_i\rangle\}_{i=1}^n$.

Proof. Since G is connected, the Hilbert spaces attached to all its vertices are isomorphic. Observe further that

$$\mathbf{H} \equiv \sum_{(a \leftrightarrow b, \mathbf{U}) \in R} (|a\rangle\langle a| \otimes \mathbb{1}_n + |b\rangle\langle b| \otimes \mathbb{1}_n - |a\rangle\langle b| \otimes \mathbf{U}^\dagger - |b\rangle\langle a| \otimes \mathbf{U}), \quad (2.8)$$

hence the claim follows. \square

One can extend the notion of an associated Hamiltonian to a non-connected ULG as well by taking a direct sum of the Hamiltonians for each component of the graph (equivalently, one can assume that the Hilbert spaces associated to different components of the graph are mutually orthogonal and take the new Hilbert space to be their direct sum). Either way, such extension yields a block-diagonal associated Hamiltonian. •

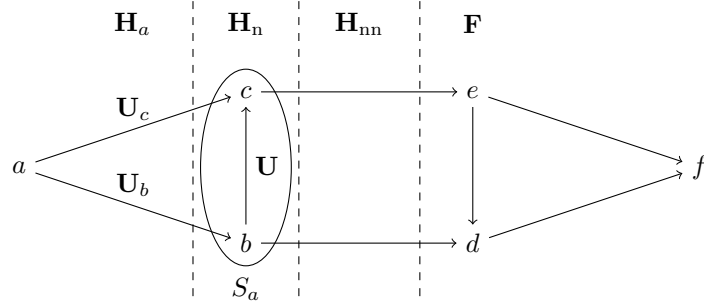


Figure 2.10: Terms of \mathbf{H} , see eq. (2.10), grouped according to how far the corresponding edges are from the chosen vertex a . For this ULG to be simple, the labels of edges forming the triangle abc must satisfy $\mathbf{U}_c^\dagger \mathbf{U} \mathbf{U}_b = \mathbb{1}$. We have omitted the labels of all other edges.

2.2.3.2 Semi-Classical Unitary Labelled Graphs

A ULG is semi-classical if its Hamiltonian is equal to a graph Laplacian (see definition 2.20), after a unitary change of basis.

Definition 2.40. *A ULG G is semi-classical if its associated Hamiltonian can be expressed as $\mathbf{H} = \mathbf{W}(\Delta \otimes \mathbb{1}_n)\mathbf{W}^\dagger$, where \mathbf{W} is some unitary operator, Δ is the Laplacian of G , and $\mathbb{1}_n$ acts on the n -dimensional Hilbert space attached to each vertex of G .*

- We note that the specific form of \mathbf{W} is left open on grounds of generality; in our case, however, \mathbf{W} will be block-diagonal in the time register basis. This definition can be easily extended to non-connected ULGs.

The following lemma is important for analysing the spectrum of any Hamiltonian coming from a simple ULG. It reduces the problem to analysing instead the spectrum of the corresponding graph Laplacian.

Lemma 2.41. *Any simple ULG is semi-classical.*

Proof. Denote the UGL by $G = (S, R)$ where S and R are the sets of vertices and rules, respectively. If G has disjoint components, \mathbf{H} is block-diagonal and we can deal with each block separately, hence we can assume without loss of generality that G is connected and all its vertices have isomorphic attached Hilbert spaces \mathcal{H} .

Pick an arbitrary vertex $a \in S$ and denote its set of neighbours by S_a . Using proposition 2.39, rewrite R in a form where a only has outgoing edges. Following eq. (2.8), define the term that encodes rule $(a \leftrightarrow b, \mathbf{U}_b) \in R$ as follows:

$$\mathbf{h}_{ab} := |a\rangle\langle a| \otimes \mathbb{1} + |b\rangle\langle b| \otimes \mathbb{1} - |a\rangle\langle b| \otimes \mathbf{U}_b^\dagger - |b\rangle\langle a| \otimes \mathbf{U}_b, \quad (2.9)$$

where the subscript of \mathbf{U}_b identifies the vertex with incoming edge. Then the terms of \mathbf{H} can be grouped

as follows (see fig. 2.10):

$$\begin{aligned} \mathbf{H} &= \sum_{\substack{(a \leftrightarrow b, \mathbf{U}_b) \\ b \in S_a}} \mathbf{h}_{ab} + \sum_{\substack{(b \leftrightarrow c, \mathbf{U}_c) \\ b, c \in S_a}} \mathbf{h}_{bc} + \sum_{\substack{(b \leftrightarrow d, \mathbf{U}_d) \\ b \in S_a \wedge d \notin S_a \cup \{a\}}} \mathbf{h}_{bd} + \mathbf{F} \\ &=: \mathbf{H}_a + \mathbf{H}_n + \mathbf{H}_{nn} + \mathbf{F}, \end{aligned} \tag{2.10}$$

where \mathbf{F} denotes the rest of the terms and all sums range over R , with some restrictions on the endpoints of the edges. Our strategy now is to apply a sequence of unitary transformations to bring the Hamiltonian \mathbf{H} to the desired form, a few terms at a time.

First, for the given vertex $a \in S$, define the following unitary:

$$\mathbf{W}_a := \prod_{\substack{(a \leftrightarrow b, \mathbf{U}_b) \\ b \in S_a}} (|b\rangle\langle b| \otimes \mathbf{U}_b + (\mathbb{1} - |b\rangle\langle b|) \otimes \mathbb{1}).$$

Observe that all terms in the product commute. Moreover, $\mathbf{W}_a^\dagger \mathbf{F} \mathbf{W}_a = \mathbf{F}$ and, by eq. (2.9),

$$\mathbf{W}_a^\dagger \mathbf{h}_{ab} \mathbf{W}_a = (|a\rangle\langle a| + |b\rangle\langle b| - |a\rangle\langle b| - |b\rangle\langle a|) \otimes \mathbb{1} \quad \bullet$$

for all $b \in S_a$, so \mathbf{W}_a takes care of all terms of \mathbf{H}_a simultaneously.

For the terms in \mathbf{H}_n , pick any edge $(b \leftrightarrow c, \mathbf{U}) \in R$, with $b, c \in S_a$, and note from eq. (2.9) that

$$\mathbf{W}_a^\dagger \mathbf{h}_{bc} \mathbf{W}_a = |b\rangle\langle b| \otimes \mathbb{1} + |c\rangle\langle c| \otimes \mathbb{1} - |b\rangle\langle c| \otimes \mathbf{U}_b^\dagger \mathbf{U}^\dagger \mathbf{U}_c - |c\rangle\langle b| \otimes \mathbf{U}_c^\dagger \mathbf{U} \mathbf{U}_b.$$

However, since abc is a cycle and the ULG is simple, the product of unitaries along the cycle must be $\mathbb{1}$, i.e. $\mathbf{U}_c^\dagger \mathbf{U} \mathbf{U}_b = \mathbb{1}$ (see fig. 2.10), so the formula simplifies to

$$\begin{aligned} \mathbf{W}_a^\dagger \mathbf{h}_{bc} \mathbf{W}_a &= (|b\rangle\langle b| + |c\rangle\langle c| - |b\rangle\langle c| - |c\rangle\langle b|) \otimes \mathbb{1} \\ &= (|b\rangle - |c\rangle)(\langle b| - \langle c|) \otimes \mathbb{1}. \end{aligned}$$

By a similar argument, we can show that the rules in \mathbf{H}_{nn} change their unitary by a factor of \mathbf{U}_b when outgoing, or \mathbf{U}_b^\dagger when incoming, respectively. We are left with a ULG with a new set of rules, namely, every edge that is either attached to a or between two different neighbours of a is trivial, i.e. the edge unitary is the identity operator $\mathbb{1}$.

We will apply the same procedure to different vertices until all edges become trivial. More specifically, we consider an arbitrary sequence of subsets $S_1 \subset S_2 \subset \dots \subset S_m \subset S$, starting from any vertex $\{a_1\} = S_1$ and ending with the set of all vertices S , such that each subsequent S_{k+1} can be obtained from

S_k by including all neighbours of some vertex $a_k \in S_k$. The overall unitary is then $\mathbf{W} = \prod_{k=1}^m \mathbf{W}_{a_k}$, where the product is over the sequence of vertices $a_1, a_2, \dots, a_m \in S$. Each successive unitary \mathbf{W}_{a_k} is obtained from the current set of rules R_k , where $R_1 = R$ is the original set while all rules in the final set are trivial. Our goal is to show that, at every step k , we can guarantee that each rule in R_k has a trivial unitary whenever both endpoints of the corresponding edge are in S_k .

We proceed by induction. Since there are no edges between vertices in $S_1 = \{a_1\}$, the induction basis is trivial. Assuming the inductive hypothesis holds for k , we take $a_k \in S_k$ and apply the unitary \mathbf{W}_{a_k} that acts non-trivially to all neighbours of a_k (recall that S_{k+1} is formed by S_k together with the neighbours of a_k). As discussed above, \mathbf{W}_{a_k} trivialises all edges between a_k and any of its neighbours. By simplicity of the ULG, it trivialises also all edges between any two different neighbours of a_k . Moreover, it does not affect any edges within S_k (they are trivial already by the inductive assumption). In other words, all edges between vertices in S_{k+1} are trivial, thus completing the induction.

Since all edge unitaries have now been transformed to the trivial unitary $\mathbb{1}$, all terms in $\mathbf{W}^\dagger \mathbf{H} \mathbf{W}$ are of the form

$$(|a\rangle\langle a| + |b\rangle\langle b| - |a\rangle\langle b| - |b\rangle\langle a|) \otimes \mathbb{1} = (|a\rangle - |b\rangle)(\langle a| - \langle b|) \otimes \mathbb{1}, \quad (2.11)$$

for some $a, b \in S$. Comparing this to eq. (2.4), the overall Hamiltonian is in fact equivalent to the Laplacian Δ of G , i.e. $\mathbf{W}^\dagger \mathbf{H} \mathbf{W} = \Delta \otimes \mathbb{1}$. \square

Lemma 2.42. *Let G be a simple ULG with vertices S and rules R . Write $G = G_1 \oplus \dots \oplus G_N$, $G_i = (V_i, E_i)$ for the associated Laplacian of the induced classical graph (S, R') with $R' := \{a \leftrightarrow b : (a \leftrightarrow b, \mathbf{U}) \in R\}$, and pick an arbitrary $v_i \in V_i \forall i$. Let further $n_i := \dim \mathcal{H}_i$ and choose a basis $\{|e_{i,j}\rangle\}_j$ of \mathcal{H}_i for all connected graph components G_i . Then the ground space $\ker \mathbf{H}$ is spanned by the set $\{|\Psi_{i,0}\rangle, \dots, |\Psi_{i,n_i}\rangle\}_{i=1}^N$, where*

$$|\Psi_{i,j}\rangle := \frac{1}{\sqrt{|V_i|}} \sum_{s \in V_i} |s\rangle \otimes |q_s\rangle \quad \text{and} \quad |q_s\rangle = \begin{cases} |e_{i,j}\rangle & \text{if } s = v_i, \\ \mathbf{U} |q_r\rangle & \text{if } (r \leftrightarrow s, \mathbf{U}) \in E_i. \end{cases}$$

Furthermore, the $|\Psi_{i,j}\rangle$ form a basis of $\ker \mathbf{H}$.

Proof. Because the ULG is simple, by lemma 2.41, there exists a unitary \mathbf{W} and a classical Laplacian Δ such that $\mathbf{W}^\dagger \mathbf{H} \mathbf{W} = \Delta \otimes \mathbb{1}_n$, and hence $\ker \mathbf{H} = \ker \Delta \otimes \mathbb{1}_n$. By lemma 2.25, the ground space of $\Delta \otimes \mathbb{1}_n$ has a basis given by

$$|\Phi_{i,j}\rangle := \frac{1}{\sqrt{|V_i|}} \sum_{s \in V_i} |s\rangle \otimes |e_i\rangle, \quad i = 1, \dots, N \quad \text{and} \quad j = 1, \dots, n.$$

Observe that

$$\mathbf{W} |\Phi_{i,j}\rangle = \frac{1}{\sqrt{|V_i|}} \sum_{s \in V_i} \prod_{a \in V_i} \mathbf{W}_a |s\rangle \otimes |e_i\rangle = |\Psi_{i,j}\rangle,$$

which can be easily verified. \square

2.2.3.3 Kitaev's Geometrical Lemma for Unitary Labelled Graphs

Analogous to section 2.2.1.6, we extend the notion of penalising vertices to ULGs. First we state an immediate corollary from lemma 2.30.

Corollary 2.43. *Take a connected simple ULG with Hilbert space \mathcal{H} for all vertices $s \in S$. Pick a non-empty subset of vertices $P \subsetneq S$ and write the penalised associated Hamiltonian $\mathbf{H}_P(G) := \mathbf{H}(G) + \mathbf{P}(G, P) \otimes \mathbb{1}_{\dim \mathcal{H}}$. Then $\lambda_{\min}(\mathbf{H}_P(G)) = \Omega(1/|S|^3)$.*

Proof. G is simple and $\mathbf{H}(G)$ has the same spectrum as $\Delta(G)$, up to multiplicity. Now use lemma 2.30 on $\Delta(G) + \mathbf{P}(G, P)$. \square

A more interesting case is when one does not want to penalise the entire Hilbert space attached to a vertex, but only a subspace. This is captured in the following lemma.

Lemma 2.44. *Take a connected simple ULG with Hilbert space \mathcal{H} for all vertices $s \in S$. Pick a non-empty subset of vertices $P \subsetneq S$ and a set of projectors $\Pi = \{\Pi_p\}_{p \in P}$ on \mathcal{H} . For $\mathbf{H}_P(G, \Pi) := \mathbf{H}(G) + \sum_{p \in P} |p\rangle\langle p| \otimes \Pi_p$, we have $\lambda_{\min}(\mathbf{H}_P(G, \Pi)) \geq \mu \Omega(1/|S|^3)$, where $\mu = 1 - \max\{|\lambda_{\max}(\Pi_i^c \mathbf{U}_{ij} \Pi_j^c)| : p_i, p_j \in P, i \neq j\}$ and \mathbf{U}_{ij} is the product of unitaries of a path connecting vertices p_i and p_j .*

Proof. First note that the \mathbf{U}_{ij} are well-defined, since the ULG is simple and connected. Construct \mathbf{W} such that the root of the sequence $S_1 \subset S_2 \subset \dots$ is one of the penalised vertices, namely r with projector Π_r . \bullet

Then

$$\begin{aligned} \mathbf{W}^\dagger \mathbf{H}_P(G, \Pi) \mathbf{W} &= \Delta(G) \otimes \mathbb{1} + |r\rangle\langle r| \otimes \Pi_r + \sum_{p \neq r} |p\rangle\langle p| \otimes \mathbf{R}_p \Pi_p \mathbf{R}_p^\dagger \\ &=: \Delta \otimes \mathbb{1} + \mathbf{A} \end{aligned}$$

where the \mathbf{R}_p are the product of unitaries connecting vertex p with the root r . Following the notation of lemma 2.30, we want to calculate the angle between the kernels of the Laplacian and the penalty terms. We write $\Pi_{\mathbf{A}} := \mathbb{1} \otimes \mathbb{1} - \sum_p |p\rangle\langle p| \otimes \Pi_p$ for the projector onto the kernel of the penalty terms. Then

$$\mathbf{W}^\dagger \Pi_{\mathbf{A}} \mathbf{W} = |r\rangle\langle r| \otimes \Pi_r^c + \sum_{p \neq r} |p\rangle\langle p| \otimes \mathbf{R}_p \Pi_p^c \mathbf{R}_p^\dagger + \sum_{v \notin P} |v\rangle\langle v| \otimes \mathbb{1}.$$

Noting that the kernel of $\Delta(G) \otimes \mathbb{1}$ is spanned by $\{|\Psi_V\rangle \otimes |\phi\rangle : |\phi\rangle \in \mathcal{H}\}$, we get

$$\begin{aligned} \cos \theta &= \max_{|\phi\rangle} \langle \Phi_V | \langle \phi | \mathbf{W}^\dagger \Pi_A \mathbf{W} | \Phi_V \rangle | \phi \rangle = \frac{1}{|V|} \max_{|\phi\rangle} \sum_{|v\rangle, |v'\rangle} \langle v | \langle \phi | \mathbf{W}^\dagger \Pi_A \mathbf{W} | \Phi_V \rangle | v' \rangle | \phi \rangle \\ &= \frac{|V| - |P|}{|V|} + \frac{1}{|V|} \max_{|\phi\rangle} \langle \phi | \Pi_r^c + \sum_{p \neq r} \mathbf{R}_p \Pi_p^c \mathbf{R}_p^\dagger | \phi \rangle \\ &\leq \frac{|V| - |P|}{|V|} + \frac{|P| - 1}{|V|} (1 + \cos \vartheta) \leq 1 - \frac{1}{|V|} (1 - \cos \vartheta), \end{aligned}$$

where we used eq. (2.5) in lemma 2.29 in the last line with a bound on the angle between subspaces $\cos \vartheta = \max_{p \neq r} \angle(\text{supp } \Pi_r^c, \text{supp } \mathbf{R}_p \Pi_p^c \mathbf{R}_p^\dagger)$ and $|P| \geq 2$. We can bound this further by

- $$\begin{aligned} \cos^2 \vartheta &= \max_{p \neq r} \max_{\substack{|\xi\rangle \in \text{supp } \Pi_r^c \\ |\eta\rangle \in \text{supp } \mathbf{R}_p \Pi_p^c \mathbf{R}_p^\dagger}} |\langle \xi | \eta \rangle|^2 = \max_{p \neq r} \max_{|\xi\rangle, |\eta\rangle} |\langle \xi | \Pi_r^c \mathbf{R}_p \Pi_p^c \mathbf{R}_p^\dagger | \eta \rangle|^2 \\ &\equiv \max_{p \neq r} \max_{|\xi\rangle, |\eta\rangle} |\langle \xi | \Pi_r^c \mathbf{R}_p \Pi_p^c | \eta \rangle|^2 \leq \max_{p \neq r} |\lambda_{\max}(\Pi_r^c \mathbf{R}_p \Pi_p^c)|^2 =: \lambda_{\max}^2. \end{aligned}$$

The rest follows lemma 2.30:

$$2 \sin^2 \frac{\theta}{2} = 1 - \cos \theta \geq 1 - 1 + \frac{1}{|V|} (1 - \lambda_{\max}) = \frac{1 - \lambda_{\max}}{|V|},$$

and the claim follows. \square

2.2.4 Quantum Thue Systems

2.2.4.1 Thue Systems

Let us briefly recall the idea behind classical Thue systems, also known as string rewriting systems.

Definition 2.45. *A Thue system—TS for short—is a tuple (Σ, R) of a finite alphabet Σ and a finite symmetric binary relation $R \subset \Sigma^* \times \Sigma^*$, where $\Sigma^* := \bigcup_{i=0}^{\infty} \Sigma^i$ denotes the set of all strings over the alphabet Σ .*

The binary relation R is usually written as a set of rewrite rules $a \leftrightarrow b$ and they are naturally extended to other strings in Σ^* : if $a \leftrightarrow b \in R$, then $c \leftrightarrow d$ in R if there exist $u, v \in \Sigma^*$ such that $c = uav$ and $d = ubv$. Thue systems with this extension—denoted by R^* —are a special case of abstract reduction systems and well-studied as computational models—see [Tho10].

Thue systems are multiway systems, i.e. starting from an initial string $a \in \Sigma^*$, exactly one substring is replaced at a time—in particular, this means that there might be branching points when the substitution is not unique. For our purpose, it is enough to consider length-preserving substitutions, i.e. any space

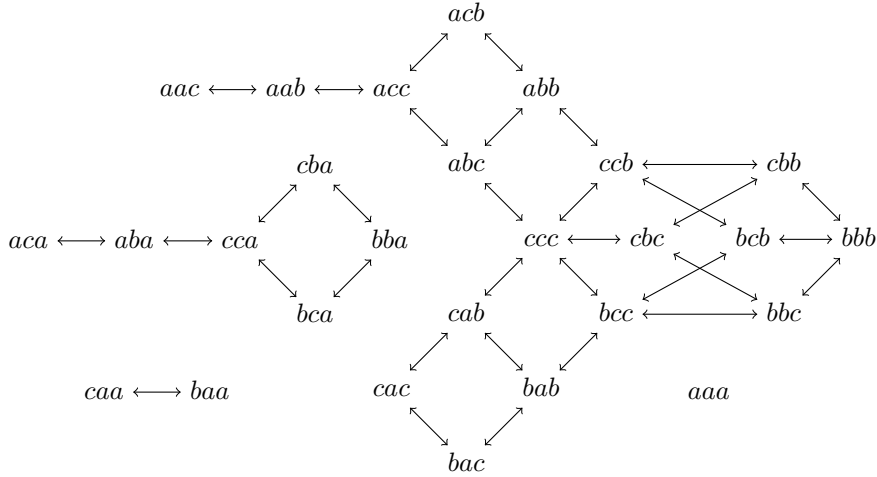


Figure 2.11: Undirected graph associated to strings of length 3 for the Thue system in example 2.47.

Σ^i should be invariant under R ; we denote the length of any string s with $|s|$. In this case, there exists a natural representation of the Thue system over strings of length N as a finite, undirected and not necessarily connected graph.

Definition 2.46. Let (Σ, R) be a TS and $N \in \mathbb{N}$. The associated graph $G = (V, E)$ for strings of length N has vertices $V := \Sigma^N$ and edges $E := \{(a, b) : a \leftrightarrow b \in R^*\}$. The Laplacian of the TS is defined as the discrete Laplacian of the associated graph G . For brevity we just write $G = (\Sigma, R)$.

Example 2.47. Take the alphabet $\Sigma := \{a, b, c\}$ and $R := \{c \leftrightarrow b, ab \leftrightarrow cc\}$. For example, starting from string acb , we can obtain a chain $acb \leftrightarrow abb \leftrightarrow ccb \leftrightarrow bcb \leftrightarrow bbb$. The entire graph for strings of length 3 for this example is shown in fig. 2.11.

Definition 2.48. Let (Σ, R) be a TS. We call a nonempty subset $U \subseteq \Sigma^*$ a valid evolution if U is closed under the transition rules R^* . We call U irreducible if there exists no valid evolution $U' \subsetneq U$.

Example 2.49. In example 2.47, the set $\{caa, baa, aaa\}$ is a valid evolution, but it is not irreducible. The only irreducible evolutions for strings of length 3 in this example are the sets formed from the connected components.

It is immediate to see this one-to-one correspondence between connected associated graph components and irreducible evolutions.

We want to introduce a sense of locality to TS relations.

Definition 2.50. A TS (Σ, R) is k -local, where k is the maximum length of any of its replacement rules. •

Observe how this definition is well-defined, as we required R to be finite, cf. definition 2.45.

2.2.4.2 Quantum Thue Systems and their Hamiltonian

We begin by generalising the notion of Thue systems to the case where our alphabet has special quantum symbols, with rewriting rules being unitary operators between them. To work with these two alphabets, consider $\Sigma = \Sigma_{cl} \sqcup \Sigma_q$ as the union of two disjoint—classical and quantum—alphabets. For a string $s \in \Sigma^*$, write $|s|_q$ for the number of letters from Σ_q in s . This allows the following definition.

Definition 2.51. *A quantum Thue system (QTS) is a quadruple $(\Sigma, R, \{\mathbf{U}_r\}_{r \in R}, \mathcal{H})$ of a bipartite alphabet $\Sigma = \Sigma_q \sqcup \Sigma_{cl}$, a relation R , a unitary operator \mathbf{U}_r for each rule $r \in R$ and a finite-dimensional Hilbert space \mathcal{H} with the following properties:*

- (Σ, R) is a TS,
- $|\cdot|_q$ is invariant under any rule $r \in R$,
- $\mathbf{U}_r \in \mathbf{U}(\mathcal{H}^{\otimes |r|_q})$ for all $r \in R$.

The invariance of $|\cdot|_q$ under a rule $r = s_1 \leftrightarrow s_2$ allows to abbreviate $|r|_q := |s_1|_q = |s_2|_q$, which indicates the number of quantum letters the rule r acts on.

We can again use the QTS to form sequences of strings: starting from a string $s \in \Sigma^*$, apply rules consecutively as for TSs. In addition, to each string $s \in \Sigma^*$, we attach a Hilbert space $\mathcal{H}_s := \mathcal{H}^{\otimes |s|_q}$: starting from some vector $v \in \mathcal{H}_s$, each time a rule r is applied to a substring, the corresponding unitary acts on the subspace wherever the rule matches, acting as identity everywhere else.

Analogous to fig. 2.11, we can build a graph for strings of length L for any QTS—where each edge is labelled by the acting unitary. The following lemma should therefore not come as a surprise.

Lemma 2.52. *Any k -local QTS $(\Sigma, R, \{\mathbf{U}_r\}, \mathcal{H})$ restricted to strings of a certain length $N \geq k$ is also a ULG. Furthermore, the associated Hamiltonian for strings of length N is isomorphic to a geometrically k -local and translationally-invariant Hamiltonian on a chain $(\mathbb{C}^\Sigma \otimes \mathcal{H})^{\otimes N}$ with the same spectrum.*

Proof. We explicitly define the ULG (S, R') for strings of length $N \geq k$. The vertex set $S := \Sigma^N$ is straightforward. For every $r \in R$ denoted $s_1 \leftrightarrow s_2$, define the ULG edges $(us_1v \leftrightarrow us_2v, \mathbb{1}_{\mathcal{H}}^{\otimes |u|} \otimes \mathbf{U}_r \otimes \mathbb{1}_{\mathcal{H}}^{\otimes |v|})$ for any $u, v \in \Sigma^*$ —potentially extending \mathbf{U}_r to $\mathcal{H}^{\otimes N-k}$ acting trivially on classical substrings—such that $us_1v \in \Sigma^N$. It is straightforward to verify that this defines a valid ULG.

The second claim follows from the canonical isomorphism between the two Hilbert spaces $(\mathbb{C}^\Sigma)^{\otimes N} \otimes \mathcal{H}^{\otimes N} \xrightarrow{\sim} (\mathbb{C}^\Sigma \otimes \mathcal{H})^{\otimes N}$, i.e. a simple rearrangement. Conjugating the associated Hamiltonian of the ULG with this isomorphism proves the second claim. \square

As QTSs are also ULGs, we will be—without always specifying the string length restriction explicitly—using ULG terminology for QTSs, e.g. associated Hamiltonians, irreducible evolutions or speak of QTSs being simple.

Lemma 2.53. *Let the setup be as in lemma 2.52. Then the isomorphism extends to a Hamiltonian on the chain $(\mathbb{C}^{\Sigma_{cl}} \oplus (\mathbb{C}^{\Sigma_q} \otimes \mathcal{H}))^{\otimes N}$ with the same spectrum up to multiplicities.*

Proof. Any rule on a classical substring acts identically on \mathcal{H} , hence on the set of Hamiltonians with this property a conjugation of the isomorphic Hamiltonian in lemma 2.52 \mathbf{H} with the projector $(\mathbb{C}^{\Sigma} \otimes \mathcal{H})^{\otimes N} \longleftrightarrow (\mathbb{C}^{\Sigma_{cl}} \oplus (\mathbb{C}^{\Sigma_q} \otimes \mathcal{H}))^{\otimes N}$ is an isomorphism. The projector preserves the spectrum, up to multiplicities. \square

In the following, we often gloss over the fact and simply assume that the associated Hamiltonian of a QTS is local in the sense of lemma 2.53. Observe however that a ULG induced from a QTS is not necessarily simple, and it is easy to find a counterexample.

2.2.4.3 Quantum Thue Systems as a Computational Model

To use a QTS for computation, we need to mark some strings that have special meaning, e.g. are input or output of the computation.

Definition 2.54. *For a QTS with alphabet Σ and Hilbert space \mathcal{H} , a marker is any tuple (s, π) where $s \in \Sigma^*$ and π is a projector on some subspace of $\mathcal{H}^{\otimes |s|_q}$. The set of markers on strings of length k —called k -local markers—is denoted $\mathcal{M}^{(k)}$, and $\mathcal{M} := \bigcup_k \mathcal{M}^{(k)}$.*

That is, we can specify a string s and a configuration of the quantum part of this string as a specific state in the computation. It is useful to think of using one marker s_{inp} to mark a string as the start of the computation, and a second one s_{out} to mark the end; the quantum parts of the markers— Π_{inp} and Π_{out} —then define the valid input and output of the computation.

Definition 2.55. *Let $(\Sigma, R, \{\mathbf{U}_r\}, \mathcal{H})$ be a QTS and $\Pi = (\Pi_{\text{YES}}, \Pi_{\text{NO}})$ a promise problem. We introduce an encoding function $\text{enc} : \Pi \rightarrow \Sigma^*$, input and output markers $(s_{\text{inp}}, \Pi_{\text{inp}}), (s_{\text{out}}, \Pi_{\text{out}}) \in \mathcal{M}^{(n)}$ for some $n \in \mathbb{N}$. Then the QTS*

- rejects an instance $l \in \Pi$ if there exists a chain of rules in R connecting $\text{enc}(l)$ with two strings containing s_{inp} and s_{out} , respectively, and $\langle \psi | \Pi_{\text{inp}} + \mathbf{U}^\dagger \Pi_{\text{out}} \mathbf{U} | \psi \rangle \geq \epsilon$ for all $|\psi\rangle \in \mathcal{H}$ —here $\mathbf{U} = \mathbf{U}(l)$ stands for the product of unitaries along this chain, and Π_{inp} and Π_{out} are extended trivially to the entire chain in case $|s_{\text{inp}}|_q > |\text{enc}(l)|_q$ or $|s_{\text{out}}|_q > |\text{enc}(l)|_q$.
- accepts l if there exists a $|\psi\rangle \in \mathcal{H}$ such that $\langle \psi | \Pi_{\text{inp}} + \mathbf{U}^\dagger \Pi_{\text{out}} \mathbf{U} | \psi \rangle \leq \epsilon/2$.
- decides Π if for all $l \in \Pi$, l is accepted if $l \in \Pi_{\text{YES}}$, and rejected for $l \in \Pi_{\text{NO}}$.

The rejection and acceptance threshold ϵ will depend on the class of promise problems that we want to decide. In particular, we want to allow this threshold to scale with the problem instance size, i.e. $\epsilon = \epsilon(|l|)$, and thus indirectly with the time that a computation can take, as specified in the following definition.

Definition 2.56. Let Q be a QTS that decides Π . For an instance $l \in \Pi$, the history state is defined as the irreducible evolution of the ULG containing $\text{enc}(l)$.

For a QTS with unambiguous transition rules—i.e. where the history state is a line—the size of the history state simply corresponds to the runtime of the underlying computation.

We now want to describe a simple example for a QTS which can decide the following simple promise problem.

Definition 2.57 (EVEN NATURAL NUMBER).

Instance. Natural number $n \in \mathbb{N}$.

Output. YES if n even, otherwise NO.

Example 2.58. Let the alphabet $\Sigma := \{-, \star, \|\}$, where \star is the only quantum symbol with Hilbert space \mathbb{C}^2 . We define $s_{\text{out}} = \star\|$ and $\Pi_{\text{out}} = |1\rangle\langle 1|$. Let further

$$\text{enc} : \mathbb{N} \rightarrow \Sigma^* \quad \text{where} \quad \text{enc}(n) := \star \underbrace{- - \dots -}_{n \text{ times}} \|\.$$

- $s_{\text{inp}} = \text{enc}(l)$, $s_{\text{out}} = \star\|$, and $\Pi_{\text{inp}} = \Pi_{\text{out}} = |1\rangle\langle 1|$. We have a single rule $(\star- \leftrightarrow -\star, \mathbf{R})$ where \mathbf{R} is a rotation by $\pi/2$, i.e. $\mathbf{R} := -|1\rangle\langle 0| + |0\rangle\langle 1|$. Then this QTS decides EVEN NATURAL NUMBER.

Proof. The proof is straightforward. Starting on the encoded input $\text{enc}(n)$, the TS generates a sequence

$$\star - - \dots - \|\ \mapsto \ - \star - \dots - \|\ \mapsto \ \dots \mapsto \ - - \dots - \star\|,$$

so there always exists a chain of rules that connects $\text{enc}(n)$ with a string containing s_{out} . The decision is thus made by the content of the quantum part: for n applications of the rule, starting from a vector $|\psi\rangle \in \mathbb{C}^2$, we apply \mathbf{R} n times. Now take any state $|\psi\rangle \in \mathbb{C}^2$ and write $|\psi\rangle = a|0\rangle + b|1\rangle$. Then

$$\begin{aligned} & \langle 0| \left(\Pi_{\text{inp}} + (\mathbf{R}^\dagger)^n \Pi_{\text{out}} \mathbf{R}^n \right) |0\rangle \\ &= |\langle 0| \mathbf{R}^n |1\rangle|^2 = |\langle 0| \mathbf{R}^{n \bmod 2} |1\rangle|^2 = \begin{cases} 0 & \text{if } n \text{ is odd,} \\ 1 & \text{otherwise,} \end{cases} \end{aligned}$$

and

$$\begin{aligned} & \langle 1| \left(\Pi_{\text{inp}} + (\mathbf{R}^\dagger)^n \Pi_{\text{out}} \mathbf{R}^n \right) |1\rangle \\ &= 1 + |\langle 1| \mathbf{R}^n |1\rangle|^2 \geq 1. \end{aligned}$$

Therefore, if n is even,

$$\langle \psi| \left(\Pi_{\text{inp}} + (\mathbf{R}^\dagger)^n \Pi_{\text{out}} \mathbf{R}^n \right) |\psi\rangle \geq |a|^2 + |b|^2 = 1$$

and $|0\rangle$ is an accepting state for odd n . The claim follows. \square

2.2.5 Hardness Result

2.2.5.1 A Special Kind of Quantum Thue System

We have seen that QTS can be used to answer simple problems. On the other hand, a more interesting question is whether there exists a universal QTS which can run any computation of a certain class of promise problems \mathbf{C} , i.e. is complete for \mathbf{C} . Of particular interest in this setting is the question about scaling of the defining parameters for such a QTS: how big is the alphabet, what is the locality and how does the string length of the vertices in the history state scale, i.e. for a promise problem $\Pi \in \mathbf{C}$, does there exist a function f such that for $l \in \Pi$, $|\text{enc}(l)| = O(f(|l|))$? And what about the size of the history state?

For the complexity class \mathbf{BQEXP} , we have the following lemma.

Lemma 2.59. For any \mathbf{BQEXP} promise problem Π , there exists a 2-local QTS $(\Sigma, R, \{\mathbf{U}_r\}, \mathbb{C}^2)$ which decides Π , and has the following uniform properties:

\mathbf{W}_1 *The alphabet has special characters $H \subset \Sigma$ —heads—and $B \subset \Sigma$ —boundaries, and a set of allowed pairs $\mathfrak{A} \subset \Sigma \times \Sigma$. All transition rules preserve any symbols from B and the number of symbols in H (denoted $|s|_h$ for a string s).*

\mathbf{W}_2 *Let $j : \Pi \rightarrow \mathbb{N}$ be a map with $j(l) = O(\exp \text{poly } |l|)$, where $|l|$ denotes the size of instance $l \in \Pi$. The QTS decides instance l on strings of this length, i.e. $|\text{enc}(l)| = j(l)$. Both input and output penalty are 2-local markers containing precisely one head symbol, and $s_{\text{inp}}, s_{\text{out}} \in \mathfrak{A}$.*

\mathbf{W}_3 *For any $l \in \Pi$, the history state \mathcal{M}_l is simple. All strings $s \in \mathcal{M}_l$ are of the form $s \in B \times (\Sigma \setminus B)^* \times B$ (bracketed), and have one head $|s|_h = 1$. Furthermore, all length-2 substrings of s are in \mathfrak{A} , and the size $|\mathcal{M}_l| = \text{poly}(j(l))$.*

\mathbf{W}_4 *For all other irreducible evolutions $\mathcal{M} \neq \mathcal{M}_l$, at least one of the following is true:*

- $|s|_h = 0 \forall s \in \mathcal{M}$,
- \mathcal{M} is not bracketed (i.e. with a boundary symbol on the left and right ends),
- \mathcal{M} can be broken up into $O(g(l))$ -sized connected parts—where $g(l) = \text{poly}(j(l))$ —each of which containing at least one string containing an invalid character tuple not in \mathfrak{A} .

2.2.5.2 $\mathbf{QMA}_{\text{EXP}}$ -Hardness Proof

Theorem 2.60. $(2, 42)$ -HAMILTONIAN is $\mathbf{QMA}_{\text{EXP}}$ -hard.

Proof. Following definitions 2.13 and 2.16, we need to show that there exists a 1D translationally-invariant 2-local Hamiltonian \mathbf{H} on $\mathcal{H} = (\mathbb{C}^{42})^{\otimes n}$ with $O(1)$ local terms, such that either (a) $\lambda_{\min}(\mathbf{H}) \leq \alpha$ or (b) $\lambda_{\min}(\mathbf{H}) \geq \beta$ with a polynomial promise gap $\beta - \alpha = \Omega(1/\text{poly } n)$, and deciding between (a) and (b) is at least as hard as some QMA_{EXP} -hard promise problem. The proof will be a simple combination of our previously-collected results.

1. Let Π be a promise problem in QMA_{EXP} . By definition 2.13, the verification of Π is a BQEXP problem. By fact 2.12, we can assume without loss of generality that the accept and reject probabilities in definition 2.55 are $1 - \epsilon$ and ϵ , respectively, where $\epsilon = 1/3^{\text{poly } |l|}$ to be specified below, where $|l|$ denotes the length of the problem input.
2. By corollary 2.72, we can thus create a QTS with properties as in lemma 2.59 that verifies Π : more specifically, for an instance $l \in \Pi$ and by item W2, we know that this QTS verifies l on strings of length $j(l)$.
3. By lemma 2.52, the QTS restricted to strings of length $j(l)$ is also a ULG. Denote the Hamiltonian associated to this ULG by \mathbf{H}_l , block-diagonal in the irreducible evolutions.

With Γ denoting the alphabet from definition 2.62 and $j(l)$ denoting the number of systems, we define a Hamiltonian on the Hilbert space $(\mathbb{C}^\Gamma)^{\otimes j(l)}$ as follows:

$$\mathbf{H} := \mathbf{H}_l + \mathbf{B}_{\text{heads}} + p(l)(\mathbf{P}_{\text{boundaries}} + \mathbf{P}) + \mathbf{P}_{\text{in/out}}, \quad (2.12)$$

where

- $\mathbf{P}_{\text{boundaries}}$ penalises any non-bracketed string (i.e. strings without a boundary symbol on at least one end)—cf. item W3,
- $\mathbf{B}_{\text{heads}}$ acting on a string $|s\rangle \in (\mathbb{C}^\Gamma)^{\otimes j(l)}$ gives a bonus of $|s|_h$, according to how many head symbols there are in s ,
- \mathbf{P} penalises any character tuple not in \mathfrak{A} ,
- $p(l)$ is a function used to scale the penalties, which will be specified later, but—keeping remark 2.18 in mind—must not exceed $p(l) = \text{poly } j(l)$.

Penalising non-bracketed strings follows an idea by [GI13]. With $\mathbf{P}_{\text{boundaries}}$, we give a 1-local bonus of size 1 to brackets appearing anywhere, but a penalty of 1/2 to them appearing next to any other symbol; since no transition rule ever moves the boundaries, this gives a uniform energy shift to all strings with brackets. The unique highest-bonus string will have a bracket appearing at the start and end with a bonus of size 1.

The encoding and output penalties $\Pi_{\text{inp}} = \Pi_{\text{out}} = |1\rangle\langle 1|$ are translationally-invariantly extended to the entire chain, i.e. on Hilbert space $(\mathcal{H}_{cl} \oplus \mathcal{H}_q)^{\otimes j^{(l)}}$, we act with the 2-local projector

$$\mathbf{P}_{\text{in}} := \sum_{i=1}^{j^{(l)}-1} (|s_{\text{inp}}\rangle\langle s_{\text{inp}}| \oplus \Pi_{\text{inp}})_{i,i+1},$$

and analogously for \mathbf{P}_{out} .

Completeness. Assume l is a YES-instance, and denote the history state as an eigenvector of \mathbf{H}_l with $|\Psi_l\rangle$, which by item W₃ is also an eigenstate of $\mathbf{B}_{\text{heads}}$, $\mathbf{P}_{\text{boundaries}}$ and \mathbf{P} . A direct calculation yields

$$\begin{aligned} \langle \Psi_l | \mathbf{H} | \Psi_l \rangle &= \overbrace{\langle \Psi_l | \mathbf{H}_l | \Psi_l \rangle}^{=0} + \overbrace{\langle \Psi_l | \mathbf{B}_{\text{heads}} | \Psi_l \rangle}^{=-1} + p(l) \left(\overbrace{\langle \Psi_l | \mathbf{P}_{\text{boundaries}} | \Psi_l \rangle}^{=-1} + \overbrace{\langle \Psi_l | \mathbf{P} | \Psi_l \rangle}^{=0} \right) \\ &\quad + \langle \Psi_l | \mathbf{P}_{\text{in/out}} | \Psi_l \rangle \\ &= -2 + \langle \Psi_l | \mathbf{P}_{\text{in/out}} | \Psi_l \rangle. \end{aligned}$$

By item W₂, we further know that at least one vertex in $|\Psi_l\rangle$ has the in- and output substrings s_{inp} , s_{out} , and because $|s_{\text{inp}}|_h = |s_{\text{out}}|_h = 1$, there is at most one such substring match for every vertex. As an upper bound, we can thus assume that the penalty applies exactly once in every vertex—i.e. $\langle \Psi_l | \mathbf{P}_{\text{in/out}} | \Psi_l \rangle \leq \epsilon$, and conclude $\langle \Psi_l | \mathbf{H} | \Psi_l \rangle \leq -2 + \epsilon$.

Soundness. Assume l is a NO-instance. We need to lower-bound the lowest energy eigenvalue of \mathbf{H} , and since we know that \mathbf{H} is block-diagonal in the irreducible evolutions, we can bound each block separately—the history state block given in item W₃ and any other irreducible evolution block characterised by item W₄. Without loss of generality we can therefore assume that $|\psi\rangle$ is completely supported on a single block of \mathbf{H} (but not necessarily an eigenvector).

Take any $|\psi\rangle$ with support constrained to the history state block. As in the completeness part, a direct calculation allows the estimate

$$\begin{aligned} \langle \psi | \mathbf{H} | \psi \rangle &= \langle \psi | \mathbf{H}_l | \psi \rangle + \overbrace{\langle \psi | \mathbf{B}_{\text{heads}} | \psi \rangle}^{=-1} + p(l) \left(\overbrace{\langle \psi | \mathbf{P}_{\text{boundaries}} | \psi \rangle}^{\geq -1} + \overbrace{\langle \psi | \mathbf{P} | \psi \rangle}^{\geq 0} \right) \\ &\quad + \langle \psi | \mathbf{P}_{\text{in/out}} | \psi \rangle \\ &\geq -2 + \langle \psi | \mathbf{H}_l + \mathbf{P}_{\text{in/out}} | \psi \rangle. \end{aligned}$$

We can now apply lemma 2.44 to the last expression. By item W₂ and definition 2.55, we obtain a bound $\langle \psi | \mathbf{H}_l + \mathbf{P}_{\text{in/out}} | \psi \rangle \geq (1-\epsilon)/|\mathcal{M}_l|^3$. Observe how this lower bound scales $\propto 1/|\mathcal{M}_l|^3$, whereas for YES-instances the upper bound scales constant $\propto \epsilon$. Since we want the lower bound for NO-instances— β —

and the upper bound for YES-instances— α —to be separated by at least some $\beta - \alpha = \Omega(1/\text{poly } j(l))$, cf. definition 2.16, we need to amplify the accepting probability to $\epsilon = O(1/|\mathcal{M}_l|^4) = O(1/\text{poly } j(l)^4)$. Observe that this does not exceed the allowed amplification, which is only limited to $O(1/3^{\text{poly } |l|})$.

We proceed to show lower bounds for all other minimum valid evolutions, following item W_4 . Assume we are in a block with o heads, which is well-defined by item W_1 . The bonus term $\mathbf{B}_{\text{heads}}$ vanishes on this subspace while all other operators in eq. (2.12) are positive semi-definite, so we obtain a lower bound of $\langle \psi | \mathbf{H} | \psi \rangle \geq 0$ for any state solely supported there.

Analogously, non-bracketed blocks can be bounded by a direct calculation, as $\mathbf{P}_{\text{boundaries}}$ penalises all vertices equally: any non-bracketed state $|\psi\rangle$ for a block with h heads satisfies $\langle \psi | \mathbf{H} | \psi \rangle \geq -h + p(l)$. It thus suffices to set $p(l) \geq j(l)$, as the number of possible heads on a string is limited by its length, i.e. $h \leq j(l)$.

The last blocks remaining are the ones with $g(j(l))$ -sized connected parts with invalid tuples, where $g(n) = \text{poly } n$ as defined in item W_4 . First observe that this part of the ULG is not necessarily simple, so we remove the transitions which allow non-trivial loops without breaking the graph up into multiple parts. We *then* split this graph into $g(j(l))$ -sized connected components by temporarily removing further edges from it, which yields a Hamiltonian for a sparser graph \mathbf{H}' . Since adding any edges back in corresponds to adding a positive semi-definite matrix to \mathbf{H}' , it suffices to lower-bound the spectrum of \mathbf{H}' on this subspace. Note that we do not remove vertices or change any penalties, so in particular all the diagonal operators in eq. (2.12) remain untouched.

Hence assume $|\psi\rangle$ has support in one of the slices of size upper-bounded by $g(j(l))$ with h heads, such that at least one vertex picks up a penalty from \mathbf{P} . Again applying lemma 2.44, we obtain a bound $\langle \psi | \mathbf{H} | \psi \rangle \geq \langle \psi | \mathbf{H}' | \psi \rangle \geq -h - 1 + p(l) \times \Omega(1/g(j(l))^3)$. We therefore have to scale p to e.g. $p(l) \geq g(j(l))^5$, which is still allowed by remark 2.18 (namely, p is polynomial in l). This concludes the proof. \square

What remains to be shown is the existence of a QTS as in lemma 2.59. The next section will provide an explicit construction, finalising the proof of our main result. This construction is meant as a proof-of-concept—the model we present can be modified in numerous ways and is likely not optimal. It does, however, make heavy use of our newly-developed methods such as branching, thus reducing the local dimension of the underlying Hamiltonian to 42, as compared to the hitherto best result by [GI13] which is larger by at least several orders of magnitude.

2.2.6 Turing's Wheelbarrow

2.2.6.1 Introduction

Turing's Wheelbarrow is our constructive proof of a QTS with properties as mentioned in lemma 2.59. The QTS will be optimised for local dimension and locality—every transition rule will be 2-local and act on

strings from an alphabet Γ with 48 characters. We describe the QTS by explicitly writing out all transition rules of the QTS and then prove the properties from lemma 2.59. Finally, in section 2.2.7 we reduce its local dimension down to 42.

The conceptual idea of the Wheelbarrow QTS is the following. To build a QTS which can decide a promise problem $\Pi \in \text{BQEXP}$, we first prefix the original circuit C_l deciding an instance $l \in \Pi$ by another circuit which verifies that a number of ancillas necessary for C_l are correctly initialised to $|0\rangle$. On some extra ancillas, we write out the problem instance l , and also leave an unconstrained section of qubits available for C_l . This witness section, problem instance and the leftover ancillas are then fed into C_l , and the output wire contains $|\text{out}\rangle = \cos((p_a + p_{\text{out}})/3) |0\rangle + \sin((p_a + p_{\text{out}})/3) |1\rangle$ for the amplitudes p_a —all ancillas being 0—and p_{out} —the circuit output of C_l on the ancillas and problem instance. This overall circuit, denoted C'_l , is shown in fig. 2.12.

It is clear that this augmented circuit family $(C'_l)_{l \in \Pi}$ is in the same uniformity class as the original circuit family $(C_l)_{l \in \Pi}$, and we can thus define these circuits with output $|\text{out}\rangle$ to be a separate **BQEXP** problem Π' . By lemma 2.33 and its proof, this new promise problem can be decided by a family of **BQEXP** QRMs with the special property that the head motion and all internal QRM states are classical—cf. fig. 2.7.

Using the Solovay-Kitaev theorem [NCro, appdx. 3], the head unitary of such a QRM can in turn be efficiently rewritten as a circuit R using the following small set of gates.

Remark 2.61. TOFFOLI, SWAP and a classically-controlled quantum-universal unitary together with at least one classical and quantum ancilla is universal for quantum computation and exactly universal for classical computation.

In particular, S-K tells us that since the head circuit $R = R(l)$ depends on the problem instance l —as it needs to write the instance out—and the size of this circuit is $|R(l)| = \text{poly } |l|$. The Wheelbarrow QTS which we construct will then be able to execute this head cyclically on a ring of qubits, where the execution is halted once the QRM terminates: as the QRM motion is deterministic, the runtime will be $\exp \text{poly } |l|$ -bounded, as required for a **BQEXP** computation.

The first step is to bootstrap the QRM head \mathbf{U} . Starting from an initially empty string, we use a number of rules to translate the string length N into a circuit description of \mathbf{U} on the left side of the string. This section will have size $\approx \log_6 N$, as we need 6 instruction symbols—a classically-controlled unitary \mathbf{U} , TOFFOLI \mathbf{T} , ancilla-checking symbol \mathbf{A} , swap \mathbf{S} , left-shift symbol \mathbf{D} and halt \mathbf{H} . The remaining right side of the string will act as classical and quantum tape that the computation runs on. Fig. 2.13 outlines how a circuit can be translated into such a 6-ary circuit description.

The QRM is then executed: for every round, a program bit is taken from the left side of the string, moved towards the tape and then applied to the leftmost two data qubits. The leftmost data qubit is then picked up and carried to the right, where it is deposited. The revert action is similar, only that the rightmost data

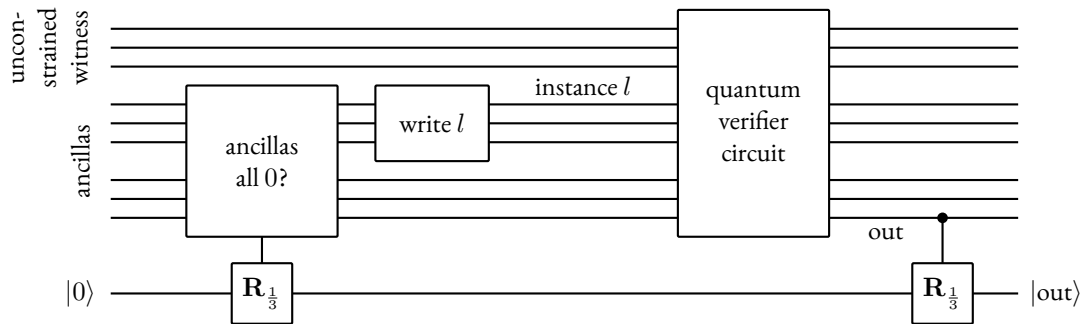


Figure 2.12: An augmented quantum verifier circuit. The circuit uses one ancilla $|0\rangle$ to verify that as many ancillas as necessary for the computation are set to 0, rotating the single guaranteed $|0\rangle$ ancilla by $\pi/3$ if this is not the case. On some ancillas, the problem instance l is written out. Another rotation by $\pi/3$ is applied depending on the output of the verifier circuit. The overall output state then takes the form $|\text{out}\rangle = \cos((p_a + p_{\text{out}})/3) |0\rangle + \sin((p_a + p_{\text{out}})/3) |1\rangle$.

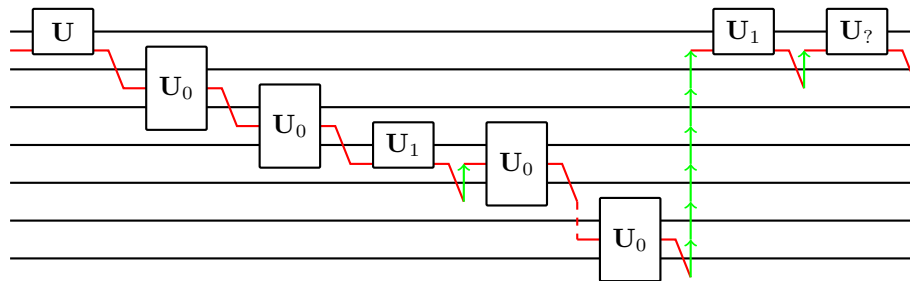


Figure 2.13: Example on how to translate a sample circuit section into a program description, where U_0 , U_1 and $U_?$ can stand for any unitary gate. Starting from the top left, the description here is $1\ 0\ 0\ 1\ \diamond\ 0\ 0\ \diamond\ 0\ 0\ \diamond\ \diamond\ \diamond\ \diamond\ 1\ \diamond\ ?$. The dashed line stands for a normal identity $0\ \diamond\ 0$, as $\text{CNOT}^2 = \mathbb{1}$. The rhombus \diamond is a special symbol that shifts the current gate position up by one; as in each successive step the position moves down by one by default, it suffices to only have this one special shift symbol. The last gate is a special identity to be used to initialise ancillas and penalise a section of the output.

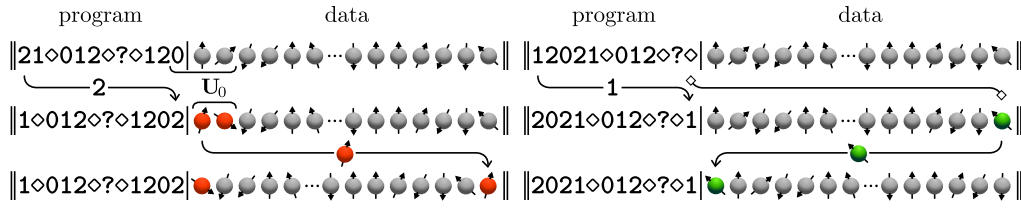


Figure 2.14: The two actions that can be performed by the wheelbarrow construction. On the left, we apply a gate U_0 corresponding to the rightmost program bit 0 . The ring of qubits is then rotated by one, which is the default downwards shift as mentioned in fig. 2.13. On the right, the special action of the \diamond symbol is depicted: it signals the rightmost qubit to move back to the left side. After either action, the program string is rotated by one.

qubit is picked up and moved to the left of the tape. Fig. 2.14 illustrates both operations. The execution runs until the underlying ring machine terminates, which can be determined using a special halt operation H which only proceeds if the tape data is not in a halting configuration.

This also explains the choice of *Turing's Wheelbarrow* as name for this QTS: qubits and program symbols are moving across the tape in two cyclic motions, mimicking a busy worker carrying and depositing information in a wheelbarrow.

2.2.6.2 Notation

For convenience, we define a special notation to describe the construction of Turing's Wheelbarrow. We begin by introducing the alphabet and tape.

Definition 2.62. Let $\Gamma := \Gamma_{cl} \sqcup \Gamma_q$ denote the alphabet consisting of 48 symbols where

$$\Gamma_{cl} := \left\{ \begin{array}{l} \parallel, \blacksquare, \boxtimes, \boxplus, \curvearrowright, \curvearrowleft, \curvearrowright, \curvearrowleft, \blacktriangleright, \blacktriangleleft, \blacktriangleright, \blacktriangleleft, \blacktriangleleft, \blacktriangleright, \\ \mathbb{U}, \mathbb{T}, \mathbb{A}, \mathbb{S}, \mathbb{H}, \diamond, \mathbb{U}, \mathbb{T}, \mathbb{A}, \mathbb{S}, \mathbb{H}, \blacklozenge, \square, \mathbb{U}, \mathbb{T}, \mathbb{A}, \mathbb{S}, \mathbb{H}, \blacklozenge, \blacklozenge, \\ \mathbb{!}, \vec{0}, \vec{1}, \dot{1}, \dot{1}, 0, 1, \mathbb{0}, \mathbb{1} \end{array} \right\},$$

$$\Gamma_q := \left\{ \mathbb{a}, \mathbb{a}, \mathbb{a}, \mathbb{a}, \mathbb{a} \right\}.$$

These two sets correspond to the classical and quantum symbols, respectively, and are of size $|\Gamma_{cl}| = 43$ and $|\Gamma_q| = 5$. The set of head characters is

$$H := \Gamma \setminus \{ \parallel, \blacksquare, 0, 1, \mathbb{U}, \mathbb{T}, \mathbb{A}, \mathbb{S}, \mathbb{H}, \diamond, \square, \blacklozenge, \blacklozenge, \mathbb{a} \},$$

and the boundary characters are $B := \{ \parallel, \boxtimes, \blacksquare \}$.

The number of alphabet characters can be further reduced to 39, 3 of which are quantum, see corollary 2.72.

For reasons of clarity, we use a slightly larger alphabet in this construction.

We will generally use the letters x, y, z as placeholders for program symbols—denoting any of the symbols $\mathbf{U}, \mathbf{T}, \mathbf{A}, \mathbf{S}, \mathbf{H}, \diamond$ as x , or alternatively $\mathbf{U}, \mathbf{T}, \mathbf{A}, \mathbf{S}, \mathbf{H}, \diamond$ as \mathbf{u} , which is always clear from the context. The symbol \mathbf{U} encodes a classically-controlled unitary, \mathbf{T} a TOFFOLI, \mathbf{A} an ancilla, \mathbf{S} a SWAP, \mathbf{H} a halt and \diamond a special tape revert symbol.

We now introduce the notation for transition rules.

Definition 2.63. *We write a transition rule $xy \leftrightarrow zw$ of a quantum Thue system as*

$$\frac{xy}{zw}.$$

The blue shading is used to indicate the location on the tape where the transition rule is applied. Note that, by construction, transition rules are symmetric, i.e.

$$\frac{xy}{zw} \text{ is equivalent to } \frac{zw}{xy}.$$

If the first rule is associated with a non-trivial unitary \mathbf{U} , the inverse rule is associated with the adjoint \mathbf{U}^\dagger .

As in definition 2.45, we never need to write out the values of the qubits anywhere. In fact, the only place where the associated Hilbert space comes in is when we want to apply a quantum gate to the qubits (see section 2.2.6.8). As an example, consider the action of swapping two neighbouring qubits. The Thue system itself does not notice this, e.g. we would have a transition with an explicit comment on the Hilbert space unitary, i.e.

$$\frac{a a}{a a} \text{ where the associated Hilbert spaces are swapped with } \mathbf{U} = \text{SWAP}.$$

To emphasise that the subspaces are in fact swapped, we generally use the letters a, b, c, d to label different quantum subsystems. *This is only to facilitate notation!* In principle, we could stick to the letter a and write out the swap action for every transition rule where this is relevant. But because we believe it is easier to read and most of the non-trivial unitaries that we use are swaps, we simply write

$$\frac{a b}{b a}$$

which is self-explanatory.

2.2.6.3 Transition Rules and History State

The following table contains a list of all transition rules, visualised from a starting string of the form $\| \mathbb{E} a c \cdots a c c \mathbb{G} \|$. The horizontal direction corresponds to space while the time flows from top to bottom. By default, the unitary associated to any rule—if not mentioned otherwise—is the identity. Apart from SWAP operations, the only non-trivial unitary appears in the computation step in section 2.2.6.8.

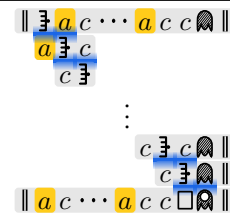
Observe that there are many possible local ambiguities within the history state, which we analyse in detail in section 2.2.6.9.

2.2.6.4 Initialisation

The initialisation is done by moving a special symbol, a “sweeper” \mathbb{E} , from one end of the tape to the other side. This ensures that the tape is actually correctly initialised, since any symbol apart from a , c or the ghost \mathbb{G} would result in a penalised configuration, see table 2.2.

Left hand side has a sweeper \mathbb{E} , right hand side an inactive ghost \mathbb{G} , and all middle symbols are qubits a or data bits $c \in \{0, 1\}$, which are opaque for the ghosts. We let the sweeper move through all middle symbols. This allows a dynamic “initialisation” of the tape: if the sweeper \mathbb{E} bumps into any symbol that is not a qubit, data bit or inactive ghost, we can penalise the configuration, singling out the proper history state.

Once the sweeper reaches the ghost \mathbb{G} at the right boundary, it activates the ghost to \mathbb{G} and transitions to the box \square .



2.2.6.5 Ghost

The ghost symbols act as general “carriage return” symbols: this saves having different return variants for each head symbol used, and is solely a way of saving local dimension. The ghost can thus be seen as a particle to the right side of any other head symbol, and which diffuses freely on the tape (i.e. randomly moves left or right). Only if the ghost is “activated”—i.e. carries a “head flag”—can it interact non-trivially with the symbols around it.

Generally, if there is an extra head symbol on the tape, the ghost is inactive (\mathbb{G} and \mathbb{G}). The ghost can itself carry the head flag, in which case we call it *active* and denote by \mathbb{G} or \mathbb{G} . The white active ghost can either turn itself into a head symbol on the left hand side, or activate the boundary. On the right boundary, it oscillates between white and black. This construction saves us a lot of symbols, since we only ever need to specify special right-moving heads, whereas the left movement of the head state is done generically by the ghost. We will often gloss over inactive ghost transitions and assume the ghost just “moves out of the way” as necessary.

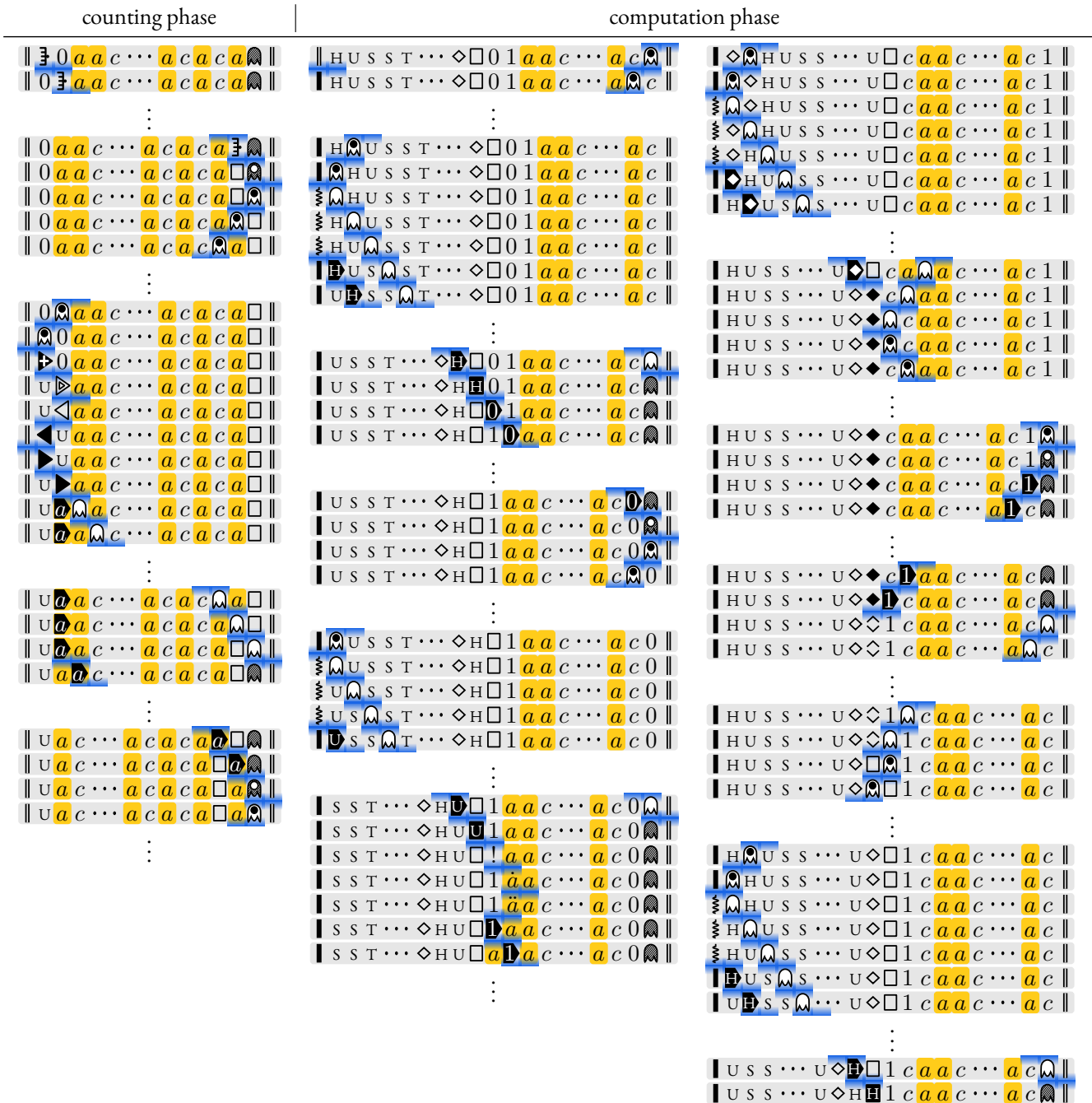




Figure 2.15: Evolution of the history state without branching.

Ghosts can change color on the right boundary. Since the black ghosts  and  are static, any incoming head from the left can detect when it has reached the boundary as it will encounter a black ghost.



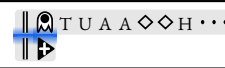
White ghosts can move through all static symbols, but *not* through heads.



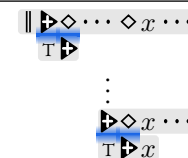
2.2.6.6 Base-6 Counter



From a high-level perspective, the base-6 counter and unary counter (next section) work together to translate the tape length into a base-6 big endian number on the left side of the tape. This base-6 number then encodes the program which we execute afterwards: we count in base 6 through the sequence T, U, A, S, H, \diamond —encoding a TOFFOLI, classically controlled unitary, ancilla, swap, halt or tape revert operation, respectively.

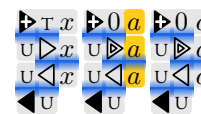
Active ghost  hits the left boundary and turns into the incrementer .





If the incrementer encounters \diamond —the highest-valued digit—it flips it to T —the lowest-valued digit. This results in an overflow that is carried over to the next digit to the right. If the next digit is again \diamond , the same procedure repeats until a different symbol or a qubit is encountered.






If the incrementer  encounters T or classical zero 0 —both of which are treated as lowest-value symbol—it increments it to the next higher-valued symbol U . To uniquely distinguish to which symbol to decrement when run in reverse, the incrementer  has to transition to the checking symbol \triangleright or \triangleright , verifying that the symbol to its right is another counter or tape symbol, respectively. We never encounter the configuration $\triangleright 1$ or $\triangleright a$, as this is penalised.

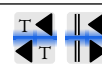



The incrementation ends with the reverter symbol .

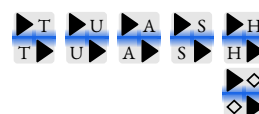
If the incrementer encounters U, A, S or H , it increments the symbol to the next higher one— A, S, H or \diamond , respectively—turning into the reverter symbol .






The reverter  moves through the lowest-valued digits T all the way to the left boundary where it turns into the right mover . Note that no digits other than T are possible to the left of  since incrementation proceeds to the next digit only in case of an overflow.



The right mover  proceeds to the right through all digits of the base-6 counter.



Eventually  encounters a qubit a or classical bit c . It turns into an inactive ghost  and picks up the qubit a or classical bit c . Afterwards the ghost  moves out of the way and we proceed to the unary counter.



For the configuration $\triangleright \square$ there is no forward transition, which means that once we entered the computation phase, this counting does not continue.

2.2.6.7 Unary Counter

The unary counter is necessary so that the base-6 counter knows when to stop. We use a block symbol \square to denote the position of the unary counter on the tape, starting from the right and moving to the left at each increment. Whenever this block is moved left once, the base-6 counter has been incremented by one as well. In this way, once the unary counter has run out of space, we have translated the tape length into a base-6 number on the left side of the tape.

A qubit a or classical bit c is carried to the right past all other qubits or classical bits.	
The position of \square indicates the value of the unary counter. As the qubit a moves through it, the block \square is pushed one position to the left—this increments the unary counter by one.	
Once the moving qubit a reaches the black ghost \blacksquare at the right boundary, the qubit a is dropped and the ghost is activated to \blacksquare .	

2.2.6.8 Computation

The tape now has the form $\| \text{H} x \cdots x \square c a a c \cdots a \blacksquare \|$, i.e. the counting is complete and by our choice of the chain length, the program description starts with a halt symbol H . The rest of the program string does not contain any H 's.

The idea behind the computation is depicted in fig. 2.4. We first take the base 6 symbol from the left end of the program description and move it to the right end (e.g. $x2345$ would become $2345x$). This symbol x can then be picked up by the box \square , which becomes activated to \blacksquare . The active box is now followed by a set of rules which applies this program action to the (qu)bits right next to it. Afterwards, the leftmost (qu)bit is carried to the right end and the procedure repeats.

The content of the tape symbols is checked on the fly using the ancilla program bit symbol \blacksquare . If it appears next to a qubit, a penalty is given for the qubit marginal being $|1\rangle\langle 1|$; for a classical bit, we penalise 0 (this is because we do not have a NOT gate, but NOT can be implemented with TOFFOLI, which maps $111 \mapsto 110$). The implementation details of all the different program bits are explained in the following table.

The computation halts once a halting program bit \blacksquare is next to a classical 1 .

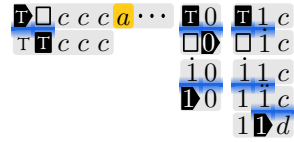
The boundary \parallel is flipped to \llcorner and can only revert to \parallel next to a halt symbol \mathbb{H} —this ensures that we can only transition back and forth between counting and computation if the program bits are in their original order. An active ghost \mathbb{G} can hit this left boundary \llcorner and activate it to $\mathbb{G}\llcorner$.



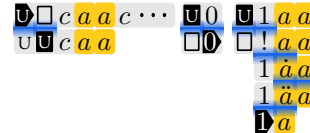
A program bit $x \in \{\mathbb{T}, \mathbb{U}, \mathbb{A}, \mathbb{S}, \mathbb{H}, \diamond\}$ is picked up as $\mathbb{G}x$ and carried to the right of the program string.



These transition rules apply the TOFFOLI gate to three classical bits $c c c$. If either $0 c c$ or $1 0 c$, the last bit remains unchanged. Only for the configuration $1 1 c$ we perform a bit flip on the last bit, i.e. $d = \neg c$. The first bit is then picked up with a carrier \mathbb{D} or \mathbb{D} .



These transition rules apply the classically-controlled unitary operation to a pair of qubits $a a$, but only for the configuration $1 a a$. The control bit is then picked up with a carrier \mathbb{D} or \mathbb{D} . Observe that this is the *only* position where we apply a unitary operation to the quantum symbols.



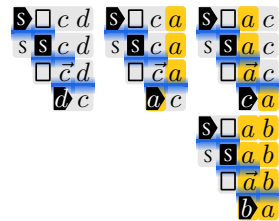
$\mathbb{A} a$ acts as identity on the qubit, but is used later on to penalise when the attached Hilbert space is $|1\rangle$, giving us the possibility for ancillas.



$\mathbb{A} 1$ acts as identity, and we will penalise $\mathbb{A} 0$, giving us a classical ancilla bit. Observe that we choose $\mathbb{1}$ here, as we can create 0 s out of nothing but 1 s with the TOFFOLI gate, but not vice versa.



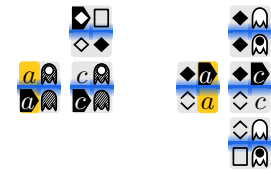
\mathbb{S} implements the SWAP gate.



\mathbb{H} acts as identity on 0 , but has no forward transition for $\mathbb{H} 1$, i.e. the operation explicitly halts the computation.



\blacklozenge implements a tape revert, i.e. moving the current tape position up by one. \blacklozenge acts like an activated boundary on the left hand side, i.e. it blocks ghosts \mathbb{G} or \mathbb{G} . An incoming ghost from the right can activate from \mathbb{G} to \mathbb{G} , after which it proceeds back to the right end. Outlined in section 2.2.6.7 but run backwards, the activated ghost will move through to the right hand side and deactivate at the boundary, while picking up a qubit \mathbb{G} . This qubit will move to the left until it encounters \blacklozenge . It drops the qubit and deactivates \blacklozenge to \diamond . As soon as the inactive ghost \mathbb{G} encounters this symbol, the ghost is reactivated and the box \square is restored.



Definition 2.64 (Turing's Wheelbarrow). Turing's Wheelbarrow is the QTS $(\Gamma, R, \{\mathbb{U}_\tau\}_{\tau \in R}, \mathbb{C}^2)$, where Γ is given in definition 2.62 and the relation R is defined by the transition rules in section 2.2.6.3 (with the conventions on notation from definition 2.63).

One can verify that Turing's Wheelbarrow, when applied to an initial string of the form $\parallel \mathbb{G} a c c a \dots c a \mathbb{G} \parallel$, where the sequence of c s and a s is such that they match the counting and computation phase, first translates

the string length into a program description on the left string side, which is then executed cyclically on the tape. We call an initial configuration of this type *valid initial configuration*.

There are, however, ambiguous transitions, which lead to branching in the graph—we discuss all possible branching points for the irreducible evolution containing this initial configuration.

2.2.6.9 Branching in the History State

We make extensive use of branching and ambiguous transitions to compress the number of symbols necessary to implement the Wheelbarrow. Therefore we need to show two things.

1. The size of the history state is poly-bounded.
2. There are no ambiguous transitions which lead to a penalised configuration.

We take fig. 2.15 as a point of reference.

Ghosts. Whenever there is a ghost on the tape, it can either be active— Ⓜ or Ⓜ , or inactive— Ⓜ or Ⓜ .

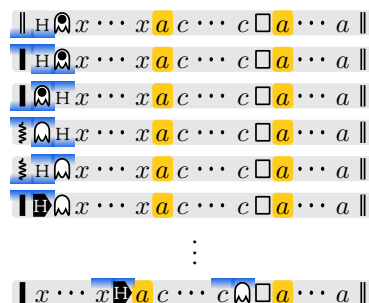
Fact 2.65. *Inactive ghosts never change non-head symbols or pass through heads.*

Therefore we will disregard any branching due to inactive ghosts, which happens because we can always move either the head or ghost at each step. This increases the history state size by an at most quadratic factor.

Counting Phase.

Fact 2.66. *Initialisation and 6-ary counter are not ambiguous in either direction.*

We have two ambiguities to analyse. Starting from an intermediary counting stage where the program string starts with H , we can prematurely transition to the computation phase:



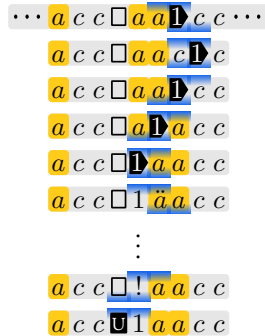
As there is no forward transition for H a , this branch is just a leg, increasing the history state by a small

constant factor ≤ 2 . The same argument holds for transitioning to \blacksquare during incrementation, i.e.

$$\frac{\parallel Hx \triangleright x \cdots x a \cdots c \square a \cdots a \parallel}{\parallel Hx \triangleright x \cdots x a \cdots c \square a \cdots a \parallel}$$

Run forward, there is no transition for $\blacksquare \blacktriangleleft$, and run backwards there is none for $\blacksquare \blacktriangleright$.

Running the carrier \blacktriangleright or \blacktriangleleft backwards for the unary counter at any point before counting is completed leads to another ambiguity, e.g.

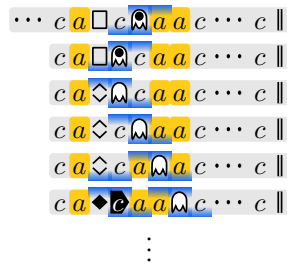


There is no backwards transition for $c \blacktriangleleft$ or $a \blacktriangleleft$ though. If there is no box \square , the branch dies off even before that. This ambiguity hence increases the history state size by another small constant factor.

Computation Phase. A similar argument as in the last section shows that a late transition into the counting phase once we are in the middle of the computation does not proceed, as there is no forward transition for a configuration $\blacktriangleright \square$. Furthermore, the same ambiguity running a carrier \blacktriangleleft or \blacktriangleright backwards holds, which we have already discussed.

Fact 2.67. *The application of gates \mathbb{H} , \mathbb{U} , \mathbb{T} , \mathbb{A} and \mathbb{S} does not introduce any branching.*

It remains to analyse the revert command, where we have a branching point for a configuration



Observe, however, that all that could happen is that the tape symbol is carried to the right, where it is dropped next to the boundary ghost \blacktriangleleft . The ghost is activated and moves back to $\blacktriangleright \blacktriangleleft$, where it deactivates.

The branch does not continue further, as there is no transition out of $a \blacklozenge$ or $c \blacklozenge$. This increases the history state size by some small constant ≤ 2 .

We define the set of tuples \mathfrak{A} as all the possible character pairs that appear in this history state—including all branches—in table 2.2.

This exhaustive analysis of all possible branching points in the history state allows us to conclude the following corollary.

Corollary 2.68. For strings of length n , the size of the irreducible evolution containing a valid initial configuration of the form $\| \exists a c c a \cdots c a \|$ —the history state—is of size $O(n^3)$, and contains no forbidden character pairs.

2.2.6.10 Simplicity of Turing’s Wheelbarrow

Let us briefly recall the idea behind simplicity in the context of QTSs. A QTS is called simple, if, for any two strings connected by more than one chain of transitions, the product of unitaries along this chain is identical. Equivalently, we can show that there are no loops in the graph connecting any strings. Regarding the QTS transition rules for Turing’s Wheelbarrow, as constructed in the last section, it is easy to see that it will not be simple. However, for our purposes, it suffices to prove the following lemma.

Lemma 2.69. Each bracketed string in Turing’s Wheelbarrow with at least one head either belongs to the history state, which is simple, or—by removing edges—can be broken up into poly n -sized valid evolutions with illegal pairs.

Proof. As no transition rule ever changes the number of heads or position of brackets, the distinction is well-defined. We can analyse each separately.

One head. We can exclude strings with illegal pairs right away. Furthermore, we can disregard configurations of non-head characters which are just allowed because there is a head symbol or ghost between them, such as $a \blacklozenge c$, as moving the head either way (which is possible, since there is only one of them) transitions to an illegal pair.

So, disregarding any head and ghost state on the string for now, the most general non-head-non-boundary string compatible with table 2.2 is

$$\underbrace{x^*}_A \blacklozenge^? \underbrace{(c \square | a \square | c | a)^+}_B .$$

It is straightforward to see that evolving this configuration backwards will transition to an illegal pair, if either

	non-heads										heads																							
			<i>x</i>	□	◇	◆	<i>c</i>	<i>a</i>	⊖	⊗	⊙	⊚	⊛	⊜	⊝	⊞	⊟	⊠	⊡	⊢	⊣	⊤	⊥	!	<i>c</i>	<i>i</i>	<i>ī</i>	<i>e</i>	<i>ā</i>	<i>ä</i>	<i>ã</i>	<i>ä</i>		
			✓				✓	✓			✓		✓		✓	✓																		
			✓										✓		✓	✓																		
<i>x</i>			✓	✓	✓	✓	✓	✓	✓	✓			✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	
□	✓						✓	✓	✓	✓	✓	✓												✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	
◇							✓	✓	✓	✓	✓	✓												✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	
◆							✓	✓	✓	✓	✓	✓												✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	
<i>c</i>	✓			✓	✓	✓	✓	✓	✓	✓	✓	✓		✓									✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	
<i>a</i>	✓			✓	✓	✓	✓	✓	✓	✓	✓	✓		✓									✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	
⊖	✓	✓	✓				✓	✓	✓																									
⊗	✓						✓	✓	✓																									
⊙	✓						✓	✓	✓																									
⊚	✓						✓	✓	✓																									
⊛	✓						✓	✓	✓																									
⊜			✓				✓	✓	✓	✓																								
⊝							✓	✓	✓	✓																								
⊞							✓	✓	✓	✓																								
⊟							✓*	✓	✓	✓																								
⊠							✓	✓	✓	✓																								
⊡							✓	✓	✓	✓																								
⊢							✓	✓	✓	✓																								
⊣							✓	✓	✓	✓																								
⊤							✓	✓	✓	✓																								
⊥							✓	✓	✓	✓																								
!							✓	✓	✓	✓																								
<i>c</i>							✓	✓	✓	✓																								
<i>i</i>							✓	✓	✓	✓																								
<i>ī</i>							✓	✓	✓	✓																								
<i>e</i>			✓				✓	✓	✓	✓	✓																							
<i>ā</i>							✓	✓	✓	✓																								
<i>ä</i>							✓	✓	✓	✓																								
<i>ã</i>							✓	✓	✓	✓																								
<i>ä</i>			✓				✓	✓	✓	✓																								

Table 2.2: All possible character tuples occurring in the history state of Turing’s Wheelbarrow. The row is the first character, the column the second—e.g. || ⊖ is allowed, whereas *x* ⊞ is not. *c* can be 0, 1, and *x* stands for any program bit U, T, A, S, H, or ◇. ✓† only allows the combination T ⊞, U ⊞, U ⊞, U ⊞ and U ⊞. ✓* only allows the combination allowed by the gates, i.e. T *c*, U *c*, A *a*, A 1, S *c*, S *a* and H 0. Observe how the lower right block is completely empty, as there can only ever be one head on the tape.

- B has multiple \square s, or \diamond and at least one \square , or neither \diamond or \square .
- A does not match the string length $\log_6 n$. Decrementation can only start if the substring A starts with the halt symbol \mathbb{H} , so it cannot happen that we start decrementing a rotated number, e.g. UHAAST instead of HAASTU , which would translate into different lengths.

Evolving this string forward then reaches the computation part, and in case the pattern of classical and qubit states in B does not match the one required for the encoded gates in A we again have an illegal pair.

We are left with the history state, and it suffices to check any transition rule containing a non-trivial unitary attached, which by construction is the computational step only, i.e.

$$\frac{\dot{a}a}{\ddot{a}a}.$$

Following the transitions forward to the next such transition, by construction, the encoded Turing machine evolution is reversible, hence there is no loop as the Turing machine changes the classical content of the tape in section B .

Multiple heads. None of the heads can pass through each other. As further boundary markers such as \mathbb{I} , \mathbb{J} and \mathbb{K} are immobile and opaque and there exists no transition out of \mathbb{M} or \mathbb{N} if not left of a boundary, we can without loss of generality assume that the tape is bracketed by either of \mathbb{I} , \mathbb{K} , \mathbb{M} , \mathbb{N} , or possibly no opaque symbol if our subsection lies at the tape ends.

If there are $h \geq 2$ head symbols on the tape, a simple argument allows us to slice the graph up into $\text{poly}(j(l))$ -sized parts: first observe that following any of the heads—with potential intermediate transitions—sweeps the entire width of the string. For any configuration of the first $h - 1$ heads, the last head will thus necessarily bump into the $h - 1^{\text{st}}$ within $O(j(l))$ steps. The same argument shows that there can be at most one ghost on the tape. \square

2.2.6.11 Special Properties

Proof of lemma 2.59. We will check the properties of lemma 2.59 one-by-one.

The deciding property follows by construction. The projectors $\Pi_{\text{inp}} = \Pi_{\text{out}} = |1\rangle\langle 1|$ are supposed to act on checking the first ancilla and output as seen in fig. 2.12, i.e. they apply to the qubit after the special identity symbol \mathbb{I} .

Item W_1 is readily verified.

Item W_2 . The encoding is given by the valid initial configuration

$$\text{enc}(l) := \mathbb{I} \mathbb{K} \underbrace{a c a c c \cdots a}_{N \text{ times}} \mathbb{M} \mathbb{N},$$

where N is a unary encoding of the QRM head circuit executing fig. 2.12 rewritten as depicted in fig. 2.13, and the sequence of \mathbf{a} s and \mathbf{c} s is such that they match the counting and computation phase. By construction, we therefore obtain, $N + 4 = \text{poly}(|l|) =: j(l)$.

Also by construction and as outlined in section 2.2.6.1, the program string on the left side of `enc` describes the head of a QRM writing out the circuit fig. 2.12. This QRM is in the same uniformity class as the original verifier's, and a constant in the size of $l \in \Pi$. We can hence pad it—using identity gates—to get the space and runtime for the QRM right, which can be as large as $\text{poly } j(l)$, as required for a BQEXP computation.

Both input and output markers $s_{\text{inp}} = s_{\text{out}} = \mathbf{A}\mathbf{a}$ are 2-local, contain one head \mathbf{A} , and $\Pi_{\text{inp}} = \Pi_{\text{out}} = |1\rangle\langle 1|$.

Item \mathbb{W}_3 We have shown the first claim in lemma 2.69 and corollary 2.68. The rest follows by direct verification.

Item \mathbb{W}_4 We can immediately sort out the no-head and not-bracketed cases. The rest follows from lemma 2.69.

This concludes the proof. □

2.2.7 Final Dimension Reduction

We want to make a few final remarks, and suggest an immediate optimisation of the Wheelbarrow construction.

The distinction between the quantum and classical tape symbols \mathbf{c} and \mathbf{a} is unnecessary, if we can ensure that there is never a quantum operation on classical symbols and vice versa. This is already proven.

The reason why we can merge these symbols is that while the QTS requires the ULG vertices to comprise only the classical alphabet symbols, we do not need to make this distinction for a ULG—as long as we can ensure that the Hilbert space dimension on each vertex in a connected component is the same. It is also clear that this does not break simplicity in lemma 2.69, as we always *know* which tape symbols are classical (the ones appearing next to classical operations, e.g. \mathbf{I}) and which ones are quantum (e.g. the one next to $\mathbf{!}$). This observation allows the following optimisation.

Remark 2.70. The Wheelbarrow construction works exactly the same when merging \mathbf{a} with \mathbf{c} , \mathbf{a} with \mathbf{c} , and $\bar{\mathbf{a}}$ with $\bar{\mathbf{c}}$.

Once we have merged the symbols, there is another merge possible. We know that TOFFOLI and some basis-changing unitary \mathbf{U} are quantum-universal, see e.g. [NC10, ch. 4.5]. This means that we can replace the classically-controlled unitary with such a one-qubit unitary, and apply TOFFOLI gates to quantum symbols as well. A similar argument as before shows that this does not break lemma 2.69, and we phrase the following remark.

Remark 2.71. *The Wheelbarrow construction works when replacing the controlled unitary with a single-qubit basis-changing unitary, and extending TOFFOLI to work on classical and quantum tape content. This makes the symbols \uparrow , \hat{a} and \hat{a} obsolete.*

Including the saved symbols from the last two remarks— c , \bar{c} , \bar{c} , \uparrow , \hat{a} and \hat{a} —we conclude with the existence proof of lemma 2.59.

Corollary 2.72. *There exists a family of simple QTSs with 2-local rules on an alphabet of size $39-3$ of which are quantum with a Hilbert space \mathbb{C}^2 —and all properties given in lemma 2.59.*

Remark 2.73. *It is straightforward to get $O(1)$ -interactions, i.e. removing the scaling polynomial $p(l)$ in theorem 2.60 if we can locally distinguish the history state at all times. This is possible e.g. by using distinct non-head symbols on the left and right hand side of the head and penalising invalid configuration using regular expressions as in [GI13]. This would increase our dimension by roughly 15.*

2.3 Chapter Summary

This work was motivated by the idea of finding a simple, translationally-invariant and physically interesting system, for which the ground state energy problem is QMA_{EXP} -hard. In [GI13], Gottesman and Irani concluded that their construction is not “particularly natural”, due to the large local dimension necessary, but that the existence of some very simple QMA_{EXP} -hard LOCAL HAMILTONIAN problem variants seems quite possible.

Our results bring us another step closer to this goal: we reprove the hardness result in [GI13] but with a local dimension of 42, whereas in [GI13]—though not explicitly specified—it was several orders of magnitude larger. To prove this result, we develop new tools and computational models which we believe are applicable to a wider range of problems.

At this point it would be interesting to see where the threshold for the translationally-invariant LOCAL HAMILTONIAN problem lies: does there exist a local dimension d_{\min} , for which the problem is in BQP, or BQEXP? We have shown that $d_{\min} < 42$, but do not believe this to be a strict bound. We therefore encourage the interested reader to construct their own version of the Wheelbarrow, which might yield an even lower local dimension, and thus tighten our bound.

Furthermore, a lot of work recently has been done to analyse non-translationally-invariant systems, and to classify interactions with locally-varying interaction strengths, e.g. [CM14]; [PM15]. In contrast to our construction, the hardness results in [PM15] resemble more a tiling construction, a subject also addressed in [GI13]. It would be an interesting approach to see if these two—fundamentally quite different—results can be combined, or if there exists yet another, completely different, method of encoding computation into the ground state of a local Hamiltonian.

Finally, we want to mention that while the research focus—as outlined in table 2.1—quickly shifted towards the 1D variant of the problem, from a physical perspective both 2D and 3D versions of this result are still of great interest. In the next chapter, we will have a look at the 3D case, and try to embed a different QMA-hard quantum Turing system into a crystal lattice; we want to emphasise that we are going beyond a simple embedding of the 1D result in this chapter. As we discuss in chapter 3, we reduce the interaction complexity significantly even beyond Turing’s Wheelbarrow by making use of the three-dimensional lattice geometry, and by “outsourcing” part of the computational overhead away from the history state, and to a static tiling problem. •

3 Hamiltonian Complexity:

Lattice Crystals

Gerade in diesem Punkt irrt die ganze Welt.

Wir entfernen uns ständig
von dem gegenwärtigen Augenblick.

—H.G. Wells, *Die Zeitmaschine*/fr-036

As we have seen in chapter 2, one major issue of QMA-hard instances of the LOCAL HAMILTONIAN problem is that, from the perspective of experimental physics and material sciences, the resulting many-body quantum systems are too contrived to be of relevance; either the local spin dimension is large, the coupling strengths vary from site to site, or the interaction graphs are not geometrically local.

While 1D results are interesting and in a sense the most fundamental models to study (as any 1D hardness result directly implies hardness of the corresponding higher-dimensional constructions by a straightforward embedding argument), most condensed matter systems are in fact two- or three-dimensional, and the comparison of local dimension between the best non-translationally invariant results in 1D and 2D — 8 [HNN13], and 2 [OT05], respectively—indicates that moving beyond 1D allows a significant reduction of the lattice spins’ dimension. It is thus a natural question to ask whether one can go beyond a simple reduction from previously-known 1D results, by exploiting these extra dimensions in a non-trivial way (i.e. beyond a simple embedding), but at the same time retaining nice physical properties such as a regular lattice structure and translational symmetries. We can even go further: is there a family of Hamiltonians on a physically realistic 3D crystal lattice with a QMA-hard ground state? This question is highly relevant, since such crystal structures are found ubiquitously in nature (e.g. face-centred cubic lattices for sodium chloride, or body-centred cubic cesium chloride crystals).

In this chapter, we prove that the LOCAL HAMILTONIAN problem remains computationally hard, even for a face-centred cubic lattice of spin-3/2 particles with geometrically 4-local translationally-invariant interactions, and open boundary conditions.

It is clear that there is always a trade-off between local dimension and interaction range: a Hermitian operator coupling k spins of dimension d each has d^{2k} real degrees of freedom. In 1D and for 2-local

interactions, the best-known construction to date is [HNN13] with 8-dimensional qudits and nearest-neighbour interactions; for each coupled pair of qudits, one Hermitian operator thus has $8^2 \times 8^2 = 16384$ free real parameters. Enforcing translational invariance, we can take Turing’s Wheelbarrow which we introduced in the last chapter—nearest-neighbour interactions between spins of dimension ≈ 50 —which would give roughly $(50^2)^2 \approx 6 \times 10^6$ parameters to choose from.

The construction in this chapter with at most 4-local interactions between spins of dimension 4 yields 4^8 degrees of freedom, a roughly two orders-of-magnitude improvement over a straightforward embedding of the best one-dimensional construction, and en par with the best non-translationally-invariant result. It also shows that there is only about three orders of magnitude left between this construction and spin systems that we encounter every day (e.g. nearest-neighbour, spin 1).

- As in chapter 2, we work with Hermitian operators acting on a multipartite Hilbert space $\mathcal{H} = (\mathbb{C}^d)^{\otimes n}$, i.e. on n qudits, each of *local dimension* d . We label subsystems of \mathcal{H} by an ordered tuple $A \subseteq \{1, \dots, n\}$. For a k -qudit Hamiltonian \mathbf{h} for some $k \leq n$ and some subset A , we denote with \mathbf{h}_A the operator that acts as \mathbf{h} on all qudits labelled by A , and as identity— $\mathbb{1}$ —everywhere else.
- If the Hilbert space \mathcal{H} is translationally-invariant—e.g. a lattice $\mathcal{H}^{\otimes \Lambda}$ —then we say that a Hamiltonian on this space exhibits translational invariance if it follows the same symmetry, i.e. that the interactions between equivalent lattice sites are identical. This allows us to define the following variant of the local Hamiltonian problem, where we follow the same naming convention as in definition 2.16.

Definition 3.1 ((k, d)-TILH-3D). Let $\Lambda(L, M, N)$ be a 3D lattice of side lengths L, M and N , all $\leq n$, with not necessarily trivial unit cell (e.g. cF, cI). Let $\mathbf{H} = \sum_{i,j,k} \mathbf{h}_{i,j,k}$ be a 3D translationally-invariant and geometrically k -local Hamiltonian on the lattice qudits $(\mathbb{C}^d)^\Lambda$.

Input. Specification of the lattice size L, M, N , and the matrix entries of \mathbf{h} , up to $O(\text{poly } n)$ bits of precision.

Promise. The operator norm of each local term is bounded, $\|\mathbf{h}\| \leq \text{poly } n$ and either $\lambda_{\min}(\mathbf{H}) \leq \alpha$ or $\lambda_{\min}(\mathbf{H}) \geq \beta$, where $\lambda_{\min}(\mathbf{H})$ denotes the smallest eigenvalue of \mathbf{H} and $\beta - \alpha \leq 1/p(n)$ for some polynomial $p(n)$.

Output. YES if $\lambda_{\min}(\mathbf{H}) \leq \alpha$, otherwise NO.

3.1 Overview of Results

The family of spin systems we study are described by a Hamiltonian on a face-centred cubic (cF) lattice as shown in fig. 3.1. More precisely, we start with a finite cubic lattice Λ , where each vertex *and each face* carries a 4-dimensional spin $\mathcal{H}_{\text{loc}} = \mathbb{C}^4$; the overall Hilbert space \mathcal{H} is then the tensor product of all spins. For a

geometrically local Hamiltonian \mathbf{h} acting on k neighbouring spins (on vertices, faces, or both), we denote with $\mathbf{h}^{\vec{x}}$ the k -local operator \mathbf{h} when offset by a lattice vector $\vec{x} \in \Lambda$, and acting trivially everywhere else; in case that $\mathbf{h}^{\vec{x}}$ protrudes out of Λ , we set $\mathbf{h}^{\vec{x}} \equiv \mathbf{0}$. For a finite index set I , we consider Hamiltonians of the form

$$\mathbf{H} = \sum_{i \in I} \left(c_i \sum_{\vec{x} \in \Lambda} \mathbf{h}_i^{\vec{x}} \right), \quad (3.1)$$

where each $\mathbf{h}_i^{\vec{x}}$ couples at most 4 spins, either within a single unit cell, or between neighbouring unit cells. By construction, this Hamiltonian is translationally-invariant, and features open boundary conditions since we do not place special interactions at faces, edges or corners of the lattice cuboid.

The index set I does not depend on the size of the lattice, and neither do any of the \mathbf{h}_i ; we allow the $c_i = c_i(|\Lambda|)$ to depend on the system size $|\Lambda| = W \times H \times D$, but require any $c_i/c_j \in [\Omega(1/\text{poly } |\Lambda|), O(\text{poly } |\Lambda|)]$. This allows us to define a variant of the LOCAL HAMILTONIAN problem where the input is given by a description of the local terms of a Hamiltonian as in eq. (3.1) (i.e. the matrix entries of the local terms $c_i \times \mathbf{h}_i$, up to polynomial precision), as well as the three side-lengths W , H and D of the lattice. Moreover, we are given two parameters $\alpha < \beta$ satisfying $\beta - \alpha = \Omega(1/\text{poly } |\Lambda|)$, and a promise that the ground state energy of \mathbf{H} is either smaller than α , or larger than β . The LOCAL HAMILTONIAN problem is then precisely the question of distinguishing between these two cases, and we prove the following main theorem.

Theorem 3.2. *The LOCAL HAMILTONIAN problem is QMA_{EXP} -complete, even for translationally-invariant 4-local interactions on a 3D face-centred cubic spin lattice (fig. 3.1) with local dimension 4, and open boundary conditions.*

We give a rigorous proof of theorem 3.2 in section 3.3; in the following, we want to give a high-level exposition of the ideas and proof techniques which we employ. As in past hardness results, we present an explicit construction of a family of QMA_{EXP} -hard instances of this variant of the LOCAL HAMILTONIAN problem. We will make use of two types of local terms, tiling and history state Hamiltonians, both of which have been studied extensively, but mostly independently of each other. In our work, we will utilise each method to its strength: the classical tiling terms will be used to encode the bootstrapping mechanism responsible for the large local dimension in prior work, while the history state terms will be used as a means of embedding the quantum computation part. First we will briefly recap these methods.

3.1.1 History State Construction

We want to give a brief summary of the more rigorous construction techniques in chapter 2 that we build on, and refer the reader to the relevant sections 2.2.1.2, 2.2.2 and 2.2.4 for more details. By definition, a promise problem Π is in QMA_{EXP} if there exists a BQEXP quantum circuit—called “verifier”—such that for any YES-instance $l \in \Pi$, there exists a poly-sized quantum state—called “witness”—which the verifier accepts

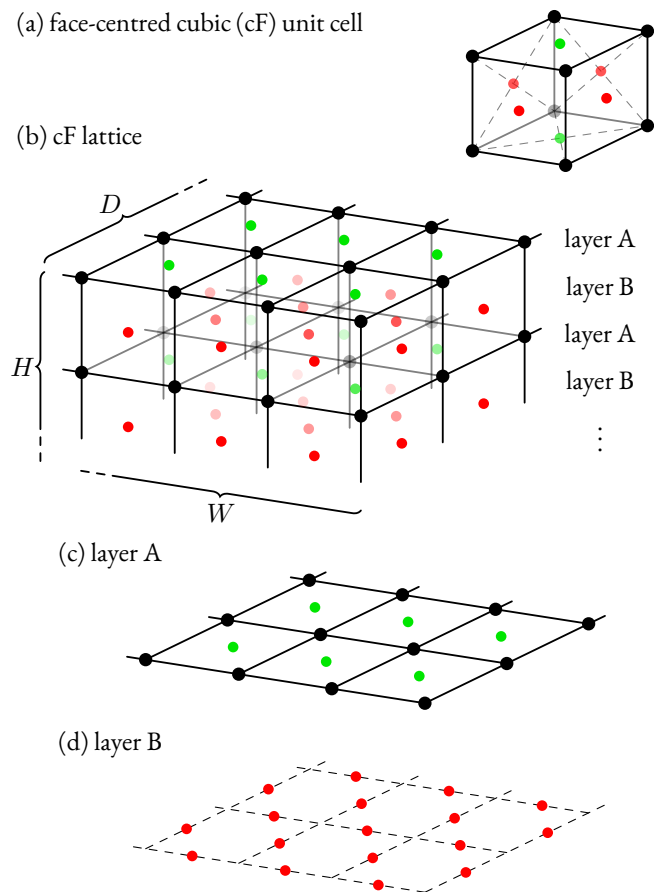


Figure 3.1: Face-centred cubic crystal lattice. All vertices and faces carry spin-3/2 particles; the red and green sublattice spins sit on the faces defined by the black lattice.

with probability $\geq 2/3$; or if l is a **NO**-instance, all poly-sized witnesses are rejected with high probability. The exact constant used here is not important, as for any polynomial p , one can always amplify a QMA_{EXP} promise problem such that the distinction works with probability $\geq 1 - 1/3^{p(|l|)}$ for an instance $l \in \Pi$ with size $|l|$ (see fact 2.12).

We further know that for any QMA_{EXP} promise problem, we can alternatively obtain a so-called *quantum ring machine* (QRM) as verifier (corollary 2.34). In brief, a QRM is a fixed unitary \mathbf{R} on $(\mathbb{C}^d)^{\otimes 2}$, which acts cyclicly on a ring of n dimension d qudits. Borrowing terminology from Turing machines (TM)—which are used to prove universality of the QRM model—we call the unitary \mathbf{R} the *head* of the QRM, and the qudit ring is essentially a TM tape with cyclic boundary conditions. Fig. 2.6 depicts such a QRM and its action in circuit notation.

We take a specific 2-qubit quantum gate \mathbf{G} and prove it to be universal, even when only applied to adjacent qubits (lemma 3.3). Together with its inverse \mathbf{G}^\dagger , we can thus use Solovay-Kitaev to approximate the QRM head unitary \mathbf{R} to within precision ϵ . Since we require that the QRM first writes out an instance $l \in \Pi$ on the ring, the resulting circuit $C_{\mathbf{R}}$ has size $|C_{\mathbf{R}}| = O(\text{poly } |l| \log^4(1/\epsilon))$, see lemma 3.5. To match the QRM evolution, we repeatedly apply $C_{\mathbf{R}}$ in a cyclic fashion, as described in fig. 2.6.

Keeping with tradition, we encode the circuit $C_{\mathbf{R}}$ as history state Hamiltonian. In its simplest form, such a Hamiltonian encodes transitions for each gate \mathbf{U}_i present in $C_{\mathbf{R}} = \mathbf{U}_T \cdots \mathbf{U}_1$; as in the last chapter, we will define a QTS which can execute said circuit, and the resulting ULG Laplacian will be the History state Hamiltonian, see definition 2.38.

In the next section, we explain how we use diagonal Hamiltonian terms to constrain the ground space of our Hamiltonian such that the circuit description for $C_{\mathbf{R}}$ is exposed at the front edge of the cuboid, in a periodically repeating fashion (see fig. 3.2). More precisely, we define a diagonal Hamiltonian \mathbf{H}_{cl} with spectral gap 1, and a degenerate ground space for which any ground state of $\mathbf{H}_{\text{cl}} + \mathbf{H}_{\text{PROP}}$ will then be in a product configuration $|\Phi_{\text{cl}}\rangle \otimes |\Psi\rangle$. Here $|\Phi_{\text{cl}}\rangle$ is a classical product state that takes a configuration as in fig. 3.2: in particular a string describing $C_{\mathbf{R}}$ is expressed, periodically, on the front edge.

This allows us to write local QTS rewriting rules, that can then be used to access this circuit description *without* any explicit knowledge of the current position within the circuit, which is implicitly given by the location on the cube where the transition rule is applied.

3.1.2 Tiling Construction

A tiling Hamiltonian is a local Hamiltonian on a lattice, where each term is a projector onto the complement of the allowed tiles at a specific lattice location; we will re-visit tiling Hamiltonians more rigorously in chapter 4, but want to give a quick overview here. As a simple example, consider just the 2D layer B-type sublattice from fig. 3.1, and assume that every spin is a qubit. We denote with white the state $|0\rangle$, and with

red shading the state $|1\rangle$. Assume the only tiles we want to allow are the four shown in fig. 3.2 (without rotated variants).

By writing a local term for each tile (where we order the corresponding Hilbert space $(\mathbb{C}^2)^{\otimes 4}$ as a tensor product of the spin on the back, right, front, and left, respectively), we can write a diagonal projector $\mathbf{h} = \mathbb{1} - \sum_{i=1}^4 \mathbf{h}_i$ on $(\mathbb{C}^2)^{\otimes 4}$ such that the ground space is spanned by quantum states corresponding to the valid tiles; as an example, for the fourth tile, we write

$$\mathbf{h}_4 := |1\rangle\langle 1| \otimes |1\rangle\langle 1| \otimes |0\rangle\langle 0| \otimes |1\rangle\langle 1|.$$

We can thus easily define a local Hamiltonian on the layer B-type sublattice which in its zero energy ground state encodes *valid* tiling patterns, where adjacent edges match, if possible; if not, the ground state energy of the Hamiltonian will be at least 1. More specifically, if P indexes all squares with four adjacent spins, then we can write the Hamiltonian as

$$\mathbf{H}_{\text{tiling}} := \sum_{\vec{p} \in P} \left(\mathbb{1}_{\vec{p}} - \sum_{i=1}^4 \mathbf{h}_i^{\vec{p}} \right) \otimes \mathbb{1} \quad (3.2)$$

where $\mathbf{h}_i^{\vec{p}} \otimes \mathbb{1}$ acts non-trivially only on the spins sitting on the edges of square \vec{p} . For the aforementioned tiles, the resulting pattern is a binary counter, which can be used to translate the depth of a lattice, D , into a binary string representation of D at the front edge (see top face in fig. 3.2).

The same method can equivalently be used to enforce a more complicated tiling pattern in three dimensions, especially when mixing penalty terms with different weights; for an extensive proof that the corresponding Hamiltonian ground space is indeed spanned by the best possible tiling we refer the reader to lemma 4.1.

3.1.3 Hard Instances for the LOCAL HAMILTONIAN Problem

We will now explain how these two techniques—tiling and history state Hamiltonians—can be combined in order to prove theorem 3.2. As a first step, we define a tiling pattern to constrain all red layer B spins of the cube—apart from the top and side layers, but including the bottom layer—to a specific symbol which is used nowhere else, and which we denote with \times . All the following terms can then be conditioned on these red spins being either in state \times , or not; this allows us to distinguish between the different faces of the cuboid in a translationally-invariant way and with open boundary conditions. This technique is commonly used in 1D (e.g. [GI13]), and we extend it to three dimensions.

- As explained in the last section, we then define four tiles which self-assemble to a binary counter; this allows us to translate the depth of the cube D to a string representation of D on the top front edge. Using

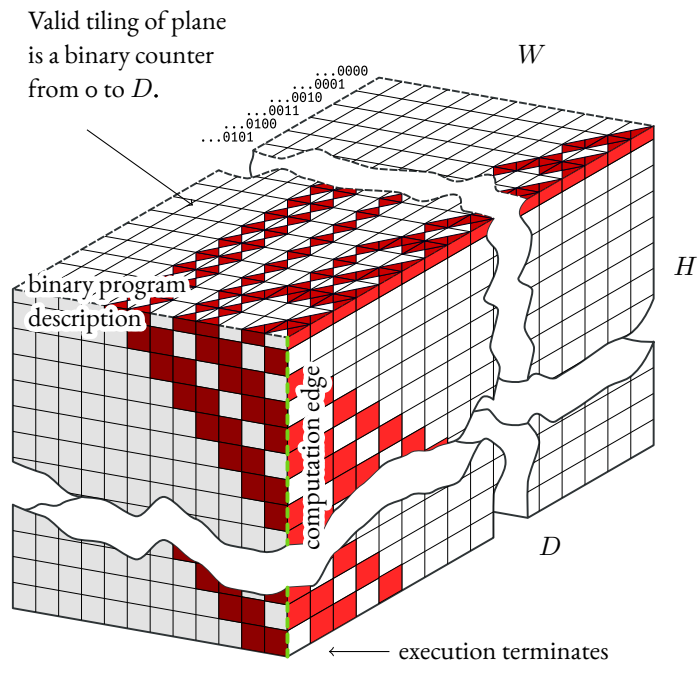
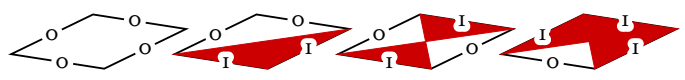


Figure 3.2: Structure of the ground state imposed by classical bonus and penalty terms. Shown here is the lattice as in fig. 3.1; a cut through the top layer of the cuboid shows the layer B red sublattice depicted in fig. 1 (d). Coloured triangles on the top layer denote a spin on the tile edge in configuration $|1\rangle$, a white tile edge stands for configuration $|0\rangle$; the tiles used on the top layer are the following four:



The same colour coding is used for the squares around the sides, which label the red cF spins on the sides of the unit cells. The dashed green front edge denotes the computation edge, where gates will be applied in the history state construction. Observe how the same binary pattern is repeated periodically along the computation edge.

similar tiles on the sides of the lattice, we wind this binary string down and around the cube in an anti-clockwise direction; like that, the string—which is the binary program description of the QRM circuit $C_{\mathbf{R}}$ —is expressed periodically on the front edge of the lattice, which we label the *computation edge*, see fig. 3.2.

We further restrict the spins in the green layer A sublattice adjacent to this computation edge to be in a state corresponding to successive pairs of program bits. For example, if the binary program description is p_1, p_2, p_3, p_4 , the green spins depend on $(0, p_1), (p_1, p_2), (p_2, p_3), (p_3, p_4)$ and $(p_4, 0)$ respectively. A special encoding (see table 3.2) allows us to translate any such binary pair into an operation to perform on the computation edge. All constraints up to this point are diagonal in the computational basis and at most 4-local; we collect all these *static* terms on the cF lattice in the Hamiltonian \mathbf{H}_{stat} .

In order to execute the circuit encoded by the binary string, we will assume that we are working in the ground space of \mathbf{H}_{stat} ; any other states necessarily have energy ≥ 1 . On the black layer A spins, we partition the Hilbert space \mathcal{H}_{loc} into $\mathbb{C}^2 \oplus \mathbb{C}^2$; each spin either stores a qubit $|q\rangle$, or it indicates one of two “mover” symbols \blacktriangleleft and \blacktriangleright .

We write transition rules for the two arrows, which move them around the cube according to their direction, while staying on the same layer (see section 3.2.4.1). Any qubit in their path is pushed down to the next layer and cycled one to the right if passed by \blacktriangleright , or one to the left if passed by \blacktriangleleft . Once an arrow arrives at the computation edge, a transition rule conditioned on the program bit pairs (p_i, p_{i+1}) (accessible through the green spins) performs the corresponding computational step on the two adjacent qubits. The arrow is then re-set to the next lower level, and the whole procedure repeats. Once the arrow returns to the computational edge and is at the bottom-most layer, there is no further forward transition; the program terminates.

Symbolically, the operations we can perform with this basic set of instructions are the following ones. We have a quantum state of N qubits $|q_1\rangle |q_2\rangle \cdots |q_N\rangle$. In one step, we can either...

1. cycle the qubits clockwise, to $|q_N\rangle |q_1\rangle \cdots |q_{N-1}\rangle$,
2. cycle them anti-clockwise, to $|q_2\rangle \cdots |q_N\rangle |q_1\rangle$,
3. perform a universal two-qubit quantum gate \mathbf{G} on the first two qubits,
4. or perform the inverse of this gate, i.e. \mathbf{G}^\dagger .

We prove in lemma 3.3 that there exists such a gate \mathbf{G} which is universal for quantum computation, even if only applied to adjacent qubits. Analogously to before, we collect all history state terms in the Hamiltonian \mathbf{H}_{hist} .

For us, the lattice instances of interest are the ones where the binary string corresponds to a circuit approximating the head of a QMA_{EXP} verifier QRM, i.e. $C_{\mathbf{R}}$ (the fact that most program strings do not

represent such a QRM is not important). In lemma 3.5 we perform a careful analysis of the approximation errors, and show that one can indeed choose height, width and depth of the lattice (depth corresponding to the encoded program, width to the ring size, and height to the run time of the verifier) such that the history state corresponds to a witness verification for any instance $l \in \Pi$, where Π can be any promise problem in QMA_{EXP} .

What remains to be done is to penalise invalid history state configurations, such as multiple active symbols, or no active symbol; collect those terms in an operator \mathbf{P} . Finally, an input penalty Π_{in} for the computation ensures that some ancillas are correctly initialised for the computation, and the output penalty Π_{out} raises the lowest energy for **NO**-instances.

Since our history state has branches (since not all transition rules we write down are completely unambiguous), we have to show that \mathbf{H}_{PROP} defines a so-called unitary labeled graph Laplacian and invoke a recently-proven variant of Kitaev’s geometrical lemma for this case, lemma 2.44. With a rigorous proof in section 3.3, we can thus show that the overall 4-local translationally-invariant Hamiltonian

$$\mathbf{H} := \mathbf{H}_{\text{stat}} + \mathbf{H}_{\text{PROP}} + \mathbf{P} + \Pi_{\text{in}} + \Pi_{\text{out}}$$

defined on the spin-3/2 cF lattice satisfies the promise gap $\lambda_{\min}(\mathbf{H}) \leq -\Omega(1/\text{poly}|\Lambda|)$ if l is a **YES**-instance, and $\lambda_{\min}(\mathbf{H}) \geq 0$ otherwise. This finishes the construction, and the claim of theorem 3.2 follows.

3.2 Turing’s Cube

3.2.1 Single Gate Universality

In order to execute the QRM, we have to be able to cyclicly apply the QRM head unitary on a pair of qudits. Since we will be working with qubits in our construction, we embed each such qudit into a list of qubits, and approximate the 2-local qudit unitary using a special 2-local unitary gate \mathbf{G} , which can act on any two neighbouring qubits. In order to apply Solovay-Kitaev to the QRM head unitary and approximate it with a $O(\log(1/\epsilon))$ gate count (as opposed to $\sim 1/\epsilon$), we have to be able to apply both \mathbf{G} and its inverse, \mathbf{G}^\dagger ; however, in Solovay-Kitaev, the requirement is that those two gates can be applied to *any pair* of qubits, whereas in our construction—as will become clear later—we can only ever apply either gate to neighbouring qubits.

While a random choice of \mathbf{G} will work, we pick \mathbf{G} explicitly (see e.g. [Chi+10]). It then suffices to prove that \mathbf{G} is universal when applied to adjacent qubits, which is what the following lemma shows. •

	I	2	3	4	5	6	7	8	9	10	11
2	3										
3	4	5									
4	6	7	8								
5		9	10	11							
6	12	13	14	15	16						
7		17	18	19	20	21					
8			22	23	24	25	26				
9		27	28	29	30	31	32	33			
10			34	35	36	37	38	39	40		
11				41	42	43	44	45	46	47	
12		48	49		50		51	52	53	54	55
13		56	57	58	59	60	61	62		63	

Table 3.1: A linearly independent set of generators for $\mathfrak{su}(8)$ in terms of nested commutators of \mathbf{H}_1 and \mathbf{H}_2 . For example, $\mathbf{H}_{42} := i[\mathbf{H}_{11}, \mathbf{H}_5]$.

Lemma 3.3 (Ozols [Ozo16]). *Define the 2-qubit unitary $\mathbf{G} := \exp(i\mathbf{H})$ with*

$$\mathbf{H} := \sigma_x \otimes \mathbf{1} + \mathbf{1} \otimes \sigma_z + \sigma_x \otimes \sigma_x + \sigma_z \otimes \sigma_z = \begin{pmatrix} 2 & 0 & 1 & 1 \\ 0 & -1 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \end{pmatrix}.$$

Then the unitaries $\{\mathbf{G}_{k,k+1 \pmod l} : k = 0, \dots, l-1\}$ generate a dense subset of $SU(2^l)$ for all $l \geq 3$, where the subscript denotes where the unitaries act.

Proof. Since 3-qubit unitaries generate a dense subset of $U(2^l)$ when applied to adjacent qubits, it suffices to prove the claim for $l = 3$. The proof follows techniques in Lie algebra [Chi+10]. Define $\mathbf{H}_1 := \mathbf{H} \otimes \mathbf{1}_2$ and $\mathbf{H}_2 := \mathbf{1}_2 \otimes \mathbf{H}$, and let $\mathcal{L}(\mathbf{H}_1, \mathbf{H}_2)$ be the Lie algebra generated by these two elements. For $j = 3, \dots, 63$, we set $\mathbf{H}_j := i[\mathbf{H}_{r_j}, \mathbf{H}_{c_j}]$, where r_j and c_j are the row and column numbers of entry j in Table 3.1. One can verify—using a computer algebra system—that the matrices $\{\mathbf{H}_1, \dots, \mathbf{H}_{63}\}$ are linearly independent, and traceless by construction. Since $\dim \mathfrak{su}(8) = 63$, they furthermore span the entire algebra, and the claim follows. \square

3.2.2 Circuit Encoding

We work with a face-centred cubic lattice of side lengths $D \times H \times W$, as shown in fig. 3.2. At each vertex we place a 4-dimensional spin with local Hilbert space $\mathcal{H}_{\text{loc}} = \mathbb{C}^4$, and we want to define a 4-local Hamiltonian on the lattice which embeds the evolution of a QMA_{EXP} verifier. Our construction comprises the following three main steps.

1. Binary counter. We construct a 2D tileset which lives on the top face of the cuboid, and translates the cuboid depth D into a binary description of D on the top front edge, which is of size $\log_2 D$. This binary string encodes a circuit C according to table 3.2 and fig. 3.3.
2. Shuffling the program. Using another 2D tileset, we cyclicly shuffle this circuit program around the sides of the cuboid and wind it down diagonally as shown in fig. 3.2. The front edge—marked in red—is the computation edge and will periodically see the entire binary description of the program.
3. Performing gates. On the sides of the cuboid, we superpose a layer of qubits. Labelling the qubits around the top edge of the cube with $|q_1\rangle |q_2\rangle \cdots |q_N\rangle$, we define transition rules which allow us to perform one of the following four operations:
 - a) cycle the qubits clockwise, to $|q_N\rangle |q_1\rangle \cdots |q_{N-1}\rangle$,
 - b) cycle them anti-clockwise, to $|q_2\rangle \cdots |q_N\rangle |q_1\rangle$,
 - c) perform a universal two-qubit quantum gate \mathbf{G} on the first two qubits,
 - d) or perform the inverse of this gate, i.e. \mathbf{G}^\dagger , on the first two qubits.

Once any gate operation is performed, all qubits are swapped with the ones on the next-lower level while cycling them in the direction specified. On the next layer, the same procedure repeats until the execution terminates after H steps (H being the height of the cube). The history state construction for one operation above thus requires $2 \times (D + W)$ steps.

The necessary transition rules are described in detail in section 3.2.4.2, and table 3.2 describes how the binary program description on the computation edge is interpreted as one of the four actions above at each level. Fig. 3.3 shows how any circuit can be encoded in this way. Observe that due to the winding program description—which is exposed periodically at the front edge—we necessarily apply the same circuit over and over again. In between each appearance of the description of $C_{\mathbf{R}}$ on the computation edge, the string of zeroes does not implement any gates or move the tape in either direction. Naturally, this is precisely the evolution of a Quantum Ring Machine.

For suitable circuits $C_{\mathbf{R}}$, this construction is thus a history state Hamiltonian which encodes an arbitrary QRM.

Remark 3.4. If we want to encode a QRM which runs for t applications of the QRM head, we necessarily need $H \geq 2t(D + W)$ for our cube. Furthermore, if the QRM head acts on two qudits of dimension d , the circuit $C_{\mathbf{R}}$ acts on $m = \lceil \log_2 d \rceil$ qubits; we thus require $D + W \equiv 0 \pmod{m}$.

For a fixed cube depth D encoding some BQEXP QRM, we need to ensure that we can tune the remaining two free parameters W and H —width and height of the cuboid—to provide enough space and time for the

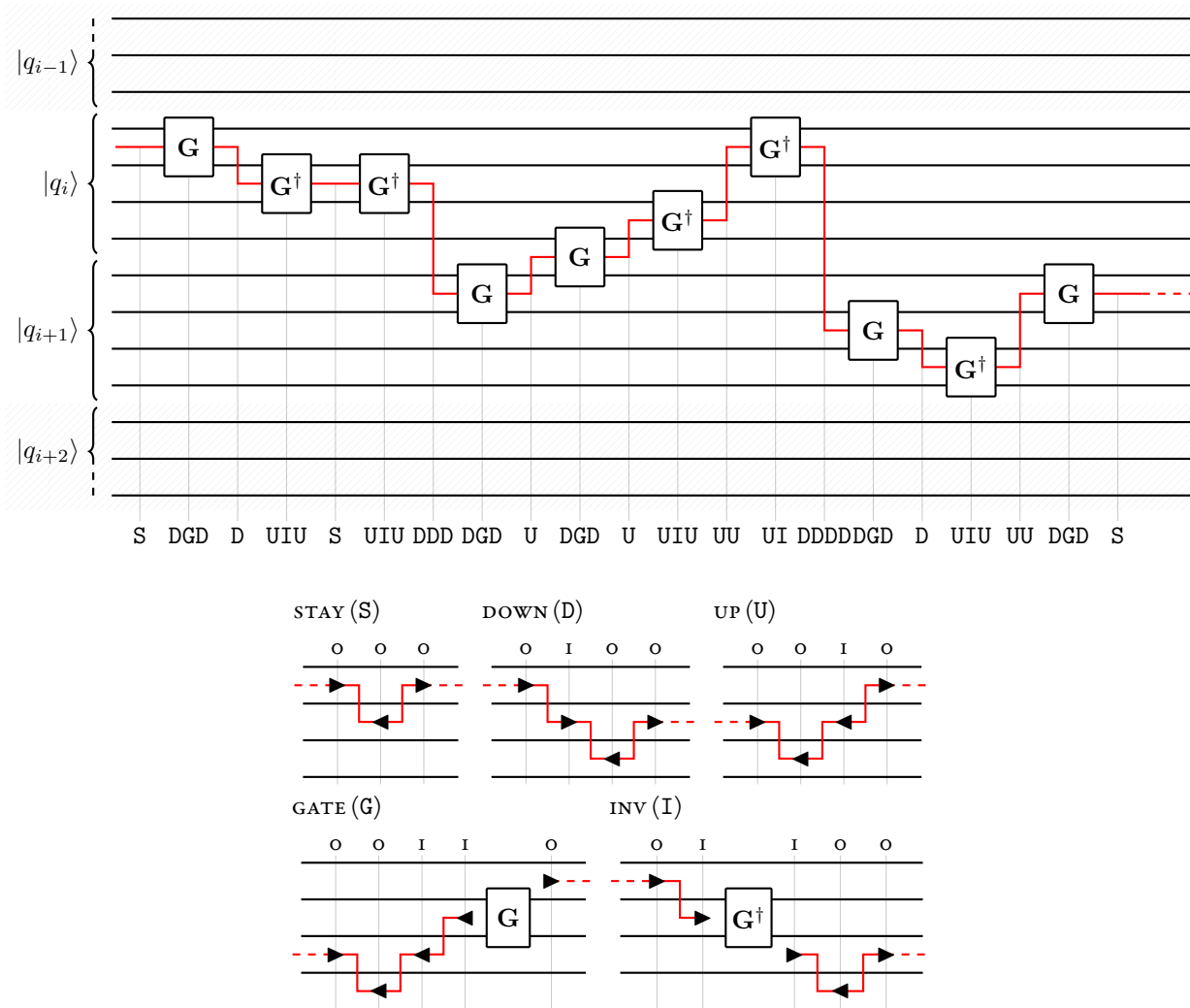


Figure 3.3: Execution order of an arbitrary circuit approximated using the universal gate G and its inverse G^\dagger . Each elementary operation start and end in a configuration \blacktriangleright , where the last program bit is a o —like this, each circuit can be constructed by a simple combination of these elementary operations, with a constant overhead. Observe that both gate application and inverse gate application do not end on the same line, which means that if we want to apply G at the current position, we have to execute DGD , and similarly UIU for G^\dagger . The specific quantum gate G that we use is proven to be universal in lemma 3.3.

computation to run and terminate, while at the same time keeping the error introduced by approximating the QRM head unitary within bounds. This is captured in the following technical lemma.

Lemma 3.5. *Take a BQEXP promise problem Π . For any precision $\delta > 0$ and instance $l \in \Pi$, there exist cube parameters $W, H, D = O(\exp \text{poly}(|l|, \log 1/\delta))$ which allow a verifier ring machine to be executed on the cube for instance l to within precision δ .*

Proof. Let $l \in \Pi$. A BQEXP witness computation for this instance l of size $|l|$ can be performed with a QRM with head unitary $\mathbf{R} \in SU((\mathbb{C}^d)^{\otimes 2})$ for some d . We require that the QRM head \mathbf{R} contains a description of instance l ; this means that d —the size of each of the two qudits that \mathbf{R} acts on—depends on the size of the instance, i.e. $d = O(\text{poly } |l|)$. Denote with t the number of steps the ring machine needs to perform to run the entire verifier computation.

1. In lemma 3.3 we show that there exists a specific 2-qubit gate \mathbf{G} which is universal for quantum computation, even when only applied to adjacent qubits.
2. Using S-K and a circuit encoding as described in fig. 3.3 using gates \mathbf{G} and its inverse \mathbf{G}^\dagger , approximate the QRM head \mathbf{R} with circuit $C_{\mathbf{R}}$ to some error $\epsilon \leq \delta/t$, where δ is the overall precision which we require for the verifier. Each qudit \mathbb{C}^d of the QRM verifier is encoded in $m = \lceil \log_2 d \rceil$ qubits. The circuit $C_{\mathbf{R}}$ thus acts on $(\mathbb{C}^2)^{\otimes 2m}$, i.e. m qubits. By [NC10], approximating an n -qubit unitary to within precision ϵ requires $O(n^2 4^n \log^c(n^2 4^n / \epsilon))$ gates (for some $c \leq 4$), if using their gateset; for our purposes it suffices to know that the number of gates required to approximate \mathbf{R} to within precision ϵ scales as $O(\text{poly}(d) \times \log^c(1/\epsilon))$.
3. The circuit description is thus of length $|C_{\mathbf{R}}| = O(\text{poly } |l| \times \log^c(1/\epsilon))$ and therefore we have to require that the depth of the cube $D = O(\exp(|C_{\mathbf{R}}|)) = O(\exp(\text{poly } |l| \times \log^c(1/\epsilon)))$.
4. The front sidelength W is increased...
 - to make the ring $r = W + D$ large enough for the computation, if it is not already, and
 - to make the ring size an integer multiple of $m = \lceil \log_2 d \rceil$.
5. Set $H = 2t(W + D)$.

With $\epsilon \leq \delta/t$ and $t = O(\exp \text{poly } |l|)$, we further have $\log^c 1/\epsilon \leq \log^c(t/\delta) = O(\text{poly}(|l|, \log 1/\delta))$, and the claim of the lemma follows. \square

Remark 3.6. *If we require cube parameters of $O(\exp \text{poly } |l|)$, we can demand a computation accuracy of at most $\delta = \Omega(1/\exp \text{poly } |l|)$.*

Proof. If we demand the two scaling parameters in lemma 3.5 to be equal, we have

$$\begin{aligned}
 & \exp \log^4 (1/\delta) = O(\exp \text{poly } |l|) \\
 \Leftrightarrow & \quad \log^4 (1/\delta) = O(\text{poly } |l|) \\
 \Leftrightarrow & \quad \log (1/\delta) = O(\text{poly } |l|) \\
 \Leftrightarrow & \quad \delta = \Omega(1/\exp \text{poly } |l|). \quad \square
 \end{aligned}$$

3.2.3 Static Lattice Constraints

3.2.3.1 Lattice Structure

We will work with a face-centred cubic lattice of 4-dimensional qudits. All interactions will be at most 4-local and translationally-invariant. The system will have open boundary conditions; in particular, we do not cut off interactions at the boundary or introduce boundary constraints of any kind. For the sake of clarity, when writing out constraints in the following, we will usually ignore parts of the sublattice, implicitly assuming that any interaction term is extended trivially everywhere else. When referring to layer A and if not explicitly mentioned, we mean the black sublattice, and layer B will be the red sublattice with side-centred vertices.

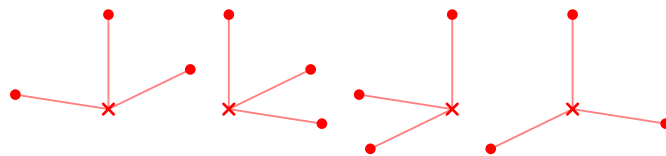
Any “static” constraint—i.e. the terms in the following four subsections—will be translated into local Hamiltonian terms diagonal in the computational basis; see section 3.2.3.6 for details.

3.2.3.2 Constraining the Lattice Bulk

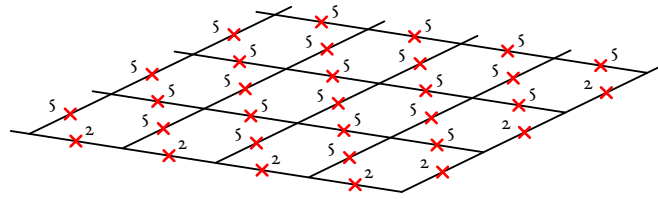
Denoting with \times a special symbol in the red sublattice, we want to constrain the lattice to be in this state in the bulk, and in its complement on the topmost red face, as well as the outermost side faces. We first give a bonus of τ to spins in the B sublattice in configuration



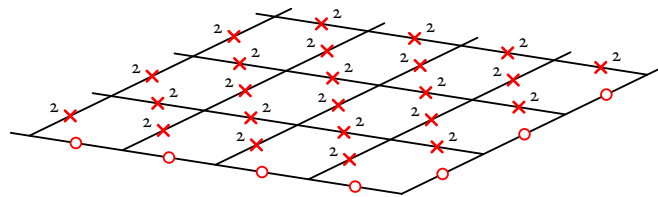
All red layers but the top one will then be in state \times . We then give a bonus of τ to all of the following configurations:



This leaves the top layer unchanged. Summarising the bonus terms so far, all other B layers, as seen from the top, are then in the configuration

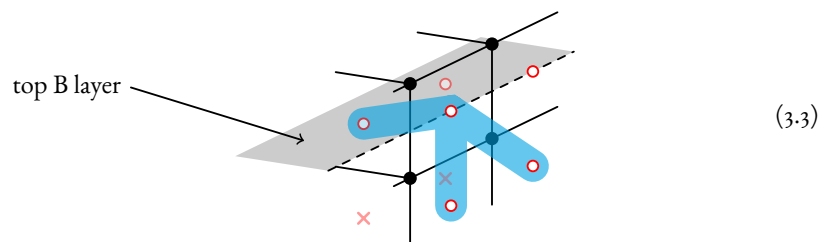


We then give a global i -local penalty to \times with strength -3 . The top B layer will thus be in the complement of \times (which we denote with \circ), while all the other B layers look like



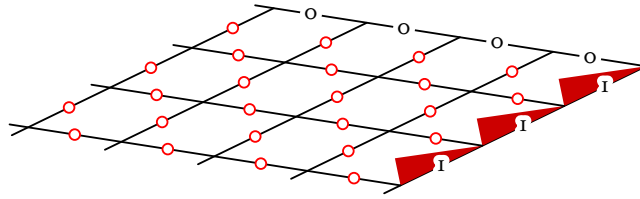
3.2.3.3 Binary Counter

The top layer of type B will carry a binary counter tiling, which translates the side length D into a binary representation on the top front edge. In order to achieve this, we need to initialise the top back edge of the cube to all 0, and the top right edge to all 1. Since we do not want to use distinct interactions on the outside layers, but have open boundary conditions, we have to find a configuration in the pre-constrained cube which only occurs on the top right and back edge, respectively. The following configuration is such an example:

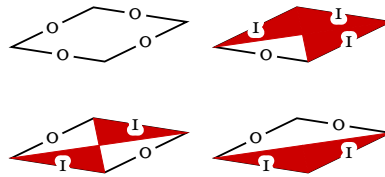


Since only the top and outer layers have red spins in configuration \circ , this four-local interaction allows us to pick out the top right boundary of the top layer, and to constrain it to state 1. A similar interaction allows

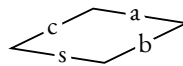
constraining the top back layer to o . The top B layer then looks like



Since this is the only B layer with spins in state o , we can use the following tiles from [Pat14] to get the desired binary counting layer.



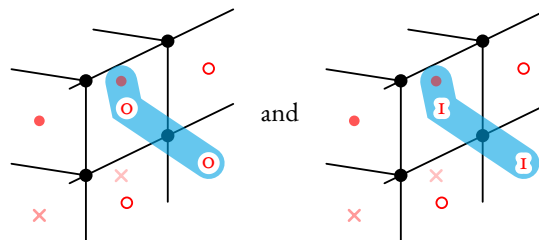
It is straightforward to verify that the general tile



obeys the rules $c = \text{carry of } a + b$ and $s = a \oplus_2 b$.

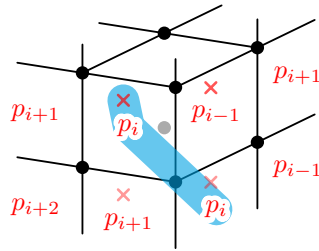
3.2.3.4 Winding Program Diagonally

We use an interaction similar to eq. (3.3) to shuffle the program around the cube in a cyclic fashion, as depicted in fig. 3.2:



Observe that, by including the red qudit one layer in, this interaction does indeed only apply to the front right face; similar interactions on the other three faces achieve the desired program copying around the cube sides. Additionally, by conditioning on if this inner qudit is either \times or o , we can apply a different rule at the top layer. In particular at the top layer of the front right face we want to flip the bit when copying down so that there are 1 's on the top layer but o 's on the layer below - see fig. 3.2.

On the corners, we use a similar shape of interaction, i.e.



and similarly for all other corners.

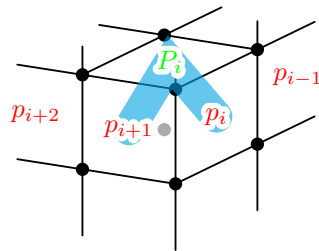
Note that in section 3.2.4.2, we will need to temporarily replace red program bits with a special symbol $!$ indicating that the application of a gate is happening in the next step, so we exclude this case from the constraints in this section (i.e. we allow *either* p_i and p_i , or p_i and $!$ to appear around the computational corner, and similarly for the diagonal face constraints bordering the computation edge).

As none of the dynamic transition rules below ever changes the number of head symbols (of which $!$ is one), we can rule out the cases where there is more than one $!$ or other head symbol present at any one time—we analyse these branching cases in detail in section 3.2.5.

3.2.3.5 Constraining layer A qudits

We label the states of the green face-centred qubits of the layer A type with the alphabet $\{A, B, C, 0\}$. For all such green lattice qubits we apply a bonus of strength $1/2$ to configuration 0 , so that this state is preferred.

In order to access two sequential program bits p_i and p_{i+1} with a single three-local interaction on the computation edge, we add a strength 1 interaction which constrains the front column of the layer A green sublattice to a state $P_i \in \{A, B, C\}$ depending on the two neighbouring computation bits, i.e.



Note that this interaction will have no effect anywhere else in the lattice, as at least one of the two red program bits will be \times .

The rule which governs what state P_i is constrained to will depend on the tuple p_i and p_{i+1} , and is derived from table 3.2. The idea is that $P_i = f(p_i, p_{i+1})$ will signify what is to happen at the computation

edge. Looking at table 3.2, we see that at each stage we either:

- A. Apply a gate (either \mathbf{G} or its inverse \mathbf{G}^\dagger , depending on where the arrow is coming from),
- B. go Backwards (i.e. change the direction of the arrow),
- C. or Continue in the same direction.

Given the encoding of table 3.2, we therefore take $P_i = f(p_i, p_{i+1})$ for a function f given by

$$f(p_i, p_{i+1}) = \begin{cases} B & \text{if } p_{i+1} = 0, \\ C & \text{if } p_i = 0 \text{ and } p_{i+1} = 1, \text{ and} \\ A & \text{if } p_i = p_{i+1} = 1. \end{cases}$$

Due to the aforementioned $1/2$ bonus which applies at all green spins, the remainder of layer A is in configuration 0.

3.2.3.6 Summary of static constraints

As explained in the main text under “Tiling Construction”, we take all static constraints listed so far and translate them to diagonal and local projectors \mathbf{h}_i . This allows us to write a 4-local, translationally-invariant classical Hamiltonian $\mathbf{H}_{\text{stat}} = \sum_{\vec{x}} \sum_{\mathbf{h}_i} \mathbf{h}_i^{\vec{x}}$ (i.e. product and diagonal in the computational basis of each spin) with a ground space spanned by states with the following properties.

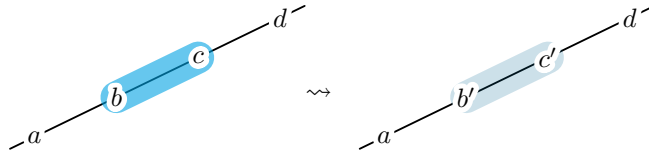
1. Any black vertex spin in layer A is unconstrained.
2. The red layer B spins will be in a state as depicted in fig. 3.2, i.e. on the top cuboid face, they represent a binary counter translating the depth D of the cuboid into a binary description of D on the top front edge. This binary string $s = p_1 \dots p_T$ is wound down diagonally around the cube, which expresses s periodically on the front computation edge. *Only* the spins adjacent to this edge are also allowed in a configuration !. In the bulk of the cube all the way to the bottom-most layer, the red spins are in state \times .
3. The green layer A is in configuration 0 everywhere but on the front edge; there, the spins there are in a configuration depending on the two adjacent program bits p_i and p_{i+1} , as outlined above.

This Hamiltonian \mathbf{H}_{stat} is gapped with a size-independent constant gap, and we can rescale the interactions so far and shift the overall energy to assume that this ground space as detailed above has energy zero, and any other configuration has energy lower-bounded by 1.

In the next sections, we will explain the history state construction, which—within this ground space of \mathbf{H}_{stat} —will represent a valid QRM evolution for the circuit represented by the binary string s .

3.2.4 Dynamic Lattice Constraints

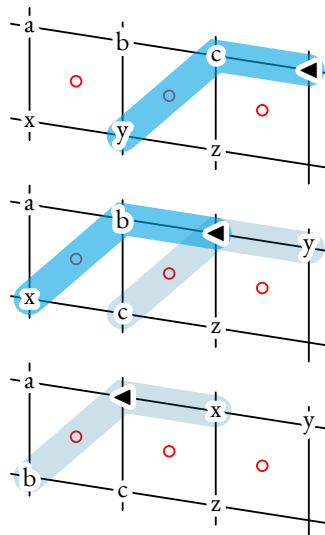
We always depict a transition rule as connected by a squiggly arrow \rightsquigarrow ; the notation is self-explanatory: the brighter blue shading indicates the original state, whereas the dull blue shading indicates the target configuration. To give an example, a transition



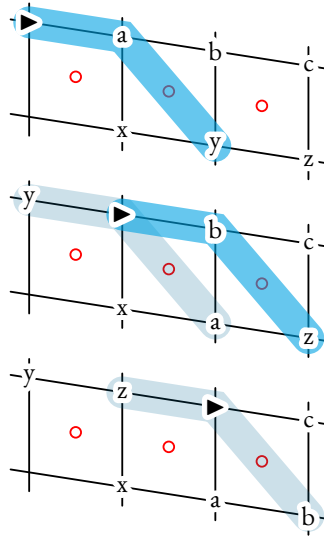
would be translated into a two-local term $\mathbf{h} = |bc\rangle\langle bc| + |b'c'\rangle\langle b'c'| - |b'c'\rangle\langle bc| - |bc\rangle\langle b'c'|$, and correspondingly with an extra quantum register if b or c were labelling vertices that carry a qubit (i.e. the black layer A sublattice vertices).

3.2.4.1 Moving Qubits

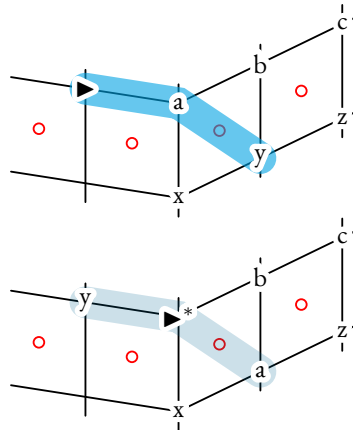
The black sublattice (A layers) comprises the alphabet $\{0, 1, \blacktriangleright, \blacktriangleleft\}$, where we treat the 0, 1-subspace as a qubit, i.e. \mathbb{C}^2 . The right and left arrows are markers to indicate where to move qubits to. As an example on the front face, we have a left moving sequence



and analogously the right moving sequence



To move qubits around a corner, we use an interaction of the form



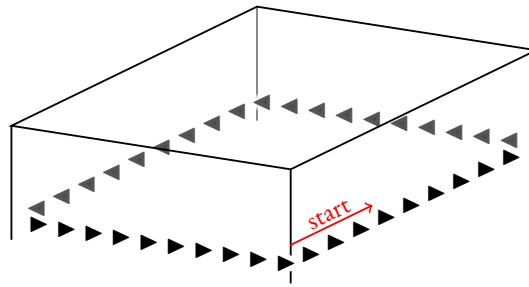
at the back, left and right corners (different rules as described in section 3.2.4.2 are used for the front edge) and similarly for going around the corner in the opposite direction.

A few remarks: first note that all the transitions defined so far are unique, i.e. given the cube bulk constrained to \times as done in section 3.2.3.2, and for every configuration with only one arrow symbol (the other cases we will penalise as a last step), there exists precisely one forward and one backwards transition. Another important point is how to modify the arrows when going around the circumference of the cube once (marked with a \blacktriangleright^* in the last transition rule); at the moment, if we left the arrow type unchanged for every corner, we would not be able to shuffle around the qubits in a circle; on the back face, we would be doing the opposite shuffling operation. Therefore, we change the arrow type according to the following

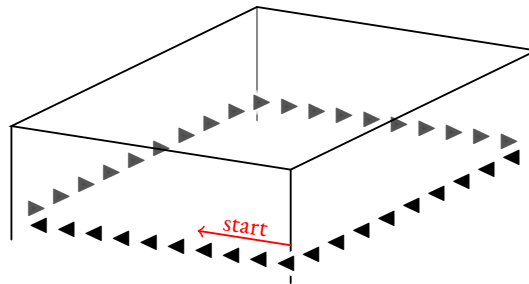
incoming tape head	program $x_{n-1}x_n$	operation on qubits	outgoing tape head
▶	oo	$\mathbb{1}$	◀
▶	oi	$\mathbb{1}$	▶
▶	io	$\mathbb{1}$	◀
▶	ii	\mathbf{G}^\dagger	▶
◀	oo	$\mathbb{1}$	▶
◀	oi	$\mathbb{1}$	◀
◀	io	$\mathbb{1}$	▶
◀	ii	\mathbf{G}	◀

Table 3.2: Program encoding. The arrow symbols ▶ and ◀—i.e. the heads moving on the tape—indicate in which direction the ring is moving. Relative to the tape, the current head is thus moving in the opposite direction. With this encoding, any circuit can be executed with the available operations, see fig. 3.3.

scheme:



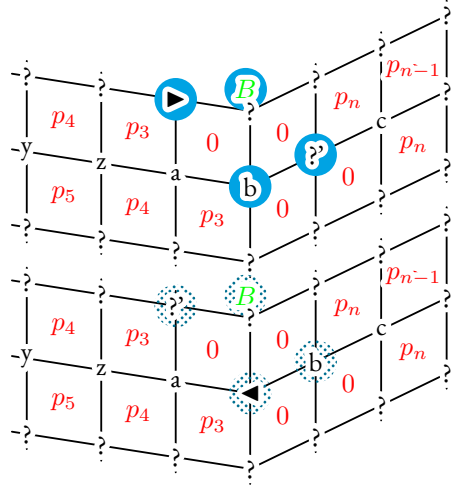
and



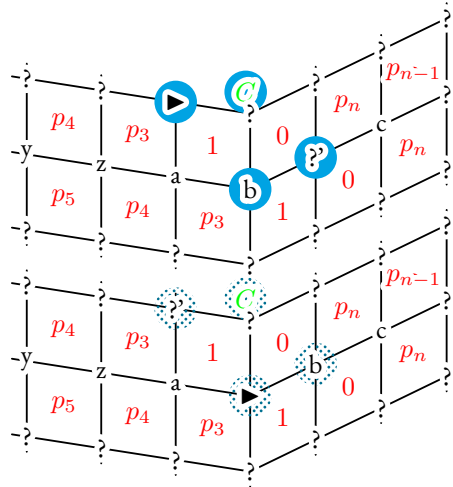
3.2.4.2 Computation

In order to execute any circuit as in fig. 3.3, we have eight elementary operations available, all of which are listed in table 3.2. It is easy to see that there exists a symmetry between the right- and left-moving arrow; we will thus explain the right-moving arrows (including the application of gate \mathbf{G}) in detail and leave the reverse direction as an exercise to the reader.

Case 1. Consider $p_1 p_2 = 00$ or 10 , so that $f(p_1, p_2) = B$. The transition is conditioned to only happen if the green qubit in the layer A sublattice is in the B state. Move $?$ up, b to the right and flip the arrow. This corresponds to simply reverting the direction as in table 3.2.

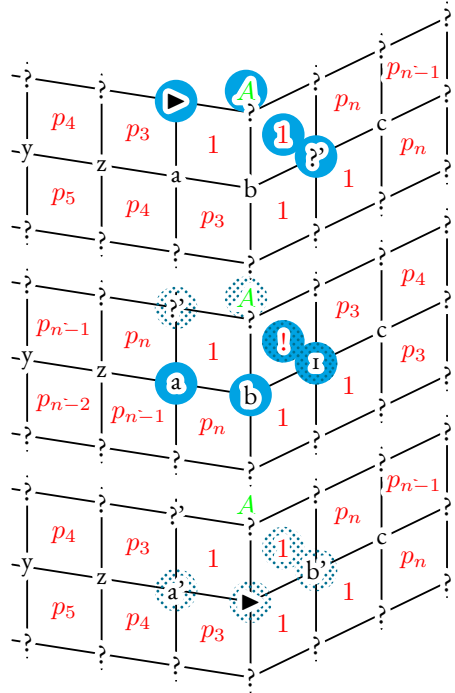


Case 2. Consider $p_1 p_2 = 01$ so that $f(p_1, p_2) = C$. We perform the same action as above, but keep the arrow direction.



Case 3. Consider $p_1 p_2 = 11$ so that $f(p_1, p_2) = A$. We want to execute a gate, which requires one intermediate step. We place the computation marker on the right hand side of the computation edge. This signals that the next step is to perform a gate \mathbf{G} on a and b . Here $|a'\rangle |b'\rangle := \mathbf{G} |a\rangle |b\rangle$. The program is

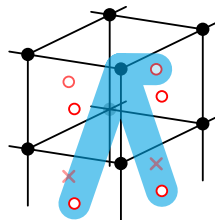
restored and the arrow left in the right moving configuration, as required by table 3.2.



We now move the arrow once around the tape and then arrive at the computational corner from the other side. Observe—as mentioned—that the encoding in table 3.2 is mirror-symmetric, so by reversing all the rules above one can implement the same rules—while applying G^{-1} instead of G when $P_i = A$ for an arrow incoming from the right.

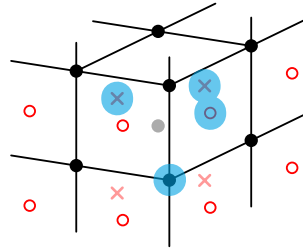
3.2.4.3 Computational Input and Output Constraints

Since the instance is specified within the QRM head, it suffices to provide the computation with a single ancilla $|0\rangle$ as input; in case we need more ancillas than available on the front edge, we can augment our verifier as in fig. 2.12. Due to the configuration of the red layer B sublattice, it is straightforward to find a local configuration which only ever appears on a top right corner; more specifically, we utilise the constraint interaction

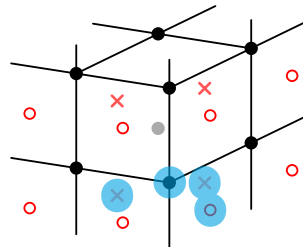


to enforce that the black symbol is either in an arrow configuration, or o, respectively. The rest of the tape is left unconstrained.

Since there is nothing special about the bottom-most layers A and B, we need to use a pair of interactions to enforce the last black qubit to an accepting state. This can be readily achieved using



constraining the black qubit to state $|0\rangle$, and



giving a bonus to the complement configuration. Everywhere but on the bottom-most layer, the two penalties precisely cancel; however, on the last layer, only the projection onto $|0\rangle$ survives, which thus acts as output penalty once the computation is terminated.

3.2.4.4 Multiple Heads Penalty

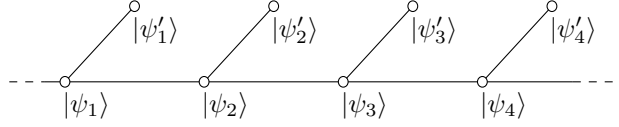
Since we only want to allow precisely one head on the computational layer, we will penalise any configuration where two heads are next to each other. This finishes our construction.

3.2.5 Analysis of History State Branching

In this section, we want to analyse all transition rules and show that the parts where they are ambiguous do not break the evolution of the computation. First note that all constraints in section 3.2.3 are static, i.e. there are no possibilities for any ambiguities in the configuration. We will call configurations that obey all those static constraints and have precisely one head symbol on the computational layer—i.e. exactly one of \blacktriangleright , \blacktriangleleft or $!$ —*valid* configurations.

We will go through each dynamic penalty in section 3.2.4 separately.

1. In section 3.2.4.1, the transition rules for the faces are unambiguous, since they depend on the red symbol to be in a configuration \circ .
2. The rules for moving around a corner, however, *can* happen on a face: in this case, the arrow symbol is moved one layer into the bulk. Observe though that none of the movement transitions can apply to the arrow when it is inside of the bulk (apart from moving it back out with a reverse transition), so the computation branches, but the leg does not proceed: we obtain an evolution of the form



where all the primed states are redundant, but at most enlarge the overall evolution by a factor of 2.

3. In section 3.2.4.2, the computation transitions are unambiguous; observe in particular that there is no transition rule that simply copies the arrow around the computation edge (by construction, see section 3.2.4.1).
4. Finally, the input and output constraints are static again.

This allows us to formulate the following two branching lemmas.

Lemma 3.7. *Any valid history state for the given transition rules is of size $O(\text{poly}(W, D, H))$, where W , D and H are the cuboid's width, height and depth.*

Proof. Follows by construction; the head can perform at most $O(H \times (W + D))$ unique transitions. \square

Lemma 3.8. *In case there is more than one head symbol (i.e. \blacktriangleleft , \blacktriangleright or \blacktriangle) present, the minimal valid evolution splits up into poly-sized slices, each of which carries at least one penalty from two directly adjacent heads.*

Proof. The argument is the same as in theorem 2.60. One can keep all but one of the head symbols fixed; the one left free to move is necessarily meeting another head symbol within poly many steps. \square

3.3 QMA_{EXP}-Hardness Proof

In this section, we provide a rigorous proof of theorem 3.2. We want to point out that the Hilbert space structure of this lattice Hamiltonian \mathbf{H}_{PROP} is not a product space between clock and computation space $\mathcal{H}_{\text{clock}} \otimes \mathcal{H}_{\text{comp}}$, which would result in a ground state of the standard history state form $\sum_t |t\rangle |\psi_t\rangle$. The reason for this is that depending on which sub-lattice a spin sits on, its local Hilbert space $\mathcal{H}_{\text{loc}} = \mathbb{C}^4$ decomposes differently. The red and green spins can be regarded as being completely in the clock space,

- as all transition rules which act on them are completely classical, i.e. they never move any of the red and green spins out of a computational basis state. The black spins, however, decompose into a direct sum $\mathcal{H}_{\text{clock}} \oplus \mathbb{C}^2$, the latter space carrying a qubit, and the clock part being reserved for the two arrow symbols \blacktriangleleft and \blacktriangleright , which are part of the clock.

In order to analyse the spectrum, we note that there exists an isometric transformation between our Hamiltonian and Hilbert space, and one which respects the product space structure, which in particular will allow us to view the Hamiltonian as a ULG Laplacian and apply lemma 2.44. In the language of Quantum Thue Systems (cf. definition 2.51) we can state the following lemma.

Lemma 3.9. *The transition rules in section 3.2.2 define a Quantum Thue System, and the induced ULG is simple.*

Proof. Verifying that the rules define a Quantum Thue System is straightforward by a simple re-ordering of the spins. Simplicity of the corresponding unitary labelled graph follows from lemma 3.7, and we refer the reader to definition 2.37. \square

Without further ado, we now proceed to the proof of theorem 3.2, which we re-state here in a rigorous, but concise fashion.

Theorem 3.10. *(4, 4)-TILH-3D is QMA_{EXP} -complete.*

Proof. Containment in QMA_{EXP} is clear (see theorem 2.17). To show that the Hamiltonian instances of the cube construction define a QMA_{EXP} -hard family, we will employ techniques proven there which should simplify the analysis.

Let $\Pi = (\Pi_{\text{YES}}, \Pi_{\text{NO}})$ be a QMA_{EXP} promise problem, as in definition 2.13. By lemma 3.5, we know that we can pick a constant error threshold $\delta > 0$ such that for any instance $l \in \Pi$ there exists a cube which allows a verifier circuit for this instance to be executed on the sides. Since we will require probability amplification later on in the proof (fact 2.12), we set $\delta = f(|l|)$ for some function f to be specified later, and also assume that the original verifier's acceptance probability is $\epsilon_l \leq f(|l|)$.

We translate all static and dynamic penalties into a Hamiltonian as explained in chapter 2 and eq. (2.7), and denote the corresponding Hamiltonian operator with

$$\mathbf{H} = \mathbf{P} + \mathbf{H}_{\text{prop}} = \mathbf{P}_{\text{in}} + \mathbf{P}_{\text{out}} + \mathbf{P}_{\text{static}} + \mathbf{P}_{\text{heads}} + \mathbf{H}_{\text{prop}},$$

where $\mathbf{P}_{\text{static}}$ comprises all static constraints for the cube (cube structure, binary counter and winding of program), $\mathbf{P}_{\text{heads}}$ penalises any two head symbols next to each other, and such that $\mathbf{P}_{\text{in/out}}$ represent the input and output penalties, respectively.

Soundness. We first address the case when $l \in \Pi_{\text{YES}}$. Denote with $|\Psi_l\rangle$ the valid history state, i.e. the unique uniform superposition ground state of \mathbf{H}_{prop} started out in a valid initial configuration with a single left-moving head in the top left row, and such that no initial or static penalty is violated. Then

$$\begin{aligned}
\langle \Psi_l | \mathbf{H} | \Psi_l \rangle &= \langle \Psi_l | \mathbf{P}_{\text{in}} | \Psi_l \rangle &&= 0 \quad (\text{i}) \\
&+ \langle \Psi_l | \mathbf{P}_{\text{out}} | \Psi_l \rangle \\
&+ \langle \Psi_l | \mathbf{P}_{\text{static}} | \Psi_l \rangle &&= 0 \quad (\text{ii}) \\
&+ \langle \Psi_l | \mathbf{P}_{\text{heads}} | \Psi_l \rangle &&= 0 \quad (\text{iii}) \\
&+ \langle \Psi_l | \mathbf{H}_{\text{prop}} | \Psi_l \rangle &&= 0 \quad (\text{iv}).
\end{aligned}$$

Term (i) is zero because $|\Psi_l\rangle$ satisfies all input constraints, (ii) because $|\Psi_l\rangle$ is a valid history state, (iii) because $|\Psi_l\rangle$ has precisely one active head symbol, and (iv) since $|\Psi_l\rangle$ is a ground state of \mathbf{H}_{prop} .

What remains to be analysed is the output penalty $\langle \Psi_l | \mathbf{P}_{\text{out}} | \Psi_l \rangle$. If we write $|\Psi_l\rangle = \frac{1}{\sqrt{T}} \sum_{t \in T} |t\rangle |\psi_t\rangle$ where T is the normalisation constant for the history state (i.e. the number of unique vertices in the ULG evolution represented by $|\Psi_l\rangle$), which we know by lemma 3.7 to be $T = O(\text{poly}(W, D, H))$ —i.e. the number of computational steps taken, including branching, cannot be larger than a polynomial in the cube width, depth and height. Then

$$\begin{aligned}
\langle \Psi_l | \mathbf{P}_{\text{out}} | \Psi_l \rangle &= \frac{1}{T} \left(\sum_{t, t' \in T} \langle t | \langle \psi_t | |T\rangle \langle T| \otimes \Pi_{\text{out}} |t'\rangle | \psi_{t'} \rangle \right) \\
&= \frac{1}{T} \langle \psi_T | \Pi_{\text{out}} | \psi_T \rangle \\
&= \frac{1}{T} \mathbb{P}(\text{circuit rejects}) \\
&\leq \frac{1}{T} (\epsilon_l + \delta) \\
&= \frac{2f(|l|)}{T}.
\end{aligned}$$

Completeness. If $l \notin \Pi_{\text{YES}}$, we have to show that for any $|\psi\rangle$, $\langle \psi | \mathbf{H} | \psi \rangle$ is bounded away from the YES-case by a $1/\text{poly}$ gap. If any static constraint is violated, we can immediately bound $\mathbf{H} \geq 1$. So we can assume that the state $|\psi\rangle$ is in a valid configuration.

Note that the number of head symbols \blacktriangleright , \blacktriangleleft or $!$ is always preserved for any transition rule. This means that \mathbf{H}_{prop} —and therefore also \mathbf{H} —is block-diagonal in the static cube configuration and the number of head symbols on the computational layer, we can address each case separately.

- i. In case of multiple head symbols we observe that each head necessarily sweeps the entire surface of the cube. Mark an arbitrary head symbol, and define $\mathbf{H}'_{\text{prop}}$ to be \mathbf{H}_{prop} with any transitions

for the other heads removed. Any such transition rule as in eq. (2.9) is positive semi-definite, which necessarily means $\mathbf{H}_{\text{prop}} \geq \mathbf{H}'_{\text{prop}}$ (spectrum wise, by which we mean $\mathbf{H}_{\text{prop}} - \mathbf{H}'_{\text{prop}}$ is psd itself). This new operator $\mathbf{H}'_{\text{prop}}$ might be non-local, but we only need it to lower-bound the spectrum of \mathbf{H}_{prop} .

The marked head symbol will then encounter another head in at most $\text{poly}(W, H, D)$ many steps (it cannot take longer than visiting the entire surface of the cuboid, see lemma 3.8). At that point, it will pick up a penalty. Utilising our variant of Kitaev's lemma (lemma 2.44), we conclude

$$\begin{aligned} \mathbf{H} &= \mathbf{P}_{\text{in/out}} + \mathbf{P}_{\text{static}} + \mathbf{P}_{\text{heads}} + \mathbf{H}_{\text{prop}} \\ &\geq \mathbf{P}_{\text{heads}} + \mathbf{H}'_{\text{prop}} \\ &\geq \Omega(1/\text{poly}(W, H, D)). \end{aligned}$$

We thus need to set f to a function which allows a polynomial separation (in the system size) between YES and NO instance; by remark 3.6, this is always possible.

2. The same argument lets us bound

$$\mathbf{H} \geq \mathbf{P}_{\text{in/out}} + \mathbf{H}_{\text{prop}} = \Omega(1/\text{poly}(W, H, D))$$

in case of a single head valid history state since l is a NO-instance.

3. What remains to be analysed is the zero head case. There are two standard approaches: we can either increase the number of symbols on the computational layer, such that on one side of a head—i.e. behind it in direction of the computation—we take one kind of symbols, and on the other side we take the other set; constructions like this can be constrained by a regular expression (without repetition of symbols, cf. [GI3, lem. 5.2]) and thus penalised with local terms.

Since our benchmark tries to reduce the local dimension of the system, we instead add a bonus term \mathbf{B} to the Hamiltonian \mathbf{H} , and such that $\mathbf{B}|\psi\rangle = -g(|l|) \times h|\psi\rangle$ where h is the number of head symbols for any basis state $|\psi\rangle$ of \mathbf{H} , and g is a function chosen such that there is again a $1/\text{poly}$ separation between the zero head state and the ground state for YES-instances, but such that multiple head configurations stay bounded away from 0. It is clear that \mathbf{B} can be implemented by a 1-local term of the form $-g(|l|)|\text{head}\rangle\langle\text{head}|$.

To be more precise and to determine how quickly g has to grow, assume that the construction up to now satisfies $\lambda_{\min} \leq 1/A$ for the YES case, and $\lambda_{\min} \geq 1/B$ in the NO case (excluding zero heads), where $B = O(\text{poly}(W, H, D))$, and $A \geq 4B \times W \times H \times D$. Choose $g = 2/A$; since there can

be at most $W \times H \times D$ heads on the cuboid's faces, we obtain the bounds

$$\lambda_{\min} \leq 1/A - 2/A = -1/A$$

if $l \in \Pi_{\text{YES}}$. Otherwise, for at least one head, •

$$\lambda_{\min} \geq 1/B - 4(W \times H \times D)/A \geq 1/B - 1/B \geq 0.$$

Finally, the zero head case can be trivially lower-bounded by $\mathbf{H} \geq 0$.

We have thus shown a promise gap of $1/\text{poly}$ in the system size: for $l \in \Pi_{\text{YES}}$, $\mathbf{H} \leq -\Omega(1/\text{poly}(W, H, D))$, and $\mathbf{H} \geq 0$ otherwise. The claim of theorem 3.2 follows. □

3.4 Chapter Summary

The quest for ever-more physically realistic families of QMA hard local Hamiltonians has arguably led us to increasingly contrived constructions. The increase in complexity necessary when going from non-translationally-invariant constructions to translational invariance is striking [GI13], and the same holds true for the effort to bring the local dimension back within reasonable range (chapter 2). On the other hand, almost always some fundamental new piece of machinery had to be developed, advancing our knowledge about circuit Hamiltonians: such as allowing branching to happen in the computational path, or using easier-to-implement computational models (Quantum Ring Machines), of independent interest e.g. in the context of adiabatic quantum computation ([WL15]).

In our case, we combine our construction with Wang tiles, which to our knowledge have not ever been used for this purpose. This “outsourcing” of part of the computation to a classical constraint satisfaction problem saves a significant amount of overhead for the control machinery surrounding the actual quantum verification procedure. Furthermore, the single universal quantum gate could be of independent interest in other applications, as it is reasonable to imagine a physical set-up where gates can only be applied to adjacent qubits in a circuit.

In fact, our 3D construction showcases that the embedded computation need not be highly obscure, and can, in contrast, even be quite elegant, as is evident by the much lower required local dimension and the therefore much smaller number of possible interactions necessary. By moving beyond simple spatial lattices, we can show that such structures *support* the emergence of more complex behaviour, despite the intrinsic symmetry of the crystal lattices we employ. By making use of these novel features, we are able to reduce the local dimension by two orders of magnitude as compared to the best result known to date.

We suggest three concrete open problems.

1. While our cube crystal structure is three-dimensional, we do not exploit its bulk structure beyond making use of its different sides. But there are small universal machines in higher dimensions (e.g. 2D or 3D Turing machines, Turmites, or cellular automata) which might be of use for improving this result further. This also leaves open the question of the required local dimension necessary for any 2D construction.
2. Including classical computation parts with Wang tiles is one step, but are there other, fundamentally different sets of local interactions even suitable to encode parts of a quantum computation?
3. A bottom-up approach proving a *lower* bound on the local dimension (or locality) of the interactions would be an alternative route to new insights into the LOCAL HAMILTONIAN problem. We want to emphasise that there is not much space left for any optimisation: as mentioned in the introduction, our construction allows each coupling to have $\approx 10^4$ free parameters; by the same benchmark, physically realistic spin lattices found in nature allow somewhere around $(3 \times 3)^2 \approx 80$ different couplings.

Recent results show that e.g. 1D gapped Hamiltonian ground states can be approximated efficiently (i.e. in randomised poly-time, cf. [LVV15]), but since history state constructions have a spectral gap that closes inverse-polynomially with the runtime of the encoded computation, a lower bound on the required local dimension remains open.

4 Size-Driven Quantum Phase Transitions

His conclusion was that things were not always what they appeared to be. The cub's fear of the unknown was an inherited distrust, and it had now been strengthened by experience. Thenceforth, in the nature of things, he would possess an abiding distrust of appearances.

—Jack London, *White Fang*

As we have seen in chapters 1 to 3, quantum systems can have complex dynamical and static properties. In this chapter, we want to study if we can exploit the idea of embedding computation into the ground state of a local many-body system even further, to describe materials with novel physical behaviour, and move even further towards a theory that has potentially directly-observable consequences for our real world.

The thermodynamic limit of many-body quantum Hamiltonians is the predominant mathematical tool used to study macroscopic properties of physical systems. A common approach is to analyse a growing sequence of finite system sizes—numerically or experimentally—and then extrapolate the characteristics of interest to the macroscopic limit [ŠB13]. This approach has been proven highly successful in numerous cases [LL80]; [Mar92]; [Bar83]; [Pir+12]; [Tag+08]. On the other hand, it has been shown that e.g. determining whether a system is gapped or gapless in the thermodynamic limit is an undecidable problem [CPW15b]. In order to correctly extrapolate the thermodynamic properties of a physical model, it is important to distinguish and recognise features that are a consequence of *finite-size effects*, i.e. properties of the model which are not present in the thermodynamic limit but appear as a by-product of conditions which only hold for systems sizes smaller than some threshold. While some finite-size effects only produce small perturbations of the real model, this is not always the case. For example, relevant finite size effects for the distinct behaviour of antiferromagnets on even or odd system sizes have been proposed in [Lou+08] and recently observed experimentally in [Gui+15].

In this work we show that finite-size effects can in fact be dominant at arbitrary length scales, to the point of completely obscuring the physics of the thermodynamic limit. This phenomenon occurs not just in pathological examples, but even e.g. for translationally invariant Hamiltonians on low-dimensional spins

arranged on a square lattice.

Main result 1: We explicitly construct models exhibiting the following exotic finite-size effects: below a threshold lattice size with sides of length N , the ground state of the Hamiltonian is a non-degenerate product state in the canonical basis, i.e. entirely classical, with a constant spectral gap above it. For system sizes greater than N , however, the low energy space is that of the Toric Code, which is in a sense as quantum as possible: the ground state exhibits topological degeneracy, and the system has anyonic excitations.

Moreover, we provide lower bounds on the threshold lattice size N for spin dimensions ≤ 10 (cf. table 4.1). Already for dimension 10 the threshold size can be as large as $5.2 \cdot 10^{36534}$.

Since in practice in a real-world experiment the ground state cannot be accessed, and only the Gibbs state at some small but non-zero temperature can be prepared, we also prove for one of our models that:

Main result 2: There exists a finite temperature below which measurement in system sizes smaller than the threshold still yields classical results up to small errors, while the thermodynamic limit converges for low temperatures to the ground state of the Toric Code [CNN16]. Even under a strong fidelity requirement of 10^{-6} , the necessary temperatures are rather mild (cf. table 4.1).

This sudden and dramatic change in the nature of the ground state may be viewed as a type of *quantum phase transition*, driven by the system size rather than a varying external field or coupling strength.

It has been known for some time that the critical values of external parameters (e.g. temperature, pressure) can depend on the size of the studied samples. Well-studied effects include rising melting points for small particles [BB76]; [GEA92], structural temperature- or pressure-dependent phase transitions between different crystal lattices in thin-film samples and in nano-crystals [TA94]; [McH97]; [Riv+11]; [Li+16], where the energetically favourable structure differs from that in the thermodynamic limit. And charge density wave order transitions or superconductivity [Xi+15]; [Yu+15], for which the critical temperature changes when approaching mono-layer sample sizes.

Crucially, these works all describe phase transitions driven by an external parameter, whose critical value *varies* depending on the size of the system. Here, we exhibit a transition which is driven by the system size itself; the transition occurs at some critical system size, without any external parameters varying at all. The effects which are most reminiscent of what we prove rigorously here are certain peculiar phenomena for mono-layer samples, or samples with 3 or 13 atom layers, for which the described phase transition cannot be observed anymore [Xi+15]; [Yu+15]; one suggested explanation is a lack of space for nucleation sites [TA94]; [Li+16].

Table 4.1 shows an overview of the explicit examples we construct. The threshold system sizes N_d from these examples show that large thresholds are achievable with d -dimensional spins. These are of course lower bounds on the maximum possible threshold size for given local spin dimension; even larger size thresholds may be achievable by other constructions. We chose for concreteness to construct a size-driven transition

d	4	6	7	8	9	10
N_d	2	15	84	420	$3.3 \cdot 10^7$	$5.2 \cdot 10^{36534}$
$T_d[\frac{\Delta}{k_B}]$	0.058	0.050	0.043	0.038	0.020	5.9μ

Table 4.1: Lower bounds on the maximum threshold lattice size $N_d \times N_d$ for different spin dimensions d , after which finite-size effect suddenly disappear and the physics of the thermodynamic limit becomes accessible. Up to dimension 8, a prime periodic Wang tiling gives the lower bound; for larger dimensions, an embedding of Busy Beaver Turing machines. The critical temperature T_d gives an estimate for the temperature at which the transition can still be discriminated with a fidelity of $1 - 10^{-6}$, as a function of the system's spectral gap Δ , which here is equal to the interaction strength since the Hamiltonians are commuting.

from classical to Toric Code; our constructions can readily be generalised to instead produce a size-driven transition to other quantum phases.

In order to prove these effects mathematically rigorously, we deliberately construct examples for which there exists an analytic solution. However, this is not true for the general case: as the structure of the Hamiltonian becomes more complex, one expects the behaviour to become more erratic. Indeed, we know that for extremely complex Hamiltonians with very large local spin dimension the behaviour can even become uncomputable [CPW15b].

It is important to emphasise that the dramatic finite-size effects exhibited here do *not* depend on any careful tuning of coupling strengths, and even occur for Hamiltonians without obvious separation of energy scales in their coupling constants or the matrix entries of the local interactions. Without this restriction, i.e. allowing interactions of magnitude $O(1)$ and $O(1/N^2)$, it is in fact trivial to construct a model whose ground state changes character at system size $O(N)$, with the spectral gap closing as $O(1/N)$. Our result is much stronger, in the sense that it does not allow such a prediction based solely on an analysis of the coupling strengths, nor from extrapolation of spectral data; in particular, the spectral gap of our model remains constant all the way up to the transition. •

4.1 Preliminaries

4.1.1 Embedding a Generalised Tiling into a Hamiltonian Spectrum

We have briefly introduced tiling problems in section 3.1.2. In this section, we rigorously formulate the embedding of the tiling problems we consider in this work into the spectrum of a local Hamiltonian. Instead of focusing only on star and plaquette interactions, we take an abstract point of view and define the notion of a *generalised tiling*. Assume $\mathcal{G} = (V, E)$ is a finite undirected graph with coloured vertices, where we allow colours $C := \{1, \dots, c\}$, $c \in \mathbb{N}$. Let $\mathcal{L} := \{l : l \subset \mathcal{G}\}$ be a finite set of (local) interactions, e.g. all the 3- or 4-local star and plaquette interactions on a lattice as in fig. 4.1. For all interactions $l \in \mathcal{L}$, we allow

a finite set of *pieces* $\mathcal{T}_l := \{(c_v)_{v \in l}\}$ —where the family $(c_v)_{v \in l}$ assigns a colour to every vertex in l —and a weight function $w_l : \mathcal{T}_l \rightarrow \mathbb{R}$. Now assign a colour to each vertex in \mathcal{G} , e.g. by defining a family $(c_v)_{v \in V}$, $c_v \in C$. The *score* of this assignment is then given by

$$\text{score} := \sum_{l \in \mathcal{L}} \begin{cases} 1 - w_l(t_l) & \text{if } (c_v)_{v \in l} \text{ is a valid piece in } \mathcal{T}_l \\ 1 & \text{otherwise.} \end{cases}$$

For $w_l(t_l) < 1$, we can thus give a score penalty, and $w_l(t_l) > 1$ gives a bonus to piece t_l at site l . An assignment $w_l(t_l) = 1$ is neutral and gives neither bonus nor penalty. Observe that *not* including a piece in the piece set \mathcal{T}_l is equivalent to giving it a weight of 0. It is easy to see how this specialises to our tiling examples: in case of the periodic tiling and for l a plaquette interaction in the bulk, the sets \mathcal{T}_l would all be identical and correspond to the allowed 4-local tiles. The w_l then assign the bonuses or penalties, accordingly.

We formulate the following lemma.

Lemma 4.1. *Define a Hilbert space $\mathcal{H} := \bigotimes_{v \in V} \mathbb{C}^c$ over the interaction graph \mathcal{G} , assigning c -dimensional qudits to each vertex $v \in V$. Then there exists a classical Hamiltonian \mathbf{H} on \mathcal{H} , diagonal in the computational basis, with \mathcal{L} -local interactions such that the eigenvalue λ for a basis state $|\psi\rangle = \bigotimes_{v \in V} |c_v\rangle$ is given by the score of the associated generalised tiling, i.e.*

$$\lambda = \sum_{l \in \mathcal{L}} \begin{cases} 1 - w_l(t_l) & \text{if } |\psi\rangle|_l \in \mathcal{T}_l \\ 1 & \text{otherwise.} \end{cases}$$

We denote with $|\psi\rangle|_l$ the restriction of $|\psi\rangle$ to the subspace $\bigotimes_{v \in l} \mathbb{C}^c \leq \mathcal{H}$.

Proof. Define

$$\mathbf{H} := \sum_{l \in \mathcal{L}} \left(\mathbb{1} - \sum_{t \in \mathcal{T}_l} w_l(t) \Pi_t \right),$$

where $\Pi_t := \bigotimes_{v \in l} |t_v\rangle\langle t_v|$ denotes the projector onto the valid piece $t \in \mathcal{T}_l$ for interaction $l \in \mathcal{L}$, and t_v denotes the colour of vertex v for piece t . Take a computational basis state $|\psi\rangle = \bigotimes_{v \in V} |c_v\rangle$. Then

$$\begin{aligned} \mathbf{H} |\psi\rangle &= \sum_{l \in \mathcal{L}} \left(|\psi\rangle - \sum_{t \in \mathcal{T}_l} |\psi\rangle \begin{cases} w_l(t_l) & \text{if } |\psi\rangle|_l \in \mathcal{T}_l \\ 0 & \text{otherwise} \end{cases} \right) \\ &= \lambda |\psi\rangle, \end{aligned}$$

and the claim follows. □

This allows us to conclude the following corollary.

Corollary 4.2. *The ground state energy of \mathbf{H} is determined by the lowest score assignment of the associated generalised tiling problem.*

Equipped with this machinery, it suffices to formulate generalised tiling problems on the square lattices as in fig. 4.1 with 3- and 4-local interactions, such that for lattice sizes below some threshold, the lowest score assignment has a score $\leq -1/2$, and above the threshold the lowest score assignment has a score ≥ 1 . This way, combining the Toric Code Hamiltonian \mathbf{H}_{TC} via lemma 4.3 creates a model with a size-induced transition from classical to topological ground state. Observe that we require our model to be translationally invariant.

4.1.2 The Toric Code

The Toric Code Hamiltonian \mathbf{H}_{TC} is a sum of 3- and 4-local interactions

$$\mathbf{H}_{\text{TC}} := -J \sum_s \mathbf{A}^{(s)} - J \sum_p \mathbf{B}^{(p)},$$

with $\mathbf{A}^{(s)} := \prod_{i \in s} \sigma_i^x$ a product of Pauli σ^x acting on 4 spins i adjacent to vertex s as seen in fig. 4.1. The $\mathbf{B}^{(p)} := \prod_{i \in p} \sigma_i^z$ are defined analogously. We call the $\mathbf{A}^{(s)}$ *star* and the $\mathbf{B}^{(p)}$ *plaquette*-interactions, respectively. The free parameter $J > 0$ is a coupling strength and can be used to rescale the spectrum.

4.1.3 Combining Hamiltonian Spectra

Lemma 4.3. *Let \mathbf{H}_1 and \mathbf{H}_2 two local Hamiltonian defined on $\bigotimes_{u \in \Lambda} \mathbb{C}^{d_1}$ and $\bigotimes_{u \in \Lambda} \mathbb{C}^{d_2}$ for some interaction graph Λ . Let further $\mu \in \mathbb{R}$. Then there exists a Hamiltonian \mathbf{H} on $\mathcal{H} = \bigotimes_{u \in \Lambda} \mathbb{C}^{d_1} \oplus \mathbb{C}^{d_2}$ with the following properties:*

1. *Any eigenvector v of \mathbf{H} with eigenvalue $\lambda \leq \mu$ is given by an eigenvector of either \mathbf{H}_1 or \mathbf{H}_2 , extended canonically to the larger Hilbert space \mathcal{H} , with the same eigenvalue λ .*
2. *\mathbf{H} is translationally invariant if \mathbf{H}_1 and \mathbf{H}_2 are.*
3. *\mathbf{H} contains nearest neighbour interactions and otherwise leaves the interaction range of \mathbf{H}_1 and \mathbf{H}_2 intact.*

Proof. Let $\mathbb{1}_1$ and $\mathbb{1}_2$ be the identity operators on \mathbb{C}^{d_1} and \mathbb{C}^{d_2} , respectively. Let $\delta := 1 + \mu$. Define further

$$\mathbf{H}_0 := \delta \sum_{i \sim j} \mathbb{1}_1^i \otimes \mathbb{1}_2^j + \mathbb{1}_2^i \otimes \mathbb{1}_1^j,$$

where $i \sim j$ denotes any neighbouring spin pairs. Set $\mathbf{H} := \mathbf{H}_0 + \mathbf{H}'_1 + \mathbf{H}'_2$, where $\mathbf{H}'_1 := \mathbf{H}_1 \oplus \mathbf{0}$ and analogously for \mathbf{H}'_2 .

The last two claims are satisfied by construction. To prove the first point, note that \mathbf{H}_0 , \mathbf{H}'_1 and \mathbf{H}'_2 commute and thus share a common eigenbasis with spectrum $\sigma(\mathbf{H}) = \sigma(\mathbf{H}_0) + \sigma(\mathbf{H}'_1) + \sigma(\mathbf{H}'_2)$. Since $\delta > \mu$, any eigenstate of \mathbf{H} with eigenvalue $\lambda \leq \mu$ thus has to be in the kernel $\ker \mathbf{H}_0 \equiv \text{supp}(\mathbf{H}'_1 + \mathbf{H}'_2) = \text{supp} \mathbf{H}'_1 \sqcup \text{supp} \mathbf{H}'_2$, and the claim follows. \square

4.2 Size-Driven Quantum Phase Transitions

4.2.1 Hamiltonian Construction

For local spin dimension $d > 3$, we construct a local, translationally invariant spin Hamiltonian $\mathbf{H}^{(d)}$ on a 2D square lattice with open boundary conditions, such that there exists a threshold system size $N_d \times N_d$, up to which the ground state of $\mathbf{H}^{(d)}$ is entirely classical (i.e. product in the canonical basis), whereas for larger lattice sizes the ground state is that of the Toric Code. Lower bounds on the maximum possible such *transition threshold* N_d for a given local dimension d are shown in table 4.1. For $d > 6$, we give a general procedure for constructing models for which N_d grows faster than any computable function.

The *Toric Code*—introduced by Kitaev [Kito03]—is defined by a Hamiltonian on a two-dimensional spin-1/2 lattice. It is one of the simplest models exhibiting topological order [Wen13]; [Pac12].

We start out with a finite lattice as shown in fig. 4.1. To every edge marked with a dot, we assign a d -dimensional spin $\mathbb{C}^d = \mathcal{H}_{\text{TC}} \oplus \mathcal{H}_{\text{CL}}$ where $\mathcal{H}_{\text{TC}} = \mathbb{C}^2$ and $\mathcal{H}_{\text{CL}} = \mathbb{C}^{d-2}$, such that the overall Hilbert space on the lattice is a tensor product over all separate spins, i.e.

$$\mathcal{H}^{(d)} = \bigotimes (\mathcal{H}_{\text{TC}} \oplus \mathcal{H}_{\text{CL}}) \cong (\bigotimes \mathcal{H}_{\text{TC}}) \oplus (\bigotimes \mathcal{H}_{\text{CL}}) \oplus \mathcal{H}',$$

where \mathcal{H}' contains all mixed \mathcal{H}_{TC} and \mathcal{H}_{CL} terms.

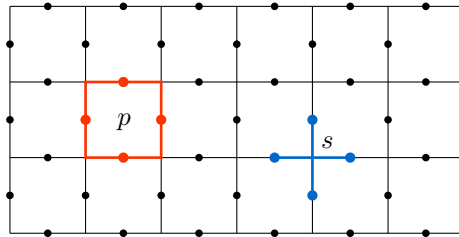


Figure 4.1: Plaquette and star interactions of the two-dimensional Toric Code Hamiltonian \mathbf{H}_{TC} . We assign a spin-1/2 particle to every lattice edge marked with a dot. \mathbf{H}_{TC} is a sum of 4-local interactions, plaquettes and stars, which are products of σ^z and σ^x operators, respectively.

We define a purely classical Hamiltonian \mathbf{H}_{CL} with support only on the subspace $\otimes \mathcal{H}_{\text{CL}}$, such that the ground state energy of \mathbf{H}_{CL} is -1 for lattice sizes $N \leq N_d$, and otherwise $\lambda_{\min}(\mathbf{H}_{\text{CL}}) \geq 1/2$. We then combine \mathbf{H}_{CL} with \mathbf{H}_{TC} in such a way that the spectrum below some energy $\lambda' > 0$ is uniquely determined by one or other of these Hamiltonians, by giving an energy penalty for any state with support on \mathcal{H}' . We define the overall Hamiltonian by $\mathbf{H}^{(d)} := \mathbf{H}_{\text{TC}} + \mathbf{H}_{\text{CL}} + \mathbf{H}'$, where

$$\mathbf{H}' := C \sum_{i \sim j} \mathbb{1}_{\text{TC}}^i \otimes \mathbb{1}_{\text{CL}}^j + \mathbb{1}_{\text{CL}}^i \otimes \mathbb{1}_{\text{TC}}^j,$$

where $i \sim j$ stands for a sum over any adjacent spins. $\mathbb{1}_{\text{TC}}$ denotes the projector on the \mathcal{H}_{TC} subspace, and analogously $\mathbb{1}_{\text{CL}}$. Note that \mathbf{H}' only contains 2-local interactions.

In this way, any state $|\psi\rangle \in \mathcal{H}^{(d)}$ supported on \mathcal{H}' will necessarily pick up an energy penalty of at least C . Choosing $C = 1 + \lambda_{\min}(\mathbf{H}_{\text{CL}})$ shifts this part of the spectrum to energies ≥ 1 . We can rescale \mathbf{H}_{TC} to have its low-energy spectrum within $[0, 1/2]$. The ground state of $\mathbf{H}^{(d)}$ will thus be given by either \mathbf{H}_{TC} or \mathbf{H}_{CL} , whichever has the smaller energy. In particular, the system will change abruptly from classical to topologically ordered with anyonic excitations when the lattice size N surpasses the threshold N_d , while keeping a constant spectral gap.

In order to construct a suitable classical Hamiltonian \mathbf{H}_{CL} , we will exploit the same locality structure as in the Toric Code—4-local star and plaquette interactions—since this does not increase the interaction range of the overall Hamiltonian $\mathbf{H}^{(d)}$. We will only consider the case of open boundary conditions, which is the most natural one in this context.

It is convenient to express the interactions as a so-called *tiling* problem with extra constraints, similar to the well-known Wang tiles. A Wang tile is simply a square tile with coloured edges, and the condition for placing two tiles next to each other is that their edge colours match. Despite this simple setup, it has been shown that the question of whether one can tile the entire plane with a finite set of Wang tiles is in fact undecidable [Ber66], which shows that tiling can encode extremely complex behaviour.

It is easy to represent the tiling problem as a ground state energy problem of a classical, translationally invariant Hamiltonian \mathbf{H}_W on the lattice in fig. 4.1, and straightforward to verify that this representation only defines a single energy scale. As shown in fig. 4.3, each tile can be regarded as a plaquette on the lattice. The condition that neighbouring tiles share the same edge colour is thus automatically met. It is clear that for c colours, we need a c -dimensional classical subspace \mathcal{H}_{CL} for each spin, i.e. $d = c + 2$. Working on this classical subspace, we want to find local Hamiltonian plaquette interactions between the spins surrounding a plaquette p , which we denote with E_p , that penalise any tile not in our set of allowed tiles \mathcal{W} . Specifically,

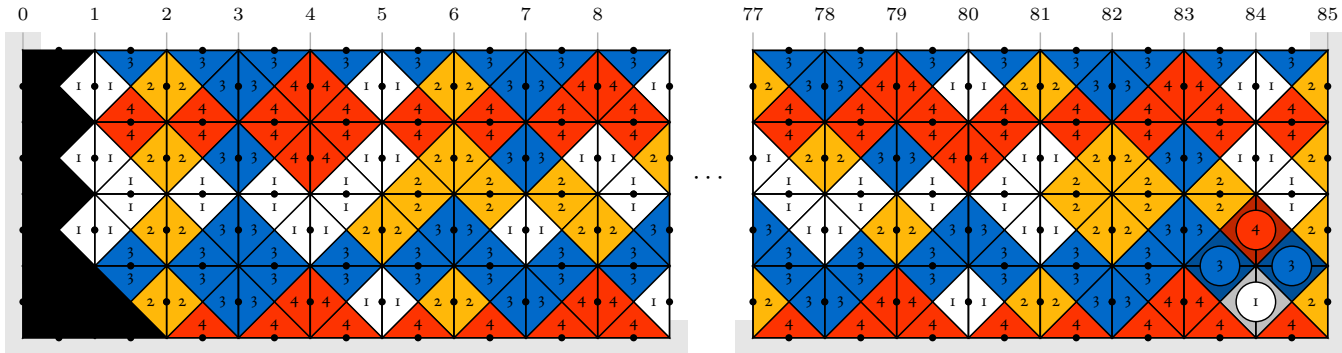


Figure 4.2: Section of the prime periodic pattern for four plus one colours, which is 84-periodic. The left edge and lower corner is enforced by giving the solid black square in the corner a bonus of 1, but penalising black to appear to the right of black on a horizontal edge: this way, the global pattern can be started uniquely with open boundary conditions. The horizontal edge colours form disjoint sets: starting from the bottom row, the colours are red $\{4\}$, blue $\{3\}$, white and yellow $\{1, 2\}$, after which the cycle continues with red. This can be achieved using the 4-local star interactions, e.g. by allowing blue to only appear next to blue and white next to white. For the top row allowing two colours white and yellow, we alternate between them whenever the colour on the vertical edge above it is white. Within each row, these colours on the vertical edge count cyclically through subsequences of length 4, 3 and $4 + 3 = 7$, respectively, which yields the overall horizontal period $\text{lcm}\{4, 3, 7\} = 84$. Every 84 tiles, the pattern necessarily exposes a unique local colour configuration, highlighted in the lower right corner. It can be penalised by a single star interaction and forces the spectrum of the associated Hamiltonian to ≥ 1 when the system size surpasses the threshold $N_5 = 84$.

we define a local classical tile interaction via

$$\mathbf{h}^{(p)} := \sum_{w \in \mathcal{W}} a_w \bigotimes_{e \in E_p} |w_{e,p}\rangle\langle w_{e,p}|, \quad (4.1)$$

where $w_{e,p}$ labels the colour on edge e of tile w placed at plaquette site p . The parameters $(a_w)_{\mathcal{W}}$ do not depend on the plaquette position, and the overall translationally-invariant tiling Hamiltonian is given by the sum over all plaquette sites in the lattice $\mathbf{H}_{\mathcal{W}} := \sum_p (\mathbf{1} - \mathbf{h}^{(p)})$. If $a_w = 1$ for all w , one can show that $\mathbf{H}_{\mathcal{W}}$ has zero energy ground state if and only if the set \mathcal{W} tiles the lattice. If we want to give an energy “bonus” to (i.e. decrease the energy of) a specific tile w , we can set $a_w > 1$. An energy penalty can be given by setting $a_w < 1$. Each tiling thus has a net score—bonuses minus penalties minus mismatching tile pairs. The net score of a specific tiling gives the energy of the corresponding state of $\mathbf{H}_{\mathcal{W}}$. In general, then, the ground state of $\mathbf{H}_{\mathcal{W}}$ will maximise the number of tiles with a bonus while avoiding as many penalties as possible.

A similar construction allows us to add extra star-shaped interactions, constraining tile edges adjacent to a corner. The overall Hamiltonian $\mathbf{H}_{\mathcal{W}} + \mathbf{H}_{\mathcal{S}}$ will then have an optimal ground state in the sense that the sum of penalties minus the sum of bonuses—for both tiles and stars—is minimised. The rigorous argument is presented in lemma 4.1.

Let us now discuss the two families of classical Hamiltonians we construct.

4.2.2 Prime Period Tiling

The key idea is to create a tiling pattern that can tile the entire plane with a very large period p . We require that a certain locally detectable sub-pattern—i.e. using a star interaction—occurs exactly once per period. By disallowing this sub-pattern, the tiling will be possible up to a square of size $p \times p$, but once the grid surpasses this size, there will be at least one pattern violation, which can be penalised locally with a Hamiltonian term.

4.2.2.1 General Construction

For the general construction, we first consider the following discrete optimisation problem. Assume we have q colours available. We want to construct a family of tuples $(r_i)_{1 \leq i \leq f}$, each of which stands for a row of colours $r_i = (r_{i1}, r_{i2}, \dots, r_{im_i})$. These rows have to satisfy three constraints.

1. There are fewer than q rows overall, i.e. $f \leq q$.
2. Each row has fewer than q colours, i.e. $m_i \leq q \forall i$.
3. For the first and last row, each colour r_{ij} is picked from the q colours available, i.e. $r_{1j}, r_{fj} \in \{1, \dots, q\}$ —for all other rows, we leave out the last, i.e. $r_{ij} \in \{1, \dots, q - 1\}$.

We can associate a period $p_i := \sum_j^{m_i} r_{ij}$ to each row i . The rows r_i are now chosen such that the objective function $p(q) := \text{lcm}\{p_i : i \leq I\}$ —i.e. the overall period—is maximised.

We now give a description on how to translate such an optimal row family $(r_i)_i$ into a set of tiles and stars that enforces a unique horizontal tile sequence with periodicity $p(q)$. More specifically, for each row, we define tiles that allow a colour pattern

$$1, \dots, r_{i1}, 1, \dots, r_{i2}, \dots, 1, \dots, r_{i(m_i-1)}, 1, \dots, r_{im_i} \quad (4.2)$$

on their vertical edge—i.e. the tiles form sub-periods, starting at 1 and counting to r_{ij} , then starting at 1 again, counting to $r_{i(j+1)}$ etc.. Cleverly choosing colours on the horizontal edges then a) make this pattern unique for each row, and b) enforce a unique stacking order of the rows overall—which in turn yields a $p(q)$ -periodic global tile pattern.

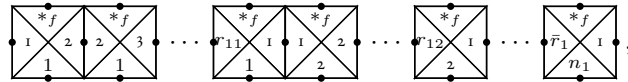
To facilitate the tile notation, we use a few shorthands.

- The last sub-period for each row has highest colour $r_{im_i} =: \bar{r}_i$.
- We sequentially enumerate the sub-periods with colours for use on the horizontal edges, i.e. $r_{11} \leftrightarrow 1, r_{12} \leftrightarrow 2, \dots, \bar{r}_1 \leftrightarrow m_1, r_{21} \leftrightarrow m_1 + 1$ etc.. The highest such label for every row is denoted with h_i , and the lowest l_i , e.g. for the first row $l_1 = 1$ and $h_1 = m_1$. More rigorously, we have the sequences $h_0 = 0, h_i = h_{i-1} + m_i$ and $l_i = h_{i-1} + 1$.
- The set of colours on the horizontal edges needed for the i^{th} row is denoted with $V_i := \{l_i, \dots, h_i\}$, respectively.

For every row r_i , we then define the tiles

$$\begin{array}{c} \begin{array}{|c|c|} \hline \bullet & \bullet \\ \hline c & c' \\ \hline \bullet & \bullet \\ \hline \end{array} & \begin{array}{l} \forall t \in V_{(i-1 \bmod f)}, \quad b \in V_{(i \bmod f)} \\ \forall c = 1, \dots, r_{ij}, \end{array} \end{array}$$

where $c' = c + 1 \pmod{r_{ij}}$. As an example, consider the first row with $V_1 = \{1, \dots, m_1\}$. We obtain a set of tiles



where $*_f$ stands for any colour allowed on the bottom of the last row, i.e. $*_f \in V_f$. All other rows are defined analogously, where the top and bottom colours are chosen successively, i.e. for the i^{th} row, we use colours V_i for the bottom and any V_{i-1} for the top.

On their own, the tiles from different sets can be mixed at will. To enforce that each row can only be

assembled from its own tile set, for each row i , we restrict to the following star configurations:

$$\begin{array}{c}
 \begin{array}{c}
 \circ \\
 \text{c} \\
 \circ \\
 \text{j} \quad \text{---} \quad \circ \\
 \text{j} \quad \text{---} \quad \text{j} \\
 \circ \\
 *
 \end{array}
 \quad 2 \leq c \leq r_{ij}, \quad j \in V_i, \\
 \\
 \begin{array}{c}
 \circ \\
 1 \\
 \circ \\
 \text{j} \quad \text{---} \quad \circ \\
 \text{j} \quad \text{---} \quad \text{j}' \\
 \circ \\
 *
 \end{array}
 \quad \text{and} \quad
 \begin{array}{c}
 \circ \\
 1 \\
 \circ \\
 \text{j}' \quad \text{---} \quad \circ \\
 \text{j}' \quad \text{---} \quad \text{j} \\
 \circ \\
 *
 \end{array}
 \quad \forall j \in V_i
 \end{array}$$

where $j' := l_i + (j + 1 \pmod{m_i})$. It is easy to verify that each row defines a unique p_i -periodic horizontal tile pattern as in eq. (4.2). As the top colours of the i^{th} row are restricted to the bottom colours of the $i - 1^{\text{th}}$ row—modulo the numbers of rows f —the rows can be stacked above each other uniquely. Every block of rows r_1, \dots, r_f , stacked vertically, thus defines a valid horizontally p -periodic tiling pattern for the plane.

In order to be able to detect this periodicity locally, we make use of the extra colour available in all but the first and last row due to constraint (3). For all $i = 2, \dots, f - 1$, we add two tiles

$$\begin{array}{c}
 \bullet \quad \bullet \\
 \diagdown \quad \diagup \\
 \bar{r}_i \quad d \\
 \diagup \quad \diagdown \\
 \bullet \quad \bullet
 \end{array}
 \quad \text{and} \quad
 \begin{array}{c}
 \bullet \quad \bullet \\
 \diagdown \quad \diagup \\
 d \quad 2 \\
 \diagup \quad \diagdown \\
 \bullet \quad \bullet
 \end{array}$$

Alternative to the row sequence $\dots, \bar{r}_i - 1, \bar{r}_i, \mathbf{1}, 2, \dots$, this allows counting $\dots, \bar{r}_i - 1, \bar{r}_i, \mathbf{q}, 2, \dots$. By adding the star penalties

$$\begin{array}{c}
 \circ \\
 1 \\
 \circ \\
 h_1 \quad \text{---} \quad \circ \\
 \circ \quad \text{---} \quad 1 \\
 \circ \\
 1
 \end{array}
 \quad \text{and} \quad
 \begin{array}{c}
 \circ \\
 q \\
 \circ \\
 h_i \quad \text{---} \quad \circ \\
 \circ \quad \text{---} \quad l_i \\
 \circ \\
 1
 \end{array}
 \quad \forall 2 < i < f - 1,$$

we ensure that whenever two consecutive rows complete a cycle in the same column, we mark the occurrence with a q instead of a 1. This way, if in the first row we finish a cycle with a 1 and observe a q right below, we know that the entire horizontal pattern has completed one period. To be more specific, every $p(q) = \text{lcm}\{p_1, \dots, p_K, p_L, p_f\}$ tiles, where $L = f - 1$ and $K = f - 2$, we have the pattern

$$\begin{array}{c}
 \bullet \quad \bullet \quad \bullet \quad \bullet \\
 \diagdown \quad \diagup \quad \diagdown \quad \diagup \\
 \bar{c}_L \quad q \quad q \quad 2 \\
 \diagup \quad \diagdown \quad \diagup \quad \diagdown \\
 \bullet \quad \bullet \quad \bullet \quad \bullet \\
 \bullet \quad \bullet \quad \bullet \quad \bullet \\
 \diagdown \quad \diagup \quad \diagdown \quad \diagup \\
 \bar{c}_f \quad 1 \quad 1 \quad 2 \\
 \diagup \quad \diagdown \quad \diagup \quad \diagdown \\
 \bullet \quad \bullet \quad \bullet \quad \bullet
 \end{array}, \quad \text{locally detectable via} \quad
 \begin{array}{c}
 \circ \\
 q \\
 \circ \\
 h_L \quad \text{---} \quad \circ \\
 \circ \quad \text{---} \quad l_L \\
 \circ \\
 1
 \end{array}.$$

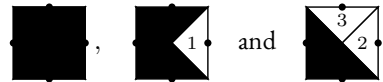
We call this sub-pattern a *period marker* and penalise it with a weight of 2.

So far we have constructed a tile set which can periodically tile the entire plane. By disallowing said period marker, we restrict the tileable square size to at most $p \times p$. Observe however that due to the freedom

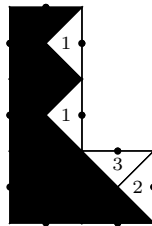
colours $q + 1$	line periods	overall period p
2^\dagger	2, 2	4
4	3, 5	15
5	4, 3, 7	84
6	5, 4, 3, 7	420
7	6, 5, 7, 11	2310
8	7, 6, 5, 11, 13	30030
9	8, 7, 11, 13, 15	120120

Table 4.2: Maximum tiling periods for a given number of colours (plus one special black colour needed for $q > 2$). The second column shows how to distribute the periods between the lines. † For 2 colours, it is easy to see that the extra black tile is redundant.

to shift sets of rows horizontally and the entire pattern vertically, there are a potentially huge number of possibilities to tile any square smaller than $p \times p$. We will thus add a special colour to fix this freedom, borrowing an idea from [GI09]. We will enforce a specific pattern in the lower left corner, which uniquely fixes the starting configuration for the bulk, without imposing any boundary condition, but instead by adding bulk interactions which will have the effect of favouring the desired configuration in the boundary. We add the following tiles:



We further disallow black appearing to the right of black using a star constraint, and give a bonus of $1/2$ to the all-black tile. It is then easy to verify that the best score tiling starts with the following configuration in the lower left corner:



It is straightforward to verify that starting from this corner configuration, the plane can be tiled uniquely up to a grid size of $p \times p$ with a net score of $-1/2$, after which the net penalty jumps to a value ≥ 1 .

In table 4.2, we list the cases $q = 2, \dots, 8$ with a solution to the associated constraint problem and the resulting overall period p . It is easy to see that in the case of 2 colours only, the extra black tile to remove degeneracies is redundant.

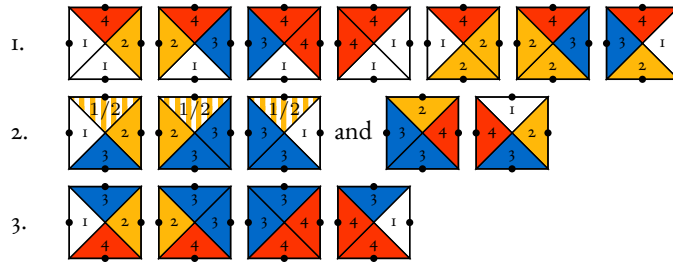
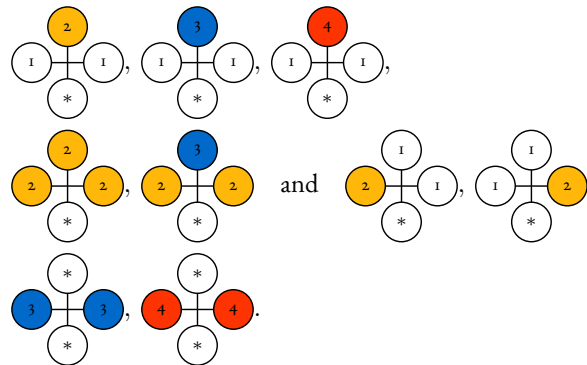


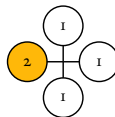
Figure 4.3: Four+1 colour tile set that defines a tiling of the plane with period 84. The striped top in the second row denotes either of the colours 1 or 2.

4.2.2.2 Five Colour Tiling Example

We give the $q = 4$ colour case as an explicit example. The tile set in fig. 4.3 defines three disjoint sets of tiles, each of which can be assembled into horizontal lines, which in turn can be stacked above each other in a unique order. To avoid mixing the tiles from different sets on one same line, we add the following star operators with parameter $b_s = 1$



The third row then unambiguously assembles to a line with a horizontal period of 4, while the vertical edges appearing on the first line periodically cycle through 1, 2, 3, 4, 1, 2, 3 with a period of 7. The horizontal edges of the second line are fixed by the line above and below, but there exists some freedom to choose the colours on the vertical edge. More specifically, we use the freedom of either counting $\dots, 2, 3, 1, 2, \dots$ or $\dots, 2, 3, 4, 2, \dots$ to detect when all three lines complete a period in the same column. We add a penalty for the configuration

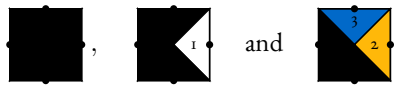


by adding the corresponding operator multiplied by $b_s = -1$, which enforces colour 4 to appear instead of colour 1 whenever the first row finishes one cycle at the same time as the second one below. The combined period p of the three lines is thus given by $\text{lcm}(7, 4, 3) = 84$, and it can be detected by penalising the configuration



with a penalty of 2 (i.e., with $b_s = -2$.)

The freedom to horizontally shift the lines relative to each other or the entire pattern vertically is fixed without adding boundary conditions with the special tiles



We choose $a_w = 2$ for the first. Starting from there, the entire plane can be tiled uniquely up to a grid size of 84×84 , after which the penalised star—eq. (4.3)—naturally occurs and the net penalty is ≥ 1 . A section of the complete 5-colour tiling can be seen in fig. 4.2.

Generalising the prime tiling to higher dimensions, we obtain the periods as given in table 4.2.

4.2.3 Turing Machine Tiling

Starting from a number of colours $c \geq 6$, it becomes possible to embed a Turing machine into a set of tile and star interactions. We improve on an idea introduced by Robinson [Rob71a]—which has been exploited in [CPW15b] to show undecidability of the spectral gap—by making use of the extra star constraints to significantly reduce the necessary local dimension. In this new construction, the transition threshold N_d grows faster than any computable function and surpasses the periodic tiling bound for $c \geq 7$.

We have introduced Turing machines in section 2.2.1, but we want to repeat the most important points here. A Turing machine is given by finite sets of states Q and symbols A with a transition function $\delta : Q \times A \rightarrow Q \times A \times \{\text{left}, \text{right}\}$, representing the set of instructions of the Turing machine. The machine is equipped with a tape, which is sequence of symbols arranged in a 1-dimensional line extending indefinitely in both directions, and initialised with a special “blank” symbol (which we will denote 0 for simplicity of notation). The machine has an internal state $q \in Q$ and a head which sits over one of the symbols of the tape: at each step, the head reads the symbols s underneath, and it will write the symbol \bar{s} , change its internal state to \bar{q} and then move in direction $d \in \{\text{left}, \text{right}\}$, where $(\bar{q}, \bar{s}, d) = \delta(q, s)$.

The machine starts in an initial state $q_0 \in Q$ and *halts* if there is no forward transition for a given tuple

states $ Q $	colours c	$S(Q)$	threshold N_d
3	6	21 [LR65]	14
4	6	107 [Bra83]	75
5	7	$\geq 4.7 \cdot 10^7$ [Mar+90]	$3.3 \cdot 10^7$
6	8	$\geq 7.4 \cdot 10^{36534}$ [Kro]	$5.2 \cdot 10^{36534}$

Table 4.3: Number of 2 symbol Turing machine states Q and tile colours $c = \max\{|A|^2 + 2, |A| + |Q|\}$ required for the embedding. $S(|Q|)$ is the Busy Beaver number, and N_d denotes the lower bound on the maximum lattice threshold size, where $d = c + 2$.

(q, s) ¹. We say a given machine is *halting* if it eventually reaches a halting state. If we restrict to machines with a fixed number of states q for its internal memory, and which read and write only two symbols, i.e. 0 and 1, then the set of possible halting machine programs is finite: there has to exist one that runs for longer, or at least as long as, any other. These machines are called *Busy Beavers*, and their running time is called the *Busy Beaver number* $S(q)$. It is known that $S(q)$ grows faster than any computable function [Rad62]².

As in the case of the periodic tiling, we find a way of embedding a Busy Beaver Turing machine into the ground state of a classical Hamiltonian: as soon as the Busy Beaver halts, there will be a penalty, since at that point there is no valid way to continue updating the tape. The tiling is thus possible up to a square size of at least $S(q)/\sqrt{2}$ ³. As we need $c = q + 2$ colours for a q state Busy Beaver, we immediately find a transition threshold of $N_{q+4} \geq S(q)/\sqrt{2}$. We will show how to construct a set of plaquette and star interactions in such a way that the ground state encodes the history of a run of a given Turing machine. The construction will involve the use of Wang tilings and of star interactions: the latter will allow us to greatly reduce the spin dimension needed by previous works which were based only on the tiling problem [Rob71a]; [Ber66].



We assume the lattice is oriented as in fig. 4.4, where the Turing machine starts in the lower left corner. We chose to encode the tape of the Turing machine at a given time step in diagonal direction across the square lattice (denoted by the thick grey lines in fig. 4.4), and movement in the orthogonal direction represents the time evolution. Moreover, we store the head position and internal state of the Turing machine by including the state q in the tape, on the right of the symbol which will be read by the machine. Using this convention, we have that—even if the tape space is finite—it is extended by one symbol in both directions at each time step, and therefore the tape available will always be sufficient for the machine to run.

We will interpret colours on horizontal and vertical edges differently—horizontal either as pair of symbols

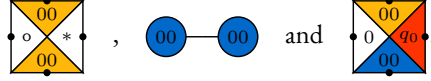
¹Such definition is equivalent to defining a special halting state h for which no further transition is defined

²A related quantity, which is also called the Busy Beaver number but is usually denoted $\Sigma(q)$, is defined as the largest number of non-blank symbols written out by the machine before terminating, and is a lower bound on $S(q)$. $\Sigma(q)$ also grows faster than any computable function.

³The constant factor of $\sqrt{2}$ is due to the fact that in our construction the tape of the Turing machine is encoded diagonally with respect to the square lattice, and the head of the machine follows a zig-zag pattern, which in the worst case scenario can only reach a distance of $S(q)/\sqrt{2}$ from the origin.

$s_1 s_2 \in A \times A$ or special boundary colour  or , and vertical either as symbol $s \in A$ or state $q \in Q$, the latter of which we highlight in red.




As we did in the periodic tiling construction, we use special bulk interaction (which will be present everywhere in the lattice) to constrain the left and bottom boundaries. In order to do so, we use the following interactions



Note that the 2-local blue term is indeed a bulk interaction, and not a “cut-off” boundary term. By giving the last of these tiles a single bonus of 2—similar to the all black tile for the periodic tiling—we obtain precisely one choice for the left and lower grid edges, namely all 0s as in fig. 4.4; in particular, the last shown tile with initial state q_0 correctly initialises the Turing machine in the lower left corner. This valid initial configuration defines the unique highest-net-bonus tiling possible.

To avoid cases where we validly tile the plane without a TM head—i.e. with net bonus zero—we use a star interaction to give a bonus of $1/2$ for any *white* symbol on a vertical edge appearing to the left or right of another arbitrary symbol, i.e.

$$1 \times \textcircled{**} - \textcircled{**} \quad , \quad \frac{1}{2} \times \textcircled{**} - \textcircled{00} \quad \text{and} \quad \frac{1}{2} \times \textcircled{00} - \textcircled{**}.$$

We further give a penalty of 1 for the white symbol appearing anywhere. This way, in the bulk, the net contribution of $2 \times 1/2 - 1 = 0$ for each of the white edges, whereas if they appeared on the left end of the plane a net penalty $\geq 1/2$ would be inflicted. A similar combination of bonus and penalty terms allows us to ensure that the lower edge is blue, and all other configurations obtain a net penalty of $\geq 1/2$ as well. Like that, there exists no configuration with net penalty $< 1/2$ without the initial q_0 tile in the lower left corner. From now on, we treat the boundary symbols  and  as equivalent to .

To implement the transitions rules we effectively need 6 different spins to interact (three for each time step): in fact, if the tape around the head reads s, q, t for some $q \in Q$ and some $s, t \in A$, then it has to be updated to \bar{q}, \bar{s}, t if $\delta(q, s) = (\bar{q}, \bar{s}, \text{left})$, while it has to be updated to \bar{s}, t, q if instead $\delta(q, s) = (\bar{q}, \bar{s}, \text{right})$.

Since we only have at our disposition 4-body interactions, in order to implement an effective 6 body interaction we will make use of the extra register we have allowed for in the horizontal edges, which will allow to “synchronise” a plaquette and a star interaction, as shown in fig. 4.4. This is done by defining, for every transition, a pair of tiles and stars, i.e. if $\delta(q, s) = (\bar{q}, \bar{s}, \text{left})$

(4.4)

where $*$ represents any symbol, and for an analogous right transition

$$(4.5)$$

Observe how the symbol pairs are necessary to uniquely couple the pair of interactions to obtain the left and right transition depicted in fig. 4.4, but are disregarded for any successive transition. The rest of the tape which is not affected by the transition rules has to be copied verbatim to the next time step. Implementing such bookkeeping tiles is straightforward, as we only need to take care of the extra—and in this situation unused—register in the horizontal edges, which is discarded when copying to vertical edges, and set to the “blank” symbol when copying from vertical to horizontal edges. More precisely, for every $a, b \in A$, we define

$$(4.6)$$

Overall, this construction thus requires $c = \max\{|A|^2 + 2, |A| + |Q|\}$ colours. It is easy to verify that starting from the initial tile, a square can be uniquely tiled with net bonuses 1 if and only if the Turing machine does *not* halt within its boundaries. All other tilings necessarily violate at least one constraint and thus have a net penalty $\geq 1/2$. A sample evolution can be seen in fig. 4.4.

The maximum number of steps any *halting* Turing machine with $|Q|$ states and 2 symbols can take before halting is called the *Busy Beaver number* and is denoted by $S(|Q|)$. Defined in [Rad62], it is known to grow faster than any computable function. The staggering threshold sizes N_d in table 4.1 show that there is no hope to address the question of extrapolating physical properties of a general system solely with an increase in computational power.

4.2.4 Thermal Stability

There exists a finite temperature T_d below which the thermal state of the Hamiltonian will still be very close to the ground state for any system size up to the threshold N_d , meaning that any measurement will still reveal a classical state up to very small errors. The temperature T_d depends only inverse-logarithmically on the threshold size, and therefore its scaling with d will be mild in the case of the prime periodic tiling. Table 4.1 lists the temperatures corresponding to the various local dimensions d , which is a linear function of Δ/k_B , where Δ is the spectral gap of the Hamiltonian and k_B the Boltzmann constant. Since our models are commuting Hamiltonians, the spectral gap Δ is simply equal to the strength of the interactions between the spins.

For the prime periodic tiling, we also show that if we go to the thermodynamic limit at finite temperature,

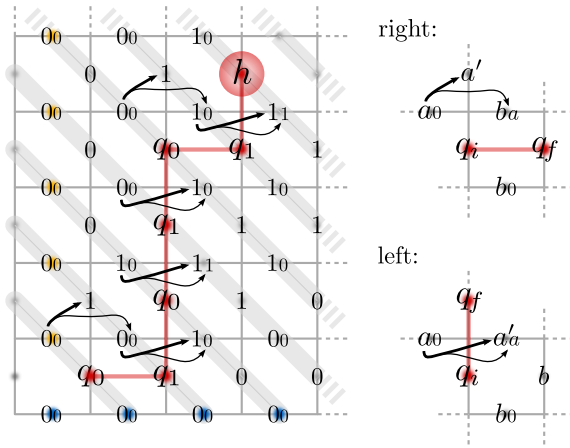
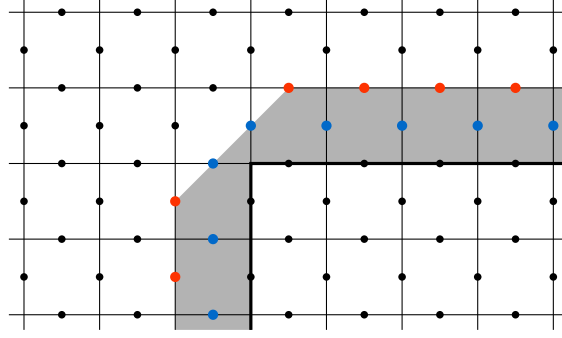


Figure 4.4: Embedding of a Turing machine into a tiling problem with extra star constraints. We chose a representation in which the Turing machine head sits in between the tape symbols, reading and writing the symbol on its left. Every grey horizontal slice shows the tape at one evolution step: it is initialised to $\dots, 0, q_0, 0, \dots$, where q_0 is the initial machine state, and every successive step is uniquely defined by the transition rules. Shown here is the 2 state Busy Beaver which halts after 6 steps. Since there is no valid tile with a halting state, the system necessarily frustrates for lattices larger $6/\sqrt{2} \approx 4$ tiles on each edge. Each right transition $(q_i, a) \mapsto (q_f, a', \text{right})$ or left transition $(q_i, a) \mapsto (q_f, a', \text{left})$ is translated into a pair of 4-local plaquette and star interactions, as depicted to the right. Observe how in both cases the tile part of the interaction has to know the initial symbol a , which is why the star operator creates a temporary copy of it. This copy is shown as small symbols and numbers to the right of the actual tape content and ignored in any following transition. As the available space grows by one symbol in both directions at each step, there is always enough tape available for the Turing machine. The coloured terms are used to initialise this spare tape to 0. Away from the head, additional interactions are used to copy currently unused tape segments forward. The exact construction with all interaction terms is explained in detail in the appendix.



and then send the temperature to zero (a procedure which is a more mathematically correct description of real implementations [AL81]), we will recover only ground states of the Toric Code. This shows that there is a complete disagreement between the mathematical predictions from the thermodynamic limit, and any measurement performed on systems below the threshold size. The Busy Beaver construction can be modified in order to show the same property, using the original construction of Robinson [Rob71b], at the cost of greatly increasing the local dimension.

Let us be more precise, and recall that a state in the thermodynamic limit is given by a linear, positive and normalised functional ω on the algebra of quasi-local observables \mathcal{A} , which is the (norm closure of the) inductive limit of the finite matrix algebras $\mathcal{B}_\Lambda = \mathcal{B}(\otimes_{i \in \Lambda} \mathcal{H}_i^{(d)})$, where Λ is an ascending sequences of finite lattices converging to \mathbb{Z}^2 [BR97]; [AL81].

Given a local Hamiltonian \mathbf{H} and a finite region Λ , we define its (exterior) boundary $\partial\Lambda$ as the set of sites in the complementary of Λ for which there is an interaction term in \mathbf{H} acting nontrivially on sites of Λ and $\partial\Lambda$ simultaneously, as shown in section 4.2.4. $\bar{\Lambda}$ is defined as $\Lambda \cup \partial\Lambda$ and, for a region R , \mathbf{H}_R will denote the restriction of \mathbf{H} to all interactions which are totally contained in R . A *ground state* is then defined as a state functional ω , such that for any finite Λ , and any local observable $A \in \mathcal{B}_\Lambda$,

$$\omega(A^\dagger[\mathbf{H}_{\bar{\Lambda}}, A]) \geq 0. \quad (4.7)$$

This definition can be obtained by taking the zero temperature limit in the definition of finite temperature equilibrium states as defined by the KMS condition (i.e. the limit of increasing-volume Gibbs ensembles satisfying the KMS condition, cf. [HHW67, eq. 4.2]). Loosely speaking, it expresses the intuitive understanding that any local perturbation should not decrease the energy of a ground state (cf. [CNN16]).

Note that since both A and $\mathbf{H}_{\bar{\Lambda}}$ have finite support, it is possible to rewrite eq. (4.7) in terms of the reduced density matrix of ω over $\bar{\Lambda}$, which we denote by $\rho_{\bar{\Lambda}}$:

$$0 \leq \omega(A^\dagger[\mathbf{H}_{\bar{\Lambda}}, A]) = \text{tr}(\rho_{\bar{\Lambda}} A^\dagger[\mathbf{H}_{\bar{\Lambda}}, A]),$$

or equivalently

$$\mathrm{tr}(\rho_{\bar{\Lambda}} A^\dagger \mathbf{H}_{\bar{\Lambda}} A) \geq \mathrm{tr}(\rho_{\bar{\Lambda}} A^\dagger A \mathbf{H}_{\bar{\Lambda}}) \quad (4.8)$$

for all Λ and all $A \in \mathcal{B}_\Lambda$. In turn, this implies that

$$\mathrm{tr}(\Phi(\rho_{\bar{\Lambda}}) \mathbf{H}_{\bar{\Lambda}}) \geq \mathrm{tr}(\rho_{\bar{\Lambda}} \mathbf{H}_{\bar{\Lambda}}), \quad (4.9)$$

for any completely positive, trace preserving linear map Φ , as can be seen by applying eq. (4.8) to the Kraus operators of $\Phi(\cdot) = \sum_i A_i \cdot A_i^\dagger$.

We will argue that the only ground states of the periodic tiling Hamiltonian $\mathbf{H}^{(d)}$ are the ground states of the Toric Code. This in turn implies, that if we take first the limit of N going to infinity, and then we send the temperature to zero, we recover only ground states of the Toric Code.

Key to our argument is that part of our Hamiltonian—i.e. \mathbf{H}_0 —is a ferromagnetic Ising-type interaction, where spin up and down are now the tiling and Toric code subspaces, respectively. We will follow the same proof technique used to show that the 2D Ising model with an external magnetic field has a unique ground state [AL81, ex. 5] to show that any ground state in the thermodynamic limit of our model is completely in the Toric code subspace, by which we mean that for all Λ , $\omega(\Pi_{\mathrm{TC},\Lambda}^\perp) = 0$ for the projector onto the Toric code subspace $\Pi_{\mathrm{TC},\Lambda}$ supported on Λ .

Let us start with some preliminary observations, which will allow us to assume some extra properties of the ground state without loss of generality. Fix Λ and let $\{M_i\}_i$ a decomposition of the identity on $\bar{\Lambda}$ into orthogonal projectors, such that $[M_i, \mathbf{H}_{\bar{\Lambda}}] = [M_i, \Pi_{\mathrm{TC},\Lambda}^\perp] = 0$ for all i . We want to show that is sufficient to study ω restricted to the subspace corresponding to each M_i . This is the content of the following lemma.

Lemma 4.4. Let ω be a ground state. Fix Λ and let $\{M_i\}_i$ as above. Whenever $\omega(M_i) \neq 0$, denote by

$$\omega_i(A) = \frac{\omega(M_i A M_i)}{\omega(M_i)}.$$

Then ω_i is also a ground state. Moreover, if for every i it holds that $\omega_i(\Pi_{\mathrm{TC},\Lambda}^\perp) = 0$, then also $\omega(\Pi_{\mathrm{TC},\Lambda}^\perp) = 0$.

Proof. ω_i is clearly a positive linear functional on local observables so that $\omega_i(\mathbb{1}) = 1$. It can then be extended to a state on \mathcal{A} . The fact that M_i commutes with the Hamiltonian makes ω_i fulfil trivially eq. (4.7), so it is a ground state. Finally, we observe that

$$\begin{aligned} \omega(\Pi_{\mathrm{TC},\Lambda}^\perp) &= \mathrm{tr}(\rho_{\bar{\Lambda}} \Pi_{\mathrm{TC},\Lambda}^\perp) = \sum_1 \mathrm{tr}(M_i \rho_{\bar{\Lambda}} \Pi_{\mathrm{TC},\Lambda}^\perp M_i) = \\ &= \sum_i \mathrm{tr}(M_i \rho_{\bar{\Lambda}} M_i \Pi_{\mathrm{TC},\Lambda}^\perp) = \sum_i \omega(M_i) \omega_i(\Pi_{\mathrm{TC},\Lambda}^\perp), \end{aligned}$$

so that the last claim of the lemma follows. \square

We will use such lemma to make two extra assumptions. The first one allows to assume that the ground state is supported, in each site, only in one of the two subspaces (TC or tiling). For that, given a finite region $R \subset \mathbb{Z}^2$ we consider signatures $\sigma = (\sigma_i)_{i \in R}$ where each $\sigma_i \in \{\text{TC}, \text{tiling}\}$. We denote by P_σ the projector onto the set of states of signature σ . It is easy to see that they satisfy the condition of lemma 4.4. The second assumption is that $\rho_{\bar{\Lambda}}$ commutes with the Toric Code stabilisers. Again, it is sufficient to consider the projectors onto the eigenspaces of such stabilisers, and the result follows from lemma 4.4.

As a second step, we will show that for any ground state for which a square boundary is completely supported in the TC subspace, the interior will be as well; for this we will assume that all square regions have smooth edges as in section 4.2.4.

Lemma 4.5. Take two concentric square regions $\Lambda' \subsetneq \Lambda$, and a ground state ω of $\mathbf{H}^{(d)}$ with a signature σ on $\bar{\Lambda}$. Assume that $\sigma_s = \text{TC}$ on all sites s of $\partial\Lambda' \subset \Lambda \setminus \Lambda'$. Moreover, assume that $\rho_{\bar{\Lambda}}$ commutes with the Toric Code stabilisers that couple Λ' with $\partial\Lambda'$. Then $\sigma_s = \text{TC}$ all sites $s \in \Lambda'$.

Proof. Denote with $T \subseteq \Lambda'$ the set of all sites $\sigma \in \Lambda'$ that satisfy $\sigma = \text{tiling}$.

Consider the CPTP map Φ_1 acting on T that on *all* those sites, traces out the tiling sector and replaces it with the maximally mixed state on the TC subspace, i.e.

$$\Phi_1(\rho) = \text{tr}_T(\rho) \otimes \left(\frac{\mathbb{1}_T^{(\text{TC})}}{\text{tr} \mathbb{1}_T^{(\text{TC})}} \oplus 0_T^{(\text{tiling})} \right).$$

Let us now consider a map Φ_2 , acting on $\bar{\Lambda}'$, which implements the following operations: first measures the Toric Code projectors which overlap with Λ' , and then, conditioned on the syndrome of the measurement, applies a unitary operator which corrects as many code errors as possible ⁴. This can be constructed by choosing as Kraus operators of Φ_2 the product of the projector onto the different syndrome subspaces multiplied on the left with the corresponding unitary operator. We extend this map on the tiling subspace with the identity map, in order to make it a CPTP map. Then eq. (4.9) implies that

$$\text{tr}(\rho_{\bar{\Lambda}} \mathbf{H}_{\bar{\Lambda}}) \leq \text{tr}(\Phi_2 \circ \Phi_1(\rho_{\bar{\Lambda}}) \mathbf{H}_{\bar{\Lambda}}). \quad (4.10)$$

We now consider $\tilde{\rho}_{\bar{\Lambda}} = \Phi_2 \circ \Phi_1(\rho_{\bar{\Lambda}})$. For any \mathbf{h} whose support is disjoint from Λ' , since $\tilde{\rho}_{\bar{\Lambda}'}$ has the same reduced density matrix as $\rho_{\bar{\Lambda}'}$ outside of Λ' , we have that $\text{tr}(\rho_{\bar{\Lambda}} \mathbf{h}) = \text{tr}(\tilde{\rho}_{\bar{\Lambda}} \mathbf{h})$. Thus eq. (4.10) reduces to $\text{tr}(\rho_{\bar{\Lambda}} \mathbf{H}_{\bar{\Lambda}'}) \leq \text{tr}(\tilde{\rho}_{\bar{\Lambda}} \mathbf{H}_{\bar{\Lambda}'})$, where in $\mathbf{H}_{\bar{\Lambda}'}$ only those Hamiltonian terms appear whose support intersects Λ' . To finish the proof, we need to find a contradiction assuming that T is not empty. First of

⁴One possible way to implement this procedure is to follow a sequence of local steps along an oriented tree structure, as described in [DKP14]

all, notice that $\tilde{\rho}_{\bar{\Lambda}}$ is completely supported on the TC subspace in $\bar{\Lambda}'$, and that can violate at most 2 of the

- Toric Code Stabilisers (at most one plaquette and one star operator, since any pair of violations would have been destroyed by the action of Φ_2). So $\text{tr}(\tilde{\rho}_{\bar{\Lambda}}\mathbf{H}_{\bar{\Lambda}'})$ can at most be equal to 2. On the other hand, even with the bonus gained in bottom-left corners of regions supported on the tiling subspace (i.e. the bonus of $1/2$ for the all-black tile used to resolve the ground state degeneracy for the periodic tiling pattern), the penalties coming from mixed signatures in $\text{tr}(\rho_{\bar{\Lambda}}\mathbf{H}_{\bar{\Lambda}'})$ are higher (an overall penalty of at least $7/2$ for each mismatch). \square

In the next lemma, we generalise the previous one for the case in which some sites on $\partial\Lambda'$ are in the tiling sector.

Lemma 4.6. Take two concentric square regions $\Lambda' \subsetneq \Lambda$, and a ground state ω of $\mathbf{H}^{(d)}$ with a signature σ on $\bar{\Lambda}$. Moreover, assume that $\rho_{\bar{\Lambda}}$ commutes with the Toric Code stabilisers that couple Λ' with $\partial\Lambda'$. Let α be the number of sites $s \in \partial\Lambda'$ for which $\sigma_s = \text{tiling}$, and β the sum of signature mismatches within Λ' —i.e. the number of neighbouring $s, s' \in \Lambda$ for which $\sigma_s \neq \sigma_{s'}$ —plus the number of period markers within Λ' . Then $\beta \leq \frac{4}{7}(1 + 4\alpha)$.

Proof. We follow the same procedure as in the proof for lemma 4.5, obtaining a new state $\tilde{\rho}_{\bar{\Lambda}}$ on $\bar{\Lambda}$, such that

$$\text{tr}(\rho_{\bar{\Lambda}}\mathbf{H}_{\bar{\Lambda}}) \leq \text{tr}(\tilde{\rho}_{\bar{\Lambda}}\mathbf{H}_{\bar{\Lambda}}). \quad (4.11)$$

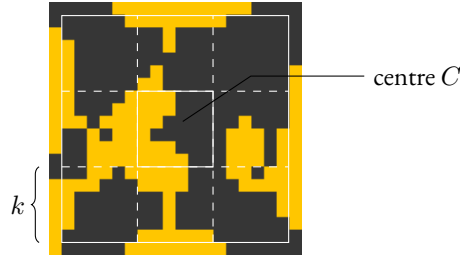
Again, let $T \subset \Lambda'$ the set of sites with tiling signature. Let us consider now the interactions \mathbf{h} in $\mathbf{H}_{\bar{\Lambda}}$ and

- compare the values $\text{tr}(\tilde{\rho}_{\bar{\Lambda}}\mathbf{h})$ and $\text{tr}(\rho_{\bar{\Lambda}}\mathbf{h})$. As in the previous lemma, if \mathbf{h} do not overlap with Λ' , then $\text{tr}(\tilde{\rho}_{\bar{\Lambda}}\mathbf{h}) = \text{tr}(\rho_{\bar{\Lambda}}\mathbf{h})$. Since $\tilde{\rho}_{\bar{\Lambda}'}$ is in the TC subspace, and can violate at most 2 stabilisers, its energy can be at most $2 + 8\alpha$ (the signature in $\partial\Lambda'$ has not changed, and each spin in the tiling subspace can violate up to 4 Ising-type interactions). On the other hand, since there are at least β signature mismatches for $\rho_{\bar{\Lambda}'}$, and each of them has an energy of at least $7/2$ (again, this is lower than 4 because of the $1/2$ bonus given to the all-black tile), we have that $\text{tr}(\rho_{\bar{\Lambda}}\mathbf{H}_{\bar{\Lambda}'}) \geq \frac{7}{2}\beta$. Inserting these two bounds into eq. (4.11), we obtain the desired bound. \square

In the following, we will show that if we pick the outer square in lemma 4.5 large enough, we are bound to find an inner concentric square—of at least a third of the outer square's size—for which we can then apply lemma 4.5 or lemma 4.6. In the pictures of the following lemma, we have coloured with black the spins which are in the TC subspace, and in yellow the ones which are not.

Lemma 4.7. Take some square Λ of side length $3k$, where $k = 10^6 N_d^2$ and consider a ground state ω of

$\mathbf{H}^{(d)}$ with a signature σ on $\bar{\Lambda}$. Subdivide Λ into $k \times k$ squares:



Then $\sigma_i = TC$ for all i in the centre C .

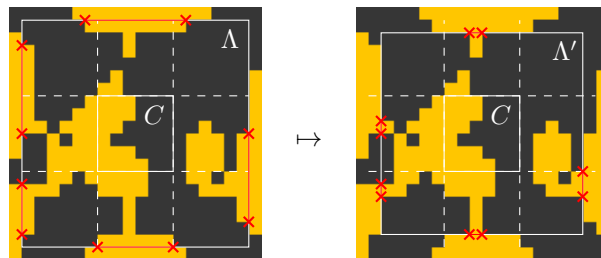
Proof. We start with a few preliminary observations, and we refer the reader to section 4.2.4 for verification. The boundary $\partial\Lambda$ contains precisely $3 \times 3k$ spins on each side, and thus $12 \times 3k = 36k$ spins overall. We denote this spin count with $|\partial\Lambda|$ in the context of this proof. By lemma 4.6 (assuming that for all $s \in \partial\Lambda : \sigma_s = \text{tiling}$), we thus know that we can have at most $\frac{4}{7}(1 + 4 \times |\partial\Lambda|) < 83k$ penalties from signature mismatches or period markers within Λ .

The overall area of Λ encompasses $9k^2$ tiles, and we count $|\Lambda| = 2 \times 9k^2 + 2 \times 3k$ spins. Every sub-square of size $N_d \times N_d$ which is not fully in the TC subspace carries a penalty ≥ 1 —note that this holds true regardless of the bonus terms present in the tiling, as the period penalty and TC-tiling mismatch are both larger. This means that at most a fraction of

$$\frac{83k}{|\Lambda|/N_d^2} = \frac{83N_d^2}{18k + 6} = \frac{83}{18 \times 10^6 + 6/N_d^2} < \frac{1}{10\,000}$$

of spins $s \in \Lambda$ can have signature $\sigma_s = \text{tiling}$. For a subset $A \subset \Lambda$, we denote this fraction with $f(A)$.

So let us assume that $f(\Lambda) < 1/10\,000$. Take Λ and shrink it uniformly by at most $k/10$, by which we mean we shrink the square on each side by one tile at a time, i.e. while keeping smooth boundaries as in fig. 4.1. This sweeps a region A which covers at least $1/10$ of the area of Λ , and thus in particular the number of spins $|A| \geq |\Lambda|/10$. Since $f(\Lambda) < 1/10\,000$, it follows that $f(A) < 1000$. This immediately implies that there exists a square $\Lambda' \subset A$ —concentric with Λ —which satisfies $f(\partial\Lambda') < 1/1000$.



Two things may happen. If in between the centre square C and Λ' there exists another square Λ'' (i.e. with $\partial\Lambda'' \subset \Lambda' \setminus C$) such that for any $s \in \partial\Lambda''$, we have $\sigma_s = \text{TC}$, then lemma 4.5 immediately implies that $\sigma_i = \text{TC}$ for all i in the centre C , and the claim follows.

It remains to analyse the case where no such square Λ'' exists. We first apply lemma 4.6 again, this time to Λ' : as $f(\partial\Lambda') < 1/1000$, and $|\partial\Lambda'| \leq |\partial\Lambda| \leq 36k$, we know that there exist *at most* $36k/1000 \times 8 \leq k/3$ spins $s \in \Lambda'$ with $\sigma_s = \text{tiling}$. But since no Λ'' exists with a boundary $\partial\Lambda''$ completely with tiling signature, there have to be *at least* $k - k/10 = 9k/10$ spins within Λ' with a tiling signature (one within the boundary for each concentric square between Λ' and C). Contradiction, and the claim follows. \square

All the above results lead trivially to

Corollary 4.8. *All ground states of $\mathbf{H}^{(d)}$ are fully supported in the TC subspace.*

4.3 Chapter Summary

By constructing two concrete classes of examples, we have shown that there exist translationally invariant, local Hamiltonians on a 2D square lattice with constant spectral gap and open boundary conditions, which belong to a topologically ordered phase in the thermodynamic limit, but appear to be classical for finite system sizes smaller than a certain threshold. This threshold grows extremely fast as a function of the local spin dimension—for one class it grows faster than any computable function—showing that even for physically realistic systems with low local dimension, erratic behaviour may occur at system sizes that are inaccessible numerically or experimentally. For such systems, physical properties in the thermodynamic limit cannot be extrapolated from sequences of finite-size instances.

The implications of these findings may be profound. Numerical simulations of lattice models play a key role in understanding the dynamics of a system, e.g. in lattice gauge theories [Kog83], fluid dynamics [Yepoi] and condensed matter physics [LLSo1]. All these simulations are extremely computationally intensive, so accessible lattice sizes are severely limited—e.g. for heavy quark simulations, current lattices have sizes reaching $96^3 \times 192$, the long direction being time [Oli+14, Ch. 18]. Our results show that there exist classes of simple, local, physical systems on a lattice of spins with moderate dimension, for which it is impossible to tell with certainty whether the system behaves the same on macroscopic scales as it does on any accessible finite size. In fact, the physical properties of this class of systems will change dramatically above some threshold size, which can even be uncomputable.

Many variations of this problem are possible. It is easy to e.g. reverse our construction and transition from topologically ordered at low system sizes to classical for large lattices, and it is clear that similar constructions using a different Hamiltonian than the Toric Code are possible. As usual when exotic models are found, we

expect that the ability to switch properties of a Hamiltonian on and off depending on the system size could also lead to interesting applications in future.

5 Beyond History States

Wir müssen wissen – wir werden wissen.

—*David Hilbert*

We'd better dream sky high,

Before we lose all hope.

A little here, a little there,

And it's gone before you know.

Because the world is made of inches,

So if you want to fly,

We've got to dream a little bigger,

If we're ever going to build those castles in the sky.

—*Geraint Luff, Sky High*

The initial demonstration of QMA-completeness of the LOCAL HAMILTONIAN problem [KSV02] was followed by a period of development during which the main goal was to broaden the class of interaction terms which suffice to make the LOCAL HAMILTONIAN problem QMA-complete [KKR06]; [OT05]. These results were motivated in part by a desire to understand the hardness of approximating physical systems that resemble those found in nature, and also by the goal of making the closely related universal adiabatic computation construction [Aha+08] more suitable for eventual physical implementation [BL08]. The success of these efforts have resulted in QMA-complete LOCAL HAMILTONIAN problems with restricted properties such as 2-local interactions [KKR06], low dimensional geometric lattices [OT05]; [Aha+09b], and translational invariance, as well as a complete classification of the complexity of the 2-LOCAL HAMILTONIAN problem for any set of interaction couplings [CM16].

This great success in classifying the hardness of physically realistic interactions stands in contrast with the relative lack of progress in resolving questions related to the robustness of quantum ground state computation, such as whether fault-tolerant universal adiabatic computing is possible, or to prove or disprove the quantum PCP conjecture [AAV13]. Such questions motivate us to seek (or to limit the possibility of) improvements to the circuit-to-Hamiltonian construction itself, which serves a foundational role in all of the results listed above. Based on ideas by Feynman [Fey86] and cast into its current form by Kitaev [KSV02],

the construction remains relatively little changed but has undergone some gradual evolution throughout its long history [Aha+09b]; [GI13]; [BT14b]; [HNN13]; [Nag14]; [BCO17], which we have extensively discussed in chapters 2 and 3.

To be robust a circuit-to-Hamiltonian construction should not only have a ground space representing valid computations, but intuitively it should penalise invalid computations with as high of an energy as possible. One way of formalising this condition is to add constraints on the input and the output of the circuit that cannot be simultaneously satisfied under the valid operation of the circuit. If the Hamiltonian enforces the correct operation of the circuit gates, then the input and output constraints that contradict each other should not be satisfiable by any state, and so the ground state energy should increase. This explains why the higher ground state energy associated with non-accepting circuits can be regarded as an energy penalty against invalid computations, which we call the “quantum UNSAT penalty”, and which is the maximal-possible promise gap achievable with a certain Hamiltonian construction.

The specific role of the $\Omega(T^{-3})$ scaling of the quantum UNSAT penalty in Kitaev’s proof is to show that the LOCAL HAMILTONIAN problem is QMA-hard with a promise gap that scales inverse polynomially in the system size. While there exists a well-defined relation between runtime T and the corresponding Hamiltonian’s system size n for any *specific* set of constructions—e.g. Kitaev’s 5-local one—the explicit scaling of this gap with T is not meaningful to the LOCAL HAMILTONIAN problem beyond the fact that it is polynomial, since the LOCAL HAMILTONIAN problem promise gap is parametrised by the number of qubits n , i.e. $1/\text{poly } n$. Nevertheless, the scaling of the quantum UNSAT penalty with T is a well defined feature of any particular circuit-to-Hamiltonian construction, and therefore we take the view that it is a reasonable metric for exploring the space of possible improvements to this construction.

5.1 Results and Overview

We analyse circuit Hamiltonians with history state ground states consisting of an arbitrary complex superposition of the time steps of the computation,

$$|\psi\rangle = \sum_{t=0}^T \psi_t |t\rangle \otimes (\mathbf{U}_t \cdots \mathbf{U}_1) |\xi\rangle, \quad (5.1)$$

where as usual $\mathbf{U}_T, \dots, \mathbf{U}_1$ are quantum gates, $|\xi\rangle$ is an arbitrary input to the computation, and $|\psi\rangle$ is a normalised state, so that in particular $\pi_t := \psi_t^* \psi_t$ is a probability distribution on $\{0, \dots, T\}$. Ground states of the form as in eq. (5.1) arise from modifications to the usual terms of the Feynman circuit Hamiltonian,

- $$\mathbf{H}_{\text{prop}} := \sum_{t=0}^T a_t |t\rangle\langle t| \otimes \mathbb{1} + \sum_{t=0}^{T-1} \left(b_t |t+1\rangle\langle t| \otimes \mathbf{U}_t + b_t^* |t\rangle\langle t+1| \otimes \mathbf{U}_t^\dagger \right), \quad (5.2)$$

where $|a_t|, |b_t| \leq 1$ for $t = 0, \dots, T$. Note that most if not all of the constructions that implement the $O(\log n)$ -local interactions with the time register using a k -LOCAL HAMILTONIAN with constant k , such as the domain wall clock which leads to a 5-local circuit Hamiltonian, can be directly applied to the modified form (5.2).

In addition to the part of the Hamiltonian which checks the propagation of the circuit, projectors $\mathbf{H}_{\text{in}} := |0\rangle\langle 0| \otimes \Pi_{\text{in}}$ and $\mathbf{H}_{\text{out}} := |T\rangle\langle T| \otimes \Pi_{\text{out}}$ can be added to \mathbf{H}_{prop} to validate specific inputs and the outputs of the computation. We define the sum of all these terms to be the ‘‘Feynman-Kitaev’’ Hamiltonian

$$\mathbf{H}_{\text{FK}} := \mathbf{H}_{\text{prop}} + \mathbf{H}_{\text{in}} + \mathbf{H}_{\text{out}}.$$

More specifically, the ground space of $\mathbf{H}_{\text{prop}} + \mathbf{H}_{\text{in}}$ will be spanned by computations starting from a valid input computation (i.e. those for which $|\xi\rangle \in \ker \Pi_{\text{in}}$ in eq. (5.1)), and \mathbf{H}_{out} will raise the energy of the state $|\psi\rangle$ in (5.1) when $\mathbf{U}_T \cdots \mathbf{U}_1 |\xi\rangle \notin \ker \Pi_{\text{out}}$. The magnitude of this frustration between the incompatible ground spaces of $\mathbf{H}_{\text{in}} + \mathbf{H}_{\text{prop}}$ and \mathbf{H}_{out} will depend on the circuit encoded by \mathbf{H}_{FK} and the specific in- and output energy penalties, i.e. the maximum acceptance probability of the circuit

$$\epsilon := \max_{\substack{|\xi\rangle \in \ker \Pi_{\text{in}} \\ |\eta\rangle \in \ker \Pi_{\text{out}}}} \langle \eta | \mathbf{U}_T \cdots \mathbf{U}_1 |\xi\rangle.$$

In the following definition we take the Π_{in} and Π_{out} that are used in the standard construction: $\Pi_{\text{out}} := |0\rangle\langle 0|$ measures a single qubit and penalises it in state $|0\rangle$ (i.e. ‘‘not accepted’’), and Π_{in} constrains a fraction of the input qubits to the $|0\rangle$ state as ancillas, some to an encoded string describing the problem instance, and leaves the rest of the input qubits unconstrained. •

For a specific set of in- and output constraints and runtime length T , we want to identify the circuit Hamiltonian best suitable to discriminate accepting and rejecting circuit paths, independent of the particular circuits used. Therefore, we let $C(\epsilon, T)$ be the set of circuits of size T for which the maximum acceptance probability is ϵ for any state obeying the in- and output constraints Π_{in} and Π_{out} , i.e.

$$C(\epsilon, T) := \{\mathbf{U}_1, \dots, \mathbf{U}_T : \max_{\substack{|\xi\rangle \in \ker \Pi_{\text{in}} \\ |\eta\rangle \in \ker \Pi_{\text{out}}}} |\langle \xi | \mathbf{U}_1 \cdots \mathbf{U}_T |\eta\rangle|^2 = \epsilon\}.$$

This leads to our definition of the quantum UNSAT penalty of a circuit-to-Hamiltonian construction, which captures how well a Hamiltonian as in eq. (5.2) can enforce the input and output penalties described above for an arbitrary circuit.

Definition 5.1. Let the $E(\mathbf{H}_{\text{FK}})$ and $E(\mathbf{H}_{\text{prop}})$ be the ground state energies of \mathbf{H}_{FK} and \mathbf{H}_{prop} respectively

and define the quantum UNSAT penalty $E_p(\epsilon, T)$

$$E_p(\epsilon, T) := \min_{\mathbf{U}_1, \dots, \mathbf{U}_T \in C(\epsilon, T)} E(\mathbf{H}_{\text{FK}}) - E(\mathbf{H}_{\text{prop}}). \quad (5.3)$$

The reason for explicitly subtracting $E(\mathbf{H}_{\text{prop}})$ is that while the standard circuit Hamiltonian constructions e.g. from Kitaev has a zero energy ground state, we do not want to restrict ourselves to this setting; the UNSAT penalty thus captures the penalisability *on top of* whatever energy the circuit Hamiltonian has as a ground state.

Note that the quantum UNSAT penalty has a closely-related quantity, the average energy of any local Hamiltonian constraints $\text{QUNSAT}_\psi = \sum_{e \in E} \langle \psi | \mathbf{h}_e | \psi \rangle / |E|$ for some set of interactions $E = \{\mathbf{h}_0, \dots, \mathbf{h}_{|E|}\}$ and a specific state ψ , as defined in the context of the detectability lemma [Aha+09a]. We use the term UNSAT *penalty* to emphasise that it is the energy difference between accepting and non-accepting computations.

Our first step in analysing the UNSAT penalty of modified Feynman Hamiltonians is to apply the same argument used in the standard construction (cf. [KSV02, sec. 14.4]) to “undo” the computation and show that \mathbf{H}_{prop} is unitarily equivalent to a Hamiltonian which acts trivially on the computational register.

- Lemma 5.2. *If $\mathbf{W} := \sum_{t=0}^T |t\rangle\langle t| \otimes (\mathbf{U}_t \cdots \mathbf{U}_1)$, then \mathbf{W} is unitary and $\mathbf{W}^\dagger \mathbf{H}_{\text{prop}} \mathbf{W} = \mathbf{H}_{\text{clock}} \otimes \mathbb{1}$, where the clock Hamiltonian $\mathbf{H}_{\text{clock}}$ is given by*

$$\mathbf{H}_{\text{clock}} := \sum_{t=0}^T a_t |t\rangle\langle t| + \sum_{t=0}^{T-1} (b_t |t+1\rangle\langle t| + b_t^* |t\rangle\langle t+1|). \quad (5.4)$$

Next we apply the same geometrical lemma used in Kitaev’s proof to lower bound the UNSAT penalty of modified Feynman Hamiltonians.

Lemma 5.3. *If the spectral gap of the corresponding clock Hamiltonian $\mathbf{H}_{\text{clock}}(T)$ —denoted $\Delta_{\mathbf{H}}(T)$ —is less than the (constant) spectral gap of $\mathbf{H}_{\text{in}} + \mathbf{H}_{\text{out}}$, then*

$$\frac{\Delta_{\mathbf{H}}(T)}{4} (1 - \sqrt{\epsilon}) \times \min\{\pi_0, \pi_T\} \leq E_p(\epsilon, T) \leq E(\mathbf{H}_{\text{clock}}(T) + |0\rangle\langle 0| + |T\rangle\langle T|). \quad (5.5)$$

The upper bound follows immediately by an operator inequality, and says that the increase in the ground state energy due to the penalty terms is at most bounded by the case when all of the frustration is in the system’s time register. The lower bound states that the UNSAT penalty can be increased either by boosting the spectral gap of the clock Hamiltonian, or by amplifying the overlap of the ground state with the beginning and ending time steps of the computation (i.e. modifying the ground state $|\psi_0\rangle$ of H such that $\langle \psi_0 | \Pi_{\text{in}} | \psi_0 \rangle$ is maximised, and analogously for Π_{out}). To see that the overlap with the endpoints of

the computation can be made arbitrarily close to one, we prove that it is in fact possible to construct a clock Hamiltonian with an arbitrary distribution as its ground state.

Lemma 5.4. *For any probability distribution μ with support everywhere on its domain $\{0, \dots, T\}$ there is a choice of coefficients $\{a_t, b_t\}_{t=0}^T$ in eq. (5.4) such that $\mathbf{H}_{\text{clock}}$ is frustration-free and has a ground space spanned by states of the form eq. (5.1) with weights $\psi_t = \sqrt{\mu_t}$.*

Using lemma 5.4 we exhibit a modified Hamiltonian for a target ground state distribution with $\pi_0, \pi_T \geq 1/4$, and show that it has a spectral gap that is $\Omega(T^{-2})$ to establish our first main result; in order to prove it, we use the *geometrical lemma* for modified Feynman Hamiltonians, a by now standard but central proof technique that is used ubiquitously in hardness constructions of the LOCAL HAMILTONIAN problem.

Theorem 5.5. *There is a frustration-free modified circuit Hamiltonian as in eq. (5.2) with a quantum UNSAT penalty $E_p(\epsilon, T)$ that is $\Omega((1 - \sqrt{\epsilon})T^{-2})$.*

A natural question is whether the lower bound given by the geometrical lemma is asymptotically tight for Feynman Hamiltonians. We show that this is not the case. As a first step we apply an improved analysis to Kitaev's original construction with uniform weights,

$$\mathbf{H}_{\text{prop}} = \sum_{t=0}^T a_t |t\rangle\langle t| \otimes \mathbb{1} + \sum_{t=0}^{T-1} \left(b_t |t+1\rangle\langle t| \otimes \mathbf{U}_t + b_t^* |t\rangle\langle t+1| \otimes \mathbf{U}_t^\dagger \right), \quad (5.6)$$

$$|\psi\rangle = \frac{1}{\sqrt{T+1}} \sum_{t=0}^T |t\rangle \otimes (\mathbf{U}_t \dots \mathbf{U}_1) |\xi\rangle \quad (5.7)$$

and we find that $E(\mathbf{H}_{FK}) = \Omega(T^{-2})$ for \mathbf{H}_{prop} as in (5.6).

Theorem 5.6. *The UNSAT penalty in Kitaev's LOCAL HAMILTONIAN construction is $\Omega(T^{-2})$.*

The proof of theorem 5.6 is based on a matrix decomposition lemma due to Jordan, which reduces the problem to lower bounding the ground energy of a discrete Schrodinger operator in each of the resulting blocks, with the lower bound following by a variational method and a suitable trial wave function.

Finally, we show that the scaling of the UNSAT penalty achieved in theorem 5.5 is the optimal scaling that can be achieved by applying the lower bound in lemma 5.3 to modified Feynman Hamiltonians of the form in eq. (5.2).

Theorem 5.7. *Let $|\psi\rangle$ be the ground state of a Hamiltonian H with eigenvalues $E := E_0 \leq E_1 \leq \dots \leq E_T$. If H is tridiagonal in the basis $\{|0\rangle, \dots, |T\rangle\}$,*

$$H := \sum_{t=0}^T a_t |t\rangle\langle t| + \sum_{t=0}^{T-1} (b_t |t+1\rangle\langle t| + b_t^* |t\rangle\langle t+1|),$$

with $|a_t|, |b_t| \leq 1$ for $t = 0, \dots, T$ then the product $\Delta_{\mathbf{H}} \cdot \min\{|\psi|_0^2, |\psi_T|^2\}$ of the spectral gap $\Delta_{\mathbf{H}} = E_1 - E$ and the minimum endpoint overlap is $O(T^{-2})$.

The rest of the chapter is organised as follows. The proofs of lemma 5.2 and lemma 5.4 can be found in section 5.2. The construction used for lemma 5.4 is given in section 5.2.3 along with some necessary background on Markov chains that will be used in the proof of theorem 5.5 in section 5.2.4. In section 5.4 we develop the quantum-to-classical mapping for arbitrary tridiagonal matrices and use it to prove theorem 5.7. In section 5.5 we describe the implications of our work for universal adiabatic computation. Finally, in section 5.6 we discuss the open problem of further increasing the quantum UNSAT penalty and relate it to some of the longstanding goals in the subject of quantum ground state computation.

5.2 Improving Circuit Hamiltonians

5.2.1 Partial Diagonalisation of Weighted History States

- We prove lemma 5.2. Since \mathbf{W} is a linear operator the calculations we need to check for lemma 5.2 are the same as in the standard unweighted case [KSV02, ch. 14.4]. As a reminder,

$$\mathbf{W}^\dagger \mathbf{W} = \sum_{t, t'=0}^T \left(|t\rangle\langle t| \otimes (\mathbf{U}_1^\dagger \cdots \mathbf{U}_t^\dagger) \right) \left(|t'\rangle\langle t'| \otimes (\mathbf{U}_1 \cdots \mathbf{U}_{t'}) \right) = \sum_{t=0}^T |t\rangle\langle t| \otimes \mathbb{1} = \mathbb{1}, \quad (5.8)$$

$$\mathbf{W}^\dagger (|t+1\rangle\langle t| \otimes \mathbf{U}_{t+1}) \mathbf{W} = |t+1\rangle\langle t| \otimes (\mathbf{U}_1^\dagger \cdots \mathbf{U}_{t+1}^\dagger) \mathbf{U}_{t+1} (\mathbf{U}_t \cdots \mathbf{U}_1) = |t+1\rangle\langle t| \otimes \mathbb{1}, \quad (5.9)$$

and so the claim of lemma 5.2 follows by linearity.

5.2.2 Kitaev's Geometrical Lemma for Weighted History States

Kitaev's geometrical lemma (see lemma 2.29) provides the starting point for the lower bound (5.5); in this section, we explicitly prove a variant (lemma 5.3) for the weighted case

We use notation from lemma 2.29. For us, $\mathbf{A} = \mathbf{H}_{\text{in}} + \mathbf{H}_{\text{out}}$, and $\mathbf{B} = \mathbf{H}_{\text{prop}}$, and in this section we use the freedom to shift the energy in eq. (3.1) to set $E(\mathbf{H}_{\text{prop}}) = 0$ (since the system can be frustrated this means the local terms may no longer be positive semi-definite, but this will not present a problem in applying the geometrical lemma above because \mathbf{H}_{prop} itself is positive semi-definite). Denote the projector onto the kernel of the penalty terms \mathbf{A} with $\Pi_{\text{pen}} := |0\rangle\langle 0| \otimes \Pi_{\text{in}}^\perp + |T\rangle\langle T| \otimes \Pi_{\text{out}}^\perp + \sum_{t=2}^{T-1} |t\rangle\langle t| \otimes \mathbb{1}$. Denote with $\mathbf{U} = \mathbf{U}_T \cdots \mathbf{U}_1$ the entire encoded quantum circuit. We first want to bound the angle θ between the kernels of the propagation and penalty Hamiltonians [KSV02]; [Cub15].

$$\begin{aligned}
\cos^2 \theta &= \max_{\substack{|\xi\rangle \in \ker A \\ |\eta\rangle \in \ker B}} |\langle \xi | \eta \rangle|^2 \\
&= \max_{\substack{|\xi\rangle \\ |\eta\rangle \in \ker B}} |\langle \eta | \Pi_{\text{pen}} | \xi \rangle|^2 \\
&\stackrel{*}{=} \max_{|\eta\rangle \in \ker B} \langle \eta | \Pi_{\text{pen}} | \eta \rangle \\
&= \max_{|\eta\rangle \in \ker B} \langle \eta | W^\dagger (W \Pi_{\text{pen}} U^\dagger) W | \eta \rangle \\
&= \max_{|\eta'\rangle \in \ker W B W^\dagger} \langle \eta' | |0\rangle\langle 0| \otimes \Pi_{\text{in}}^\perp + |T\rangle\langle T| \otimes U \Pi_{\text{out}}^\perp U^\dagger + \sum_{t=2}^{T-1} |t\rangle\langle t| \otimes \mathbf{1} | \eta' \rangle \\
&= \max_{|\phi\rangle} \sum_{s=1}^T \psi_s^* \psi_s \langle s | \langle \phi | \left(|0\rangle\langle 0| \otimes \Pi_{\text{in}}^\perp + |T\rangle\langle T| \otimes U \Pi_{\text{out}}^\perp U^\dagger + \sum_{t=2}^{T-1} |t\rangle\langle t| \otimes \mathbf{1} \right) | t \rangle | \phi \rangle \\
&= \max_{|\phi\rangle} \langle \phi | (|\psi_0^2\rangle \Pi_{\text{in}}^\perp + |\psi_T^2\rangle U \Pi_{\text{out}}^\perp U^\dagger) | \phi \rangle + 1 - |\psi_0^2| - |\psi_T^2|,
\end{aligned}$$

where we have saturated Cauchy-Schwartz in the third line (*). To bound the first inner product, we observe that if $\psi_0^2 \geq \psi_T^2$, picking $|\phi\rangle \in \ker \Pi_{\text{in}}$ gives the bound

$$\max_{|\phi\rangle} \langle \phi | (\pi_0 \Pi_{\text{in}}^\perp + \pi_T U \Pi_{\text{out}}^\perp U^\dagger) | \phi \rangle \leq \pi_0 + \pi_T \cos \vartheta,$$

where ϑ is the angle between $\text{supp } \Pi_{\text{in}}$ and $\text{supp } U \Pi_{\text{out}} U^\dagger$. This angle can be lower-bounded by the acceptance probability of the circuit:

$$\cos^2 \vartheta = \max_{\substack{|\eta\rangle \in \text{supp } \Pi_{\text{in}} \\ |\xi\rangle \in \text{supp } U \Pi_{\text{out}} U^\dagger}} |\langle \eta | \xi \rangle|^2 \leq \max_{\substack{|\eta\rangle \in \text{supp } \Pi_{\text{in}} \\ |\xi\rangle \in \text{supp } \Pi_{\text{out}}}} |\langle \eta | U | \xi \rangle|^2 \leq \epsilon.$$

Similarly, if $\psi_0^2 < \psi_T^2$, one can show an upper bound of $\pi_0 \cos \vartheta + \pi_T$. We thus obtain an overall upper bound

$$\begin{aligned}
\cos^2 \theta &\leq \max\{\pi_0, \pi_T\} + \min\{\pi_0, \pi_T\} \sqrt{\epsilon} + 1 - \pi_0 - \pi_T \\
&\leq 1 - \min\{\pi_0, \pi_T\} (1 - \sqrt{\epsilon}).
\end{aligned}$$

In Kitaev's lemma, we thus obtain a lower bound

$$2 \sin^2 \frac{\theta}{2} \geq 2 \times \frac{1 - \cos^2 \theta}{8 \cos^2 \theta} \geq \frac{1}{4} \min\{\pi_0, \pi_T\} (1 - \sqrt{\epsilon}),$$

and the claim follows.

5.2.3 Symmetrised Metropolis Hamiltonians with target ground state distributions

In this section we describe our construction which fulfils lemma 5.4. We review most concepts as needed but assume the reader has some familiarity with Markov chain transition matrices as can be found in any textbook on the subject, such as [LPW09].

Given a probability distribution π with support everywhere on its domain $\mathcal{S} = \{0, \dots, T\}$, let \mathbf{P} be the Markov chain with Metropolis transition probabilities defined on \mathcal{S} ,

$$\mathbf{P}_{t,t+1} = \frac{1}{4} \min \left\{ 1, \frac{\pi_{t+1}}{\pi_t} \right\}, \quad \mathbf{P}_{t,t-1} = \frac{1}{4} \min \left\{ 1, \frac{\pi_{t-1}}{\pi_t} \right\}, \quad \mathbf{P}_{t,t} = 1 - \mathbf{P}_{t,t+1} - \mathbf{P}_{t,t-1} \quad (5.10)$$

for all $i \in \mathcal{S}$ (setting the expressions $\mathbf{P}_{0,-1}$ and $\mathbf{P}_{T,T+1}$ to zero) and $\mathbf{P}_{t,t'} = 0$ for all $t, t' \in \mathcal{S}$ with $|t-t'| > 1$. The choice of coefficient $1/4$ implies $\mathbf{P}_{t,t} \geq 1/2$ for all t and so \mathbf{P} is positive semi-definite. The principal left eigenvector of this transition matrix $\mathbf{P} := \sum_{t,t' \in \mathcal{S}} \mathbf{P}_{t,t'} |t\rangle\langle t'|$ is $\langle \pi | = \sum_t \pi_t \langle t|$, and while \mathbf{P} is not a symmetric matrix there is a well known similarity transformation that relates \mathbf{P} to a symmetric matrix,

$$\mathbf{A} := \sum_{t,t' \in \mathcal{S}} \pi_t^{1/2} \pi_{t'}^{-1/2} \mathbf{P}_{t,t'} |t\rangle\langle t'|. \quad (5.11)$$

The two matrices are related by the fact that if $\langle v_0 |, \dots, \langle v_T |$ are the left eigenvectors of \mathbf{P} with eigenvalues $\lambda_0 = 1 \geq \lambda_1 \geq \dots \geq \lambda_T \geq 0$ then $|w_i\rangle := \sum_{t \in \mathcal{S}} \langle v_i | t \rangle \langle t | v_0 \rangle^{-1/2} |t\rangle$ satisfies $A|w_i\rangle = \lambda_i |w_i\rangle$. Therefore A has the same eigenvalues as \mathbf{P} , and in particular it has largest eigenvalue 1 corresponding to the eigenvector $|w_0\rangle$ with components satisfying $\langle t | w_0 \rangle = \langle t | v_0 \rangle^{1/2} = \sqrt{\pi_t}$. Therefore $\mathbf{H} = \mathbf{1} - \mathbf{A}$ is a nonnegative Hermitian matrix with ground state that has energy zero and components $\sqrt{\pi_t}$ in the time register basis, as claimed.

Substantial efforts been devoted to characterising spectral gaps of Markov chains. A particularly fruitful characterisation proceeds by defining a quantity called the conductance,

$$\Phi := \min_{S \subseteq \Omega} \frac{Q(S, S^c)}{\min\{\pi(S), \pi(S^c)\}}, \quad Q(S, S^c) := \sum_{x \in S, y \in S^c} \pi(x) P(x, y) \quad (5.12)$$

which determines the spectral gap within a quadratic factor,

$$\frac{\Phi^2}{2} \leq \Delta_{\mathbf{P}} \leq 2\Phi. \quad (5.13)$$

The lower bound in eq. (5.13) is known as Cheeger's inequality, and it was initially discovered in the analysis

of manifolds [Che70] before being adapted to the setting of Markov chains [SJ89]. In the next section we will use this method to lower bound the spectral gap of the Symmetrised Metropolis Hamiltonian corresponding to a particular non-uniform stationary distribution.

5.2.4 Explicit Construction of an $\Omega(T^{-2})$ UNSAT Penalty Circuit Hamiltonian

In this section let \mathbf{H}_{prop} be the Metropolis Hamiltonian corresponding to the probability distribution

$$\pi_0 = \pi_T = \frac{1}{4} \quad \text{and} \quad \pi_t = \frac{1}{2(T-1)} \quad \text{for} \quad t = 1, \dots, T-1. \quad (5.14)$$

Keeping with tradition [ANo2]; [Aha+o8]; [Aha+o9b], we exhibit \mathbf{H} as a $T+1$ by $T+1$ matrix in the time register basis,

$$H = \frac{1}{2} \begin{pmatrix} \frac{1}{T-1} & -\frac{1}{\sqrt{2T-2}} & 0 & \dots & & 0 \\ -\frac{1}{\sqrt{2T-2}} & 1 & -\frac{1}{2} & 0 & \ddots & \vdots \\ 0 & -\frac{1}{2} & 1 & -\frac{1}{2} & 0 & \vdots \\ & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & 0 & -\frac{1}{2} & 1 & -\frac{1}{2} & 0 \\ & & & 0 & -\frac{1}{2} & 1 & -\frac{1}{\sqrt{2T-2}} \\ 0 & \dots & & 0 & -\frac{1}{\sqrt{2T-2}} & \frac{1}{T-1} \end{pmatrix} \quad (5.15)$$

A few low energy eigenstates of this Hamiltonian are illustrated in figure 5.1. Since π_0 and π_T are $\Omega(1)$ it only remains to check that the spectral gap $\Delta_{\mathbf{H}}$ of the clock Hamiltonian is $\Omega(T^{-2})$, which since this spectral gap is equal to the spectral gap of the Metropolis transition matrix \mathbf{P} that is reversible with respect to π we can apply Cheeger's inequality (5.13).

The goal is to show that every subset S of \mathcal{S} has large conductance, so we divide the proof into cases corresponding to the different possibilities for the subset S . First if $S = \{0\}$ then since $\mathbf{P}_{0,1} = (8T-8)^{-1}$ so $\Phi(S)$ is $\Omega(T^{-1})$, with similar statements holding for $S = \{T\}$ and $S = \{0, T\}$. Now if $S \subseteq \{1, \dots, T-1\}$ is non-empty there must be at least one $t \in \{1, \dots, T-1\}$ such that there is a $t \in S^c$ with $\mathbf{P}_{t,t'} \geq 1/4$, and since $\pi_t = (2T-2)^{-1}$ this shows that $\Phi(S)$ is $\Omega(T^{-1})$ in this case as well. Therefore $\Delta_{\mathbf{P}}$ is $\Omega(T^{-2})$ by (5.13) and since $\Delta_{\mathbf{H}} = \Delta_{\mathbf{P}}$ this concludes the proof of theorem 5.5.

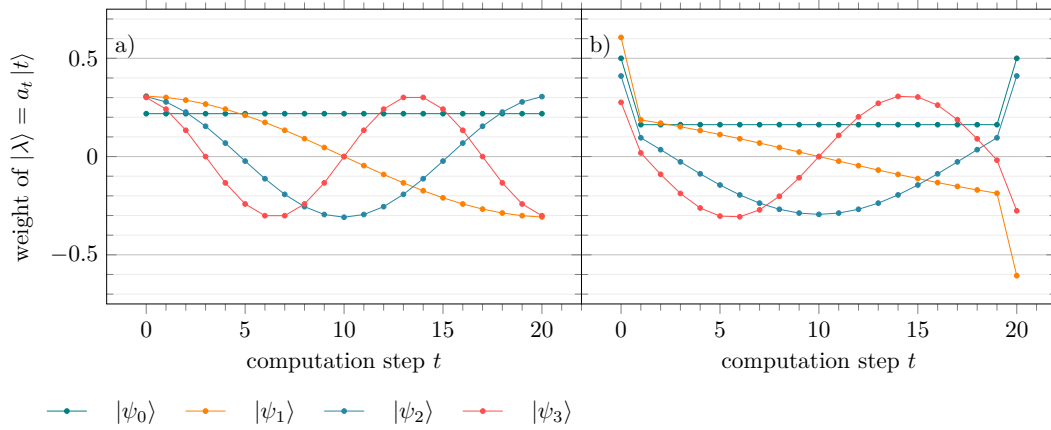


Figure 5.1: Low energy eigenstates of (a) the path graph Laplacian used in the standard circuit-to-Hamiltonian construction and (b) the symmetrised Metropolis Hamiltonian corresponding to the distribution with $\pi_0, \pi_T = 1/4$ that is used in this section.

5.3 Tightness of the Geometrical Lemma for the UNSAT Penalty

5.3.1 A Tight Bound for the Clock Hamiltonian

An interesting question is whether the scaling of $\Omega(T^{-3})$ in Kitaev’s original construction is tight or not. The Hamiltonian in his original proof is given by

$$\mathbf{H}_{\text{Kitaev}} := \mathbf{H}_{\text{prop}} + |0\rangle\langle 0| \otimes \Pi_{\text{in}} + |T\rangle\langle T| \otimes \Pi_{\text{out}},$$

where \mathbf{H}_{prop} is given by eq. (5.18) with $a_0 = a_T = 1$ and $a_t = 2$ else, and $b_t = -1$ for all t . We denote the encoded circuit with $\mathbf{U} = \mathbf{U}_T \cdots \mathbf{U}_1$. The corresponding clock- and computation registers then live on the Hilbert space $\mathcal{H} := (\mathbb{C}^{T+1}) \otimes (\mathbb{C}^2)^{\otimes d}$ for some local dimension d .

It seems to have been known for a while that the precise UNSAT penalty for Kitaev’s construction is in fact $\Omega(T^{-2})$ (see Nagaj [Nag14]), and not $\Omega(T^{-3})$ which stems from an application of the geometrical lemma alone; yet an explicit proof has never been published to the best of the authors’ knowledge. We present one in the following. We furthermore want to point out that for a modified version of Kitaev’s circuit-to-Hamiltonian construction—padding of the input and output clock states with an $\Omega(T)$ -sized identity circuit *and* spreading out the input and output penalty—yields a similar $\Omega(T^{-2})$ UNSAT penalty, but at the cost of increasing the clock register by at least a constant factor.

As a first step, and to establish some background machinery, we focus on only the *clock* part of the Hamiltonian, i.e. we disregard the encoded computation completely; the Hamiltonian we analyse has the

form $\mathbf{H} = |0\rangle\langle 0| + \mathbf{H}_{\text{clock}} + |T\rangle\langle T|$. Since \mathbf{H} is stoquastic, we can lower bound the ground state energy of \mathbf{H} by combining a suitable ansatz for the ground state with the following lemma [Far+11].

Lemma 5.8 (Farhi et al., 2011). *Let \mathbf{H} be a Hermitian operator which is stoquastic in the $|t\rangle$ basis (meaning $\langle t|\bar{H}|t'\rangle \leq 0$ for all $t \neq t'$). Let E be its lowest eigenvalue. Then*

$$E \geq \min_t \frac{\langle t|H|\phi\rangle}{\langle t|\phi\rangle} \quad (5.16)$$

for any state $|\phi\rangle$ such that $\langle t|\phi\rangle > 0$ for all t .

Noting that the state $|\phi\rangle$ in Lemma 5.8 does not need to be normalised, define

$$|\phi\rangle := \sum_{t=0}^T \sin\left(\frac{\pi(t+1)}{T+2}\right) |t\rangle. \quad (5.17)$$

The bound (5.16) can be evaluated using three cases. The first case is $t = 0$, which yields

$$\begin{aligned} \frac{\langle 0|H|\phi\rangle}{\langle 0|\phi\rangle} &= \left[\sin\left(\frac{\pi}{T+2}\right) \right]^{-1} \left[2 \sin\left(\frac{\pi}{T+2}\right) - \sin\left(\frac{2\pi}{T+2}\right) \right] \\ &= 2 \left(1 - \cos\left(\frac{\pi}{T+2}\right) \right) \\ &= \frac{\pi^2}{T^2} - O(T^{-3}). \end{aligned}$$

The next case is $1 \leq t \leq T-1$,

$$\begin{aligned} \frac{\langle t|H|\phi\rangle}{\langle t|\phi\rangle} &= \left[\sin\left(\frac{\pi(t+1)}{T+2}\right) \right]^{-1} \left[2 \sin\left(\frac{\pi(t+1)}{T+2}\right) - \sin\left(\frac{\pi t}{T+2}\right) - \sin\left(\frac{\pi(t+2)}{T+2}\right) \right] \\ &= 2 \left(1 - \cos\left(\frac{\pi}{T+2}\right) \right) \\ &= \frac{\pi^2}{T^2} - O(T^{-3}). \end{aligned}$$

The final case is $t = T$,

$$\begin{aligned} \frac{\langle T|H|\phi\rangle}{\langle T|\phi\rangle} &= \left[\sin\left(\frac{\pi(T+1)}{T+2}\right) \right]^{-1} \left[2 \sin\left(\frac{\pi(T+1)}{T+2}\right) - \sin\left(\frac{\pi T}{T+2}\right) \right] \\ &= 2 - \sin\left(\frac{\pi T}{T+2}\right) \csc\left(\frac{\pi(T+1)}{T+2}\right) \\ &= \frac{\pi^2}{T^2} - O(T^{-3}) \end{aligned}$$

Therefore we have shown that $\langle t|H|\phi\rangle/\langle t|\phi\rangle$ is $\Omega(T^{-2})$ for all t , and so applying Lemma 5.8 we can

conclude that $E(H)$ is $\Omega(T^{-2})$.

5.3.2 A Tight Bound for the Full Circuit Hamiltonian

For the purpose of this discussion, we will assume that $\text{rank } \Pi_{\text{in}}, \text{rank } \Pi_{\text{out}} \geq d/2$; this is a natural assumption e.g. if $\Pi_{\text{in}} = |1\rangle\langle 1|^{(1)} \otimes \mathbb{1}^{(2)} \otimes \dots \otimes \mathbb{1}^{(d)}$ just penalises the first qubit in a state $|1\rangle$. Note that more penalised qubits generally *increase* the rank of Π_{in} if we demand the penalties to be local operators (i.e. not something like $|1\rangle\langle 1| \otimes |0\rangle\langle 0| \otimes \dots$, but $|1\rangle\langle 1| \otimes \mathbb{1} + \mathbb{1} \otimes |0\rangle\langle 0| \otimes \mathbb{1} + \dots$).

We know that there exists a global unitary \mathbf{W} on \mathcal{H} such that $\text{spec}(\mathbf{H}_{\text{prop}}) = \text{spec}(\mathbf{W}^\dagger \mathbf{H}_{\text{prop}} \mathbf{W}) = \text{spec}(\Delta)$ up to multiplicities, where Δ is the Laplacian of a path graph of $T + 1$ vertices:

$$\Delta = \begin{pmatrix} 1 & -1 & 0 & \dots & & 0 \\ -1 & 2 & -1 & & & \\ 0 & -1 & \ddots & \ddots & & \vdots \\ \vdots & & \ddots & \ddots & -1 & 0 \\ & & & -1 & 2 & -1 \\ 0 & & \dots & 0 & -1 & 1 \end{pmatrix}$$

Then

$$\mathbf{W}^\dagger \mathbf{H}_{\text{Kitaev}} \mathbf{W} = \underbrace{\Delta \otimes \mathbb{1} + |0\rangle\langle 0| \otimes \Pi_{\text{in}}}_{:=\mathbf{A}} + \underbrace{|T\rangle\langle T| \otimes \mathbf{U}^\dagger \Pi_{\text{out}} \mathbf{U}}_{:=\mathbf{B}}. \quad (5.18)$$

We choose to work in a basis where \mathbf{A} takes the form

$$\mathbf{A} = \text{diag}(\underbrace{\Delta', \dots, \Delta'}_{\text{rank } \Pi_{\text{in}} \text{ times}}, \Delta, \dots, \Delta), \quad (5.19)$$

where $\Delta' = \Delta + |0\rangle\langle 0|$. Note that this is always possible: we simply re-order the computational register such that Π_{in} penalises the first $\text{rank } \Pi_{\text{in}}$ states. Note that $\mathbf{B} = |T\rangle\langle T| \otimes \mathbf{U}^\dagger \Pi_{\text{out}} \mathbf{U}$ does not have a simple form in this basis (apart from having only a single entry *within* each block in the time register basis, i.e. at the diagonal entry $|T\rangle\langle T|$).

If we naïvely try to diagonalise \mathbf{B} , we will again mix up the nice block-diagonal form in eq. (5.19); our goal is thus to find a unitary $\mathbf{V} = \mathbb{1}_d \otimes (\mathbf{V}' \oplus \mathbf{V}'')$ which respects the block-diagonal structure of \mathbf{A} , i.e. in particular leaves the upper left block invariant (which is automatically the case if $\dim \mathbf{V}' \leq \text{rank } \Pi_{\text{in}}$).

For this we need a variant of Jordan's Lemma. While Jordan's original paper addresses the case of orthogonal transformations between subspaces, we can state it more suitable to our needs:

Lemma 5.9 (Jordan [Jor75], Th. 1). *Let Π_0 and Π_1 be two Hermitian projectors in some Hilbert space \mathcal{H} .*

Then there exists a decomposition of \mathcal{H} into one- and two-dimensional subspaces $\mathcal{H} = \bigoplus_i \mathcal{H}_i$, such that \mathcal{H}_i is invariant under both Π_0 and Π_1 , and such that $\text{rank } \Pi_j|_{\mathcal{H}_i} \leq 1$ for all j and i .

The proof is standard. Lemma 5.9 allows us to state the following corollary.

Corollary 5.10. *Let Π_0 and Π_1 be projectors with $\text{rank } \Pi_0 = \text{rank } \Pi_1 = d/2$, and $\text{rank}(\Pi_0 + \Pi_1) = d$ is full rank. Then there exists a unitary \mathbf{V} such that $\mathbf{V}^\dagger \Pi_0 \mathbf{V} = \mathbb{1}_{d/2} \oplus 0$. Furthermore,*

$$\mathbf{V}^\dagger \Pi_1 \mathbf{V} = \begin{pmatrix} \mathbf{M}_{aa} & \mathbf{M}_{ab} \\ \mathbf{M}_{ba} & \mathbf{M}_{bb} \end{pmatrix},$$

where each $d/2 \times d/2$ block \mathbf{M}_{ij} is diagonal, and $\mathbf{M}_{ab} = \mathbf{M}_{ba} < 0$.

Proof. Applying lemma 5.9 to Π_0 and Π_1 , we know that there exists a basis in which Π_0 is diagonal, and $\Pi_1 = \bigoplus_i \mathbf{M}_i$, where the \mathbf{M}_i are 2×2 or 1×1 Hermitian matrices. Re-order the basis again such that $\Pi_0 = \mathbb{1}_r \oplus 0$ with $r = d/2$; we denote the unitary transformation from Jordan's lemma with this following reordering as \mathbf{V} .

Under the same re-ordering, the matrix $\mathbf{V}^\dagger \Pi_1 \mathbf{V}$ then necessarily has the form

$$\mathbf{V}^\dagger \Pi_1 \mathbf{V} = \begin{pmatrix} \mathbf{M}_{aa} & \mathbf{M}_{ab} \\ \mathbf{M}_{ba} & \mathbf{M}_{bb} \end{pmatrix} = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 & \xi_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \ddots & \vdots & 0 & \xi_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 & \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_r & 0 & \cdots & 0 & \xi_r \\ \hline \xi_1^* & 0 & \cdots & 0 & \mu_1 & 0 & \cdots & 0 \\ 0 & \xi_2^* & \ddots & \vdots & 0 & \mu_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 & \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \xi_r^* & 0 & \cdots & 0 & \mu_r \end{pmatrix} \quad (5.20)$$

A further global phase transformation (adjoining with a diagonal unitary, which leaves Π_0 invariant) allows us to assume that the $\xi_i = -|\xi_i|$, and the claim follows. \square

We still need to show that the resulting Hamiltonian is in particular a graph Laplacian connecting the T^{th} entry in every block of Δ in eq. (5.19) with the T^{th} entry in at least one block of Δ' , which is where the input penalty terms are located. Naturally, this will depend on the encoded circuit, and—using notation from the last proof—on the respective ranks of \mathbf{M}_{aa} and \mathbf{M}_{bb} . For this, we state the following lemma.

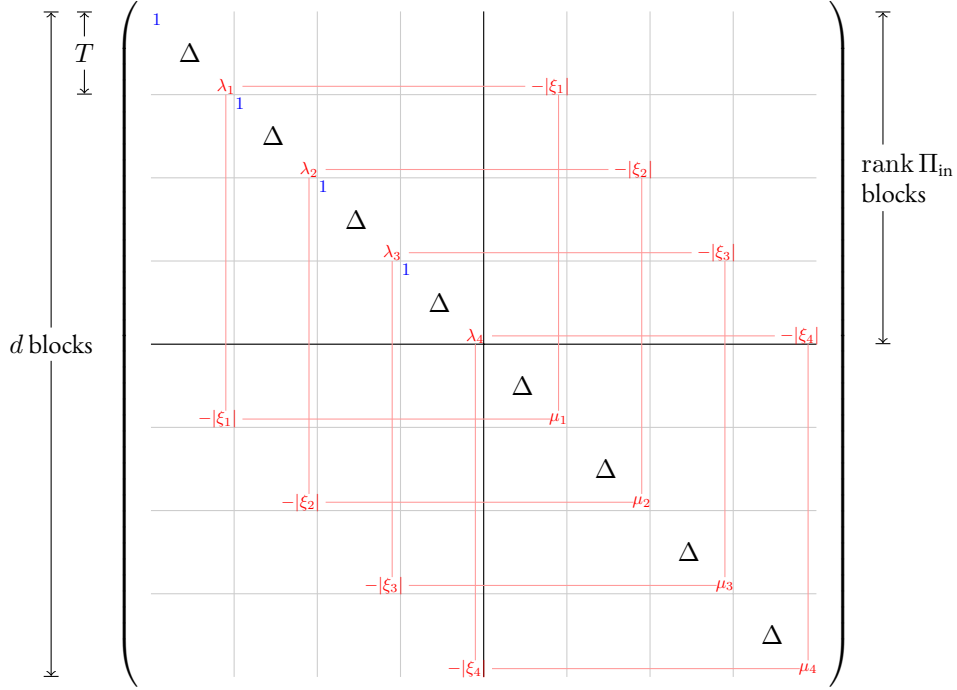


Figure 5.2: Sketch of matrix $(\mathbb{1} \otimes \mathbf{V})^\dagger(\Delta \otimes \mathbb{1} + |1\rangle\langle 1| \otimes \Pi_{\text{in}} + |T\rangle\langle T| \otimes M)(\mathbb{1} \otimes \mathbf{V})$.

Lemma 5.II. *Let \mathbf{U} encode a NO instance. Then $\Pi_{\text{in}} + \mathbf{U}^\dagger \Pi_{\text{out}} \mathbf{U}$ has full rank.*

Proof. It is easy to see that if $\ker(\Pi_{\text{in}} + \mathbf{U}^\dagger \Pi_{\text{out}} \mathbf{U}) \neq 0$, there would be a state that can be accepted with perfect probability—contradicting the assumption that \mathbf{U} encodes a NO instance. \square

This technical machinery allows us to prove theorem 5.6. Remember that we have to show that whenever \mathbf{U} encodes a NO instance, then $\lambda_{\min}(\mathbf{H}_{\text{Kitaev}}) = \Omega(T^{-2})$.

Proof. Since the circuit is a NO-instance, there exists a minimal acceptance probability for any (valid) input $|x\rangle \in \ker \Pi_{\text{in}}$, which we denote with ϵ . By lemma 5.II, we know that we can apply corollary 5.10 to Π_{in} and $\mathbf{U}^\dagger \Pi_{\text{out}} \mathbf{U}$; the unitary transformation bringing the latter into a form eq. (5.20) while leaving $\Pi_{\text{in}} = \mathbb{1} \oplus 0$ we denote with \mathbf{V} .

Now perform the similarity transform on eq. (5.18), i.e.

$$(\mathbb{1} \otimes \mathbf{V}^\dagger) \mathbf{W}^\dagger \mathbf{H}_{\text{Kitaev}} \mathbf{W} (\mathbb{1} \otimes \mathbf{V}) = \Delta \otimes \mathbb{1} + |0\rangle\langle 0| \otimes (\mathbb{1} \oplus 0) + |T\rangle\langle T| \otimes (\mathbf{V}^\dagger \mathbf{U}^\dagger \Pi_{\text{out}} \mathbf{U} \mathbf{V}).$$

This matrix is depicted in fig. 5.2. Since \mathbf{U} encodes a quantum circuit, we can immediately calculate the magnitudes of the λ_i , μ_i and ξ_i 's (see eq. (5.20) and fig. 5.2 as a reference).

1. If $|x\rangle \in \mathbb{C}^d$, then $\langle x | \mathbf{U}^\dagger \Pi_{\text{out}} \mathbf{U} |x\rangle = \|\Pi_{\text{out}} \mathbf{U} |x\rangle\|^2$ is the probability that the circuit \mathbf{U} rejects input $|x\rangle$.
2. If $|x\rangle \in \ker \Pi_{\text{in}}$ —i.e. for the \mathbf{M}_{bb} block— $\langle x | \mathbf{U}^\dagger \Pi_{\text{out}} \mathbf{U} |x\rangle \geq 1 - \epsilon$. Otherwise we have no non-trivial lower bound.
3. Assuming $\dim \text{supp } \Pi_{\text{out}} = d/2$ (discard the rest), we can thus immediately conclude that each matrix block

$$\mathbf{P}_i = \begin{pmatrix} \lambda_i & -|\xi_i| \\ -|\xi_i| & \mu_i \end{pmatrix}$$

is of rank $\mathbf{P}_i = 1$ with $1 \geq \mu_i \geq 1 - \epsilon$.

4. Since the eigenvalues of a matrix are unique, we could proceed diagonalising $\mathbf{V}^\dagger \Pi_{\text{out}} \mathbf{V}$ by further diagonalising each block \mathbf{P}_i separately, and we would obtain the same overall matrix as if we had diagonalised Π_{out} in one step. Since Π_{out} is a psd projector, we thus know that each of the \mathbf{P}_i has to be a psd projector, and we can conclude

$$\mathbf{P}_i = \begin{pmatrix} \eta_i^2 \mu_i & -\eta_i \mu_i \\ -\eta_i \mu_i & \mu_i \end{pmatrix}$$

for some $\eta_i := \lambda_i / \mu_i \geq 0$. Then $1 = \text{tr } \mathbf{P}_i = \mu_i(1 + \eta_i^2) \geq (1 - \epsilon)(1 + \eta_i^2)$, and thus

$$\eta_i \leq \sqrt{\frac{1}{1 - \epsilon} - 1} = \sqrt{\frac{\epsilon}{1 - \epsilon}} \leq \sqrt{\frac{\epsilon}{2}},$$

and thus $\lambda_i = \eta_i^2 \mu_i \leq \epsilon \mu_i / 2 \leq \epsilon / 2$, and $|\xi_i| = \eta_i \mu_i \leq \sqrt{\epsilon / 2}$.

5. From $\mu_i \leq 1$, one can obtain a trivial bound $1 \leq 1 + \eta_i^2$, and thus obviously $\lambda_i \geq 0$.

What thus remains to be analysed is the spectrum of each pairs of blocks in fig. 5.2, i.e. blocks of the form $\Delta + \Delta'$ matrix \mathbf{P}_i spread to the submatrix indexed by $T, 2T$; more precisely, each block will have twice the size of Δ , and we can write it as a stoquastic matrix on $\mathbb{C}^2 \otimes \mathbb{C}^{T+1}$:

$$\mathbf{S} = (|0\rangle\langle 0| + |1\rangle\langle 1|) \otimes \Delta + |0\rangle\langle 0| \otimes |0\rangle\langle 0| + \underbrace{\mu [|1\rangle\langle 1| + \eta^2 |0\rangle\langle 0| - \eta(|0\rangle\langle 1| + |1\rangle\langle 0|)]}_{=: \mathbf{L}} \otimes |T\rangle\langle T|, \quad (5.21)$$

where $\mu \geq 1 - \epsilon$ and $\eta \leq \sqrt{\epsilon / 2}$. For e.g. $\dim S = 8$ and reversing the second half of the time register

(i.e. reversing the basis from T to $2T$), this matrix looks like

$$\mathbf{S}_8 = \begin{pmatrix} 2 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & 2 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 2 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & \mu\eta^2 + 1 & -\mu\eta & 0 & 0 & 0 \\ 0 & 0 & 0 & -\mu\eta & 1 + \mu & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 2 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & 2 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & 1 \end{pmatrix}.$$

Using a variational argument with the state

$$|\phi\rangle := \sum_{i=1}^T \sin\left(\frac{\pi i}{2T}\right) (|i\rangle + |i+T\rangle)$$

and lemma 5.8, we can then explicitly show that $\mathbf{S} = \Omega(T^{-2})$, independent of μ and η . The bound can be evaluated using five cases, which we summarise as follows.

$t = 1$:

$$\left(2 \sin\left(\frac{\pi}{2T}\right) - \sin\left(\frac{\pi}{T}\right)\right) \csc\left(\frac{\pi}{2T}\right) = \frac{\pi^2}{4T^2} + O(T^{-3})$$

$1 < t \leq T - 1$ and $T + 1 < t < T$:

$$4 \sin^2\left(\frac{\pi}{4T}\right) = \frac{\pi^2}{4T^2} + O(T^{-3})$$

$t = T$:

$$\eta^2 \mu - \eta \mu \sin\left(\frac{\pi}{2T}\right) - \cos\left(\frac{\pi}{2T}\right) + 1 = \eta^2 \mu - \frac{\pi \eta \mu}{2T} + \frac{\pi^2}{8T^2} + O(T^{-3})$$

$t = T + 1$:

$$\mu - \eta \mu \csc\left(\frac{\pi}{2T}\right) - 2 \cos\left(\frac{\pi}{2T}\right) + 1 = -\frac{2T(\eta\mu)}{\pi} + (\mu - 1) - \frac{\pi \eta \mu}{12T} + \frac{\pi^2}{4T^2} + O(T^{-3})$$

$t = 2T$:

$$1 - \sin\left(\frac{\pi(T-1)}{2T}\right) = \frac{\pi^2}{8T^2} + O(T^{-3}) \quad \square$$

For convenience, we further summarise the findings regarding the block matrix decomposition of the full circuit Hamiltonian given in the proof of theorem 5.6 in the following lemma.

Lemma 5.12. *Kitaev's construction can be block decomposed as $\mathbf{H}_{\text{Kitaev}} = \bigoplus \mathbf{S}_i$, where each block \mathbf{S}_i has the form eq. (5.21). The ground state energy of each of these blocks is $\Omega(T^{-2})$, and thus $E(\mathbf{H}_{\text{Kitaev}}) = \Omega(T^{-2})$.*

5.4 Limitations on further improvement

The proof of Theorem 5.7 is based on applying a sharp spectral gap bound for birth-and-death Markov chains to a quantum-to-classical mapping that has been studied previously in the closely related context of universal adiabatic computation [Aha+08] and the complexity of stoquastic Hamiltonians [BBTo6]; [BT09]. A new feature of our application is the realisation that this quantum-to-classical mapping defines a Markov chain even for tridiagonal Hamiltonian matrices with arbitrary complex entries, while previous applications have been restricted to cases for which \mathbf{H} has all non-positive off-diagonal matrix entries in the time register basis.

In this section we continue with the notation of (3.1), but now we use the freedom to shift the energy to set $a_t \geq 0$ for all t , and so the ground state energy E will in general satisfy $0 \leq E < 1$. Define $\mathbf{G} := (\mathbf{1} - H)/(1 - E)$ to be a shifted and rescaled version of \mathbf{H} which is designed to satisfy $\mathbf{G}|\psi\rangle = |\psi\rangle$, where $|\psi\rangle$ labels the ground state of \mathbf{H} . For all $t, t' \in \{0, \dots, T\}$, define

$$\mathbf{P}_{t,t'} := \begin{cases} \psi_{t'} \mathbf{G}_{t,t'} \psi_t^{-1} & \text{if } \psi_t \neq 0 \text{ and } \psi_{t'} \neq 0, \\ 0 & \text{otherwise.} \end{cases} \quad (5.22)$$

In the following lemma we will show that the $\mathbf{P}_{t,t'}$ are transition probabilities for an irreducible Markov chain on $\{0, \dots, T\}$, i.e. in particular that they are all nonnegative. First, observe that if \mathbf{H} is stoquastic as in previous applications, then \mathbf{G} is a nonnegative matrix in the time register basis, ψ has nonnegative amplitudes in this basis by the Perron-Frobenius theorem and so $\mathbf{P}_{t,t'}$ is explicitly nonnegative. Here we show that even when \mathbf{G} contains arbitrary complex matrix entries (while being tridiagonal) we still have $\mathbf{P}_{t,t'} \geq 0$, because of cancellations that occur between the matrix elements of \mathbf{G} and the amplitudes of the ground state wave function in the time register basis. Continuing with the same notation used in (5.4),

Lemma 5.13. *If $\psi_0 \neq 0$, $\psi_T \neq 0$, and $b_t \neq 0$ for $t = 0, \dots, T$, then $\psi_t \neq 0$ for $t = 1, \dots, T-1$ and $\mathbf{P}_{t,t+1} = \psi_{t+1} \mathbf{G}_{t,t+1} \psi_t^{-1} \geq 0$ for all $t \in \{0, \dots, T-1\}$.*

Before proving the lemma, note that the conditions may be taken to hold without loss of generality, since $\psi_0 = 0$ or $\psi_T = 0$ immediately implies Theorem 5.7, or similarly if $b_{t'} = 0$ for some t' then $|\psi^\perp\rangle := \sum_{t=0}^T \psi_t^\perp |t\rangle$ defined by

$$\psi_t^\perp := \begin{cases} \frac{\psi_t}{\psi^2([0, t'])} & 0 \leq t \leq t' \\ -\frac{\psi_t}{\psi^2([t'+1, T])} & t' < t \leq T-1 \end{cases} \quad (5.23)$$

satisfies $\langle \psi^\perp | \psi \rangle = 0$ and $\mathbf{H}|\psi^\perp\rangle = E|\psi^\perp\rangle$, which implies $\Delta_{\mathbf{H}} = 0$ and so again Theorem 5.7 holds in this case¹.

Now turning to the proof of Lemma 5.13. From $\mathbf{H}|\psi\rangle = E|\psi\rangle$ we have that

$$a_0\psi_0 + b_0\psi_1 = E\psi_0 \quad (5.24)$$

$$b_{i-1}^*\psi_{i-1} + a_i\psi_i + b_i\psi_{i+1} = E\psi_i \quad , \quad \text{for } i = 1, \dots, T-1 \quad (5.25)$$

$$a_T\psi_T + b_{T-1}^*\psi_{T-1} = E\psi_T \quad (5.26)$$

Since $\mathbf{P}_{t,t'} = 0$ when $|t-t'| > 1$, our goal is to show $\mathbf{P}_{t,t+1} > 0$ for $t \in \{0, \dots, T-1\}$, and $\mathbf{P}_{t,t-1} > 0$ for $t \in \{1, \dots, T\}$. The first claim $\mathbf{P}_{t,t+1} > 0$ will follow by showing that E is minimised when $\psi_t \neq 0$ and $\psi_{t+1}b_t\psi_t^{-1} < 0$ for all $t = 0, \dots, T-1$. The second claim $\mathbf{P}_{t-1,t} > 0$ is then implied immediately since

$$\psi_t b_t^* \psi_{t+1}^{-1} = \left(\frac{|\psi_t|}{|\psi_{t+1}|} \right)^2 \frac{\psi_{t+1}^*}{\psi_t^*} b_t^* = \left(\frac{|\psi_t|}{|\psi_{t+1}|} \right)^2 (\psi_{t+1} b_t \psi_t^{-1})^* . \quad (5.27)$$

Rearranging eq. (5.24) yields $\psi_1 b_0 \psi_0^{-1} = E - a_0$, and since $E - a_0$ is real the value of E implied by this equation alone is minimised when the LHS is negative. This observation will be taken as the base case for an argument by mathematical induction on the finite set $\{1, \dots, T-1\}$. The inductive hypothesis is that the value of E implied by considering only equations 0 through i in the list eq. (5.25) is minimised when $\psi_i b_{i-1} \psi_{i-1}^{-1}$ is negative for $1, \dots, i$, and this will be used to show that the minimum value of E that satisfies equations 0 through $t+1$ in eq. (5.25) will be achieved when $\psi_{t+1} b_t \psi_t^{-1}$ is negative as well. Using the fact that $\psi_t \neq 0$ from the inductive hypothesis we may express eq. (5.25) as

$$b_{t-1}^* \frac{\psi_{t-1}}{\psi_t} + b_t \frac{\psi_{t+1}}{\psi_i} = E - a_t \quad , \quad \text{for } t = 1, \dots, T-1. \quad (5.28)$$

Since $E - a_t$ is real and $\psi_{t-1} b_{t-1}^* \psi_t^{-1} = (|\psi_{t-1}|/|\psi_t|)^2 (\psi_t b_{t-1} \psi_{t-1}^{-1})^*$ is negative by the inductive hypothesis, the value of E implied by equations 0 through $t+1$ in the list eq. (5.25) will indeed be minimised by taking $\psi_{t+1} b_t \psi_t^{-1}$ to be negative. This establishes the inductive claim and completes the proof of Lemma

¹Note that the same idea behind eq. 5.23 can be used to upper bound the spectral gap by the minimum of $\pi_t \pi_{t+1}$ over all $t \in \{0, \dots, T-1\}$ such that $\psi^2([0, t'])$ and $\psi^2([t'+1, T])$ are both $\Omega(1)$.

5.13.

Having established that $\mathbf{P}_{t,t'} \geq 0$ we now list several standard facts which have been previously applied to \mathbf{P} when \mathbf{G} is nonnegative, which can also be seen in the present case by direct computation:

1. \mathbf{P} is a stochastic matrix, i.e. $\sum_{t'=0}^T \mathbf{P}_{t,t'} = 1$ for all $t \in \{0, \dots, T\}$, and therefore it can be regarded as the transition matrix of a discrete time Markov chain.
2. The largest eigenvalue of \mathbf{P} is equal to 1 and it corresponds to the unique principal eigenvector $|\pi\rangle = \sum_{t=0}^T |\psi_t|^2 |t\rangle$. The probability distribution $\pi_t := \langle t|\pi\rangle$ is the stationary distribution of the corresponding Markov chain.
3. The Markov chain defined by \mathbf{P} is reversible with respect to its stationary distribution,

$$\pi_t \mathbf{P}_{t,t'} = \langle t'|\psi\rangle \langle \psi|t\rangle \mathbf{G}_{t,t'} = (\langle t|\psi\rangle \langle \psi|t'\rangle G_{t,t'}^*)^* = \pi_{t'} \mathbf{P}_{t',t}.$$

4. If $|\psi_0\rangle, |\psi_1\rangle, \dots, |\psi_T\rangle$ are the eigenvectors of \mathbf{H} with corresponding eigenvalues $E_0 < E_1 \leq \dots \leq E_T$, then $|\phi_k\rangle = \sum_{x \in \Omega} \langle \psi_0|x\rangle \langle x|\psi_k\rangle |x\rangle$ is an eigenvector of \mathbf{P} with eigenvalue $(1 - E_k)/(1 - E_0)$. Since this is the complete list of eigenvectors of \mathbf{P} we have shown that the spectral gaps of \mathbf{H} and \mathbf{P} satisfy

$$\Delta_{\mathbf{P}} = (1 - E) \Delta_{\mathbf{H}}. \quad (5.29)$$

The relation eq. (5.29) means that we can apply techniques developed for upper bounding the spectral gap of Markov chains to the problem of upper bounding the spectral gap of \mathbf{H} . A non-trivial example of such an upper bound is eq. (5.13): if the overlap of the stationary distribution with the end points $|0\rangle$ and $|T\rangle$ is constant then we can immediately see that the conductance Φ is $O(T^{-1})$ by the fact that the stationary distribution is normalised, and this implies that $\Delta_{\mathbf{H}}$ is $O(T^{-1})$. It turns out we can obtain an even tighter bound by using a characterisation of spectral gaps that applies specifically to birth-and-death chains [CS13], which we state here as a lemma.

Lemma 5.14. *If \mathbf{P} is a birth and death chain with stationary distribution π , then the spectral gap $\Delta_{\mathbf{P}}$ satisfies*

$$\frac{1}{2\ell} \leq \Delta_{\mathbf{P}} \leq \frac{4}{\ell} \quad (5.30)$$

where

$$\ell := \max \left\{ \max_{j:j \leq i'} \sum_{k=j}^{i'-1} \frac{\pi([0, j])}{\pi(k) \mathbf{P}_{k,k+1}}, \max_{j:j > i'} \sum_{k=i'+1}^j \frac{\pi([j, T])}{\pi(k) \mathbf{P}_{k,k-1}} \right\} \quad (5.31)$$

where i' satisfies $\pi([0, i']) \geq 1/2$ and $\pi([i', n]) \geq 1/2$.

In the present case we are seeking a lower bound on ℓ in order to have an upper bound on the gap. To simplify the formulas we assume that the stationary distribution of the weighted history state is symmetric around $t = T/2$ (otherwise the problem divides into two similar cases). Since we are seeking a lower bound on ℓ we can ignore the factor of $\mathbf{P}_{k,k+1} \leq 1$ in the denominator, and we are also free to replace the maximisation over j with any fixed choice of j .

With these simplifications and the choice of $j = 1$ eq. (5.31) becomes

$$\ell \geq \psi_0^2 \sum_{t=1}^{T/2-1} \frac{1}{\psi_t^2}.$$

Applying the inequality of the arithmetic and geometric means yields,

$$\sum_{t=1}^{\frac{T}{2}-1} \frac{1}{\psi_t^2} \geq \left(\frac{T}{2} - 1\right) \left(\psi_1^2 \cdots \psi_{\frac{T}{2}-1}^2\right)^{-1/k} \quad (5.32)$$

$$\geq \left(\frac{T}{2} - 1\right)^2 \left(\sum_{t=1}^{T/2-1} \psi_t\right)^{-1} \quad (5.33)$$

and so ℓ is $\Omega(\psi_0^2 T^2)$, and from eq. (5.30) we have that the spectral gap $\Delta_{\mathbf{P}}$ is $O(\ell^{-1})$ and so $\Delta_{\mathbf{H}} \cdot \psi_0^2$ is $O(T^{-2})$ as claimed.

5.5 Universal adiabatic computation

The circuit-to-Hamiltonian construction used in the proof of QMA-completeness also plays a crucial role in the proof that adiabatic quantum computation (AQC) can simulate the quantum circuit model with polynomial overhead. In this section we will review the relevant aspects of the standard universal AQC construction in order to explain how our results on modified Feynman Hamiltonians allow for a useful practical improvement in universal AQC, and also how theorem 5.7 yields a lower bound on the time needed for universal AQC using modified Feynman Hamiltonians to simulate the circuit model.

The standard version of universal AQC is based on \mathbf{H}_{prop} as given in (5.6), together with

$$\mathbf{H}_{\text{init}} := |0\rangle\langle 0| \otimes \mathbf{1} = \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & & \\ 0 & 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 & 0 \\ 0 & \cdots & & 0 & 1 \end{pmatrix} \otimes \mathbf{1} \quad (5.34)$$

The goal in universal AQC is to prepare the ground state of \mathbf{H}_{prop} by continuously varying a Hamiltonian $\mathbf{H}(s)$ as that depends on an adiabatic parameter $0 \leq s \leq 1$. First when $s = 0$ the system is initialised in the ground state of $\mathbf{H}(0) := \mathbf{H}_{\text{init}}$, which is a fiducial state that can easily be prepared, then the parameter s is slowly increased until $s = 1$ when the system Hamiltonian is $\mathbf{H}(1) := \mathbf{H}_{\text{prop}}$. Defining $\|\dot{\mathbf{H}}\| = \max_s \|\mathrm{d}\mathbf{H}/\mathrm{d}s\|$ and $\Delta = \min_s \Delta_{\mathbf{H}}(s)$, a sufficient condition for preparing the ground state of \mathbf{H}_{prop} is for the total evolution time satisfies $t_{\text{adiabatic}} = \Theta(\|\dot{\mathbf{H}}\|\Delta^{-2})$. This time scale can be shown to be bounded by $\text{poly}(n)$ for the linear interpolation schedule,

$$\mathbf{H}(s) = (1 - s)\mathbf{H}_{\text{init}} + s\mathbf{H}_{\text{prop}} \quad (5.35)$$

where specifically $\Delta = \Omega(T^{-2})$ and $\|\dot{\mathbf{H}}\| = O(1)$ so that $t_{\text{adiabatic}} = \Theta(T^4)$. The lower bound on the spectral gap was proven using a quantum-to-classical mapping, a monotonicity property of the ground state wave function as a function of s , and Cheeger's inequality.

One may notice that measuring the time register of the uniform history state (5.7) results in collapsing the computational register to its final time step $t = T$ with probability $O(T^{-1})$. To avoid repeating the adiabatic evolution many times, it is desirable to boost this probability to $\Omega(1)$. The standard trick for doing this is to pad the end of the computation with identity gates, meaning after performing the desired computational gates $\mathbf{U}_1, \dots, \mathbf{U}_T$ one adds r additional identity gates $\mathbf{U}_r = \mathbf{U}_{r+1} = \dots = \mathbf{U}_{r+T} = I$, so that measuring $t \in [T, r + T]$ suffices to project the computational register to be in the end of the computation.

First we point out that modified Feynman Hamiltonians, together with the symmetrised Metropolis Hamiltonian construction of section 5.2.3, open up a new set of trade-offs in universal adiabatic computation that may be relevant for practical implementations. Specifically, the Hamiltonian $\mathbf{H}_{\text{prop}}^*$ used in section 5.2.4 with $\Omega(1)$ probability on the endpoints can be used to increase the probability that measuring the time register will collapse the computational register of the system into the final time step of the computation. This provides an alternative to “padding the end of the computation with identity gates” that is normally used to raise the probability of sampling from the final time step of the computation. Padding the length of the computation with identity gates is relatively expensive in practical terms when the time register is encoded using local interactions (such as the domain wall clock) because the clock must be represented in unary, meaning the number of clock qubits scales linearly with the total length of the (padded) computation.

Achieving an overlap of $\delta \approx 1$ with the final step of the computation by padding the system with identity gates requires a total of $O(T/(1 - \delta))$ clock qubits, however one can instead prepare the weighted history state with $\pi_T = \delta$ using only T clock qubits. The price that one has to pay for this improvement is in an

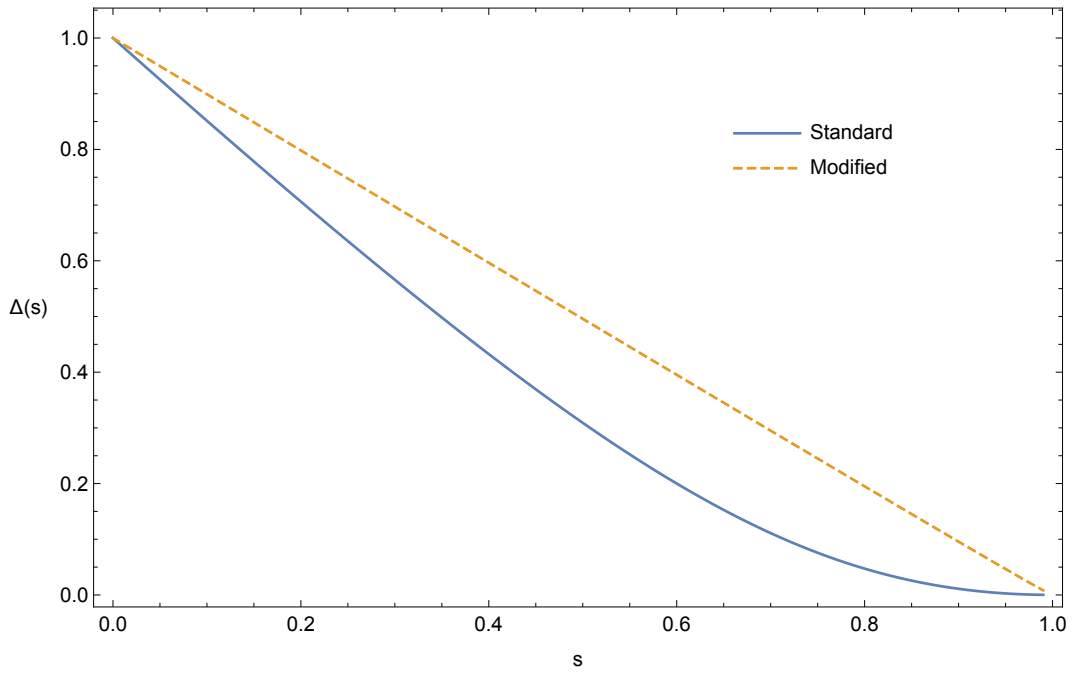


Figure 5.3: Comparison of the spectral gap as a function of the adiabatic parameter s for the standard version of universal AQC with a $1/T$ probability of measuring the history state to be in the final time step of the computation and for the modified construction corresponding to a history state with probability $1/4$ of measuring the final time step of the computation. Here $T = 100$ for both constructions, and they both follow the same adiabatic schedule with local terms of the same norm.

increase in the precision of the couplings needed to implement eq. (3.1) now must scale like $O(T^{-1})$, as seen in (5.15). This is a reasonable trade off, however, since the total number of qubits is generally the limiting factor in most experiments.

Having seen that it can be advantageous to prepare the non-uniform history state ground state of $\mathbf{H}_{\text{prop}}^*$, it remains to be shown that this can be done efficiently. A numerical comparison of the spectral gaps of $\mathbf{H}(s)$ in (5.35) and of $\mathbf{H}^*(s) = (1-s)\mathbf{H}_{\text{init}} + s\mathbf{H}_{\text{prop}}^*(s)$ can be found in figure 5.3, and it reveals that both of these linear interpolations encounter their asymptotic minimum gap at $s = 1$. We have already analysed this spectral gap in section 5.3, and shown that $\Delta_{\mathbf{H}_{\text{prop}}^*} = \Theta(T^{-2})$, such as the minimum gap for the standard construction is $\Delta_{\mathbf{H}(s=1)} = \Theta(T^{-2})$. However, it is of course desirable to bound the adiabatic run time for $\mathbf{H}^*(s)$ without any recourse to numerics. In this case the monotonicity of the ground state wave function used in the proof for the standard construction does not hold, and so we offer a different proof based on modified linear interpolation schedule,

$$\mathbf{H}^*(s) = \mathbf{H}_0 + sA\mathbf{H}_{\text{prop}}^*, \quad (5.36)$$

with $A = T^4$ and an $\Omega(1)$ spectral gap for $0 \leq s \leq 1$, so that the standard estimate of the adiabatic run time $\Omega(\|\dot{\mathbf{H}}\|\Delta_{\mathbf{H}}^{-2})$ is $\Omega(T^4)$ just as it will be for the usual version of universal AQC. At $s = 1$ the system will be in the ground state of the Hamiltonian $\tilde{\mathbf{H}}_{\text{final}} = \mathbf{H}_{\text{prop}}^* + A^{-1}\mathbf{H}_0$, and if $\{|\psi_k^*\rangle\}_{k=0}^T, \{E_k^*\}_{k=0}^T$ are the eigenstates and corresponding eigenvalues of $\mathbf{H}_{\text{prop}}^*$ then to first order in A^{-1} the perturbed ground state $|\tilde{\psi}\rangle$ satisfies

$$\| |\tilde{\psi}\rangle - |\psi_0^*\rangle \|_1 = A^{-1} \sum_{k=1}^T \frac{|\langle \psi_0 | \mathbf{H}_{\text{init}} | \psi_k \rangle|}{E_k^* - E_0^*}, \quad (5.37)$$

and using the fact that $E_k^* - E_0^* \geq \Delta_{\mathbf{H}_{\text{prop}}^*}$ and $\|\mathbf{H}_{\text{init}}\| = 1$ this implies

$$A^{-1} \sum_{k=1}^T \frac{|\langle \psi_0 | \mathbf{H}_{\text{init}} | \psi_k \rangle|}{E_k^* - E_0^*} \leq A^{-1} \Delta_{\mathbf{H}_{\text{prop}}^*}^{-1} \sum_{k=1}^T |\langle \psi_0^* | \mathbf{H}_{\text{init}} | \psi_k^* \rangle| = \mathcal{O}(A^{-1} \Delta_{\mathbf{H}_{\text{prop}}^*}^{-1} T), \quad (5.38)$$

and since $A = T^4$ we have $\| |\tilde{\psi}\rangle - |\psi_0^*\rangle \|_1 = 1 - \mathcal{O}(T^{-1})$ as claimed.

To see that $\mathbf{H}^*(s)$ in (5.36) has $\Delta_{\mathbf{H}^*(s)} = \Omega(1)$ for all s , first note that $\Delta_{\mathbf{H}(0)} = 1$. We want to show that the first excited energy is non-decreasing as a function of s , so to do this we let L be a large integer and discretise the adiabatic path into steps $s_0 = 0, s_1, \dots, s_L = 1$, with a uniform step size $s_{i+1} = s_i + L^{-1}$. Since $\mathbf{H}^*(s_{i+1}) = \mathbf{H}^*(s_i) + (\mathbf{H}^*(s_{i+1}) - \mathbf{H}^*(s_i))$ and the fact that $\mathbf{H}^*(s_{i+1}) - \mathbf{H}^*(s_i) = \mathbf{H}_{\text{prop}}$ is positive semi-definite we can apply Weyl's inequality to see that the first excited state energy is indeed non-decreasing with s (and since this holds for any L it is independent of the discretisation of the path),

and so so $E_1^*(s) \geq 1$. Next we apply the variational method, with a trial state equal to the ground state of \mathbf{H}_{prop} , to see that $\langle \psi_{\text{final}} | \mathbf{H}(s) | \psi \rangle = \langle \psi | \mathbf{H}_0 | \psi \rangle = 3/4$, and so $\Delta_{\mathbf{H}(s)} \geq 1/4$ for all s .

Finally, we note that theorem 5.7 can be interpreted as proving that the standard universal adiabatic construction plus the weighted endpoint modification made above is in a sense optimal for Hamiltonians of the form (3.1). First, the problem of upper bounding the spectral gap of universal adiabatic constructions was addressed before [GS13] by combining the quantum lower bound for unstructured search with the technique of spectral gap amplification. This previous work found a general $\tilde{O}(T^{-1})$ bound on the spectral gap of *any* adiabatic Hamiltonian, a $\tilde{O}(T^{-2})$ gap for any frustration-free adiabatic Hamiltonian, and finally an $\tilde{O}(T^{-2})$ bound on the spectral gap of modified Feynman Hamiltonians of the form (3.1) when the weights near the endpoints satisfy a reasonable assumption for any adiabatic computation. Our theorem 5.7 corroborates this last result by showing a tight $O(T^{-2})$ upper bound on the spectral gap and the minimum overlap of the weighted history state with either endpoint of the computation.

5.6 Chapter Summary

One of the main aims of the present work is to motivate new ideas in quantum ground state computation by focusing on the quantum UNSAT penalty as a metric for the improvement of circuit Hamiltonians. In this section we discuss a range of open problems related to the UNSAT penalty, in case that they are more tractable or lead to a different perspective on some of the open challenges facing this field. One difficulty is that there is at present no general abstract formulation of what it means for a Hamiltonian to have a ground space of circuit histories, as further improvement could involve alterations to the tridiagonal form of the clock Hamiltonian (one such construction allowing for computational paths that include branching, concurrency, and loops is given in [BCO17]). Therefore we describe these open problems without specifying a precise form for future circuit-to-Hamiltonian constructions e.g. how the number of local terms might scale, and so we are implicitly discussing *relative* energy penalties that are not simply made larger by e.g. increasing the overall norm of the Hamiltonian.

The classical baseline. The classical Cook-Levin theorem encodes the history of a classical circuit into the satisfying assignment of a 3-SAT formula. If the computation has T time steps, then the associated constraint satisfaction problem has $O(T)$ local terms and if each has a constant norm than the classical UNSAT penalty is immediately $\Omega(1)$. Therefore we ask: is it possible for a circuit Hamiltonian containing $O(T)$ local terms of bounded norm, which may be of a form more general than (3.1), to achieve an UNSAT penalty that is independent of the length of the computation?

Macroscopic UNSAT penalty. Building on the previous question which asks whether the UNSAT penalty can be made independent of the length of the computation, we further ask whether the UNSAT penalty can be made to scale macroscopically with the number of qubits n in the computation. Specifically, is there a circuit Hamiltonian with $O(\text{poly}(n)T)$ local terms that achieves a $\text{poly}(n)$ UNSAT penalty that is independent of T ? Such a construction could be a useful step towards fault-tolerant adiabatic computation. An intuition for this connection can be gained by considering a construction for energetically encoded fault-tolerant classical computation, whereby each logical bit could be encoded as an arrangement of spins in a self-correcting model (e.g. the 2D Ising model), so that the UNSAT penalty could have a macroscopic scaling (i.e. with the number of physical spins representing each logical bit) that is independent of T .

Constant relative UNSAT penalty. A circuit Hamiltonian with $O(m)$ local terms of bounded norm, where $m = \text{poly}(T)$, with constant relative UNSAT penalty E_p/m would yield a proof of the quantum PCP conjecture by spectral gap amplification. The reduction consists of applying the circuit Hamiltonian with constant relative UNSAT penalty to the circuit verifier that decides the ground state energy of the arbitrary input local Hamiltonian.

It is a testament to Feynman's great legacy that an idea first introduced in 1987 has had such a profound impact for wide scope of research, from condensed matter physics to quantum computation, and that despite the growth of the field of Hamiltonian complexity his original construction continues to remain essentially unchanged to date. We do not know whether or where limitations of improving the circuit-to-Hamiltonian construction will be reached, but hope that our contribution will help to push this boundary a little further.

6 Conclusion

Hamiltonian complexity theory has been a relatively active field of research over the last few years, and the interplay between complexity theory and many-body quantum physics has not come to a standstill.

In [LVV15], the authors prove that for a gapped, local and one-dimensional Hamiltonian \mathbf{H} on n dimension d qudits, approximating the ground state to within precision η can be achieved with a BPP algorithm with runtime $n^{c(d,\epsilon)} \text{poly}(n/\eta)$. Here, ϵ is the spectral gap of the Hamiltonian, and the function $c(d, \epsilon) = \exp(\log^3 d/\epsilon)$; the success probability is $1 - 1/\text{poly } n$.

This result does not depend on the local dimension, and in fact accomplishes something arguably harder than what the LOCAL HAMILTONIAN problem demands: whereas the latter only asks for an approximation of the ground state energy, [LVV15] give a matrix product state representation which is η -close to the true ground state (with regards to state fidelity, i.e. $|\langle \psi | \phi \rangle| \geq 1 - \eta$, where $|\psi\rangle$ approximates the true ground state $|\phi\rangle$). The construction crucially depends on the existence of a finite spectral gap of \mathbf{H} ; but could this result extend to the case where we allow this gap to shrink, e.g. as $\propto 1/\log n$? What if we are only interested in the energy, not in the state?

This suggests a bottom-up approach: for example, we proved (see theorem 2.2) that even for translationally-invariant, nearest-neighbour 1D spin chains with open boundary conditions, approximating the ground state energy is QMA_{EXP} hard. But the local dimension d is greater or equal to 42. What about qubits? It seems unfeasible to create a history state embedding with that little freedom to encode information into local interactions, but there might well be another path of reduction to embed computation, which does not depend on an extra clock register; remember that the no-go theorem in section 2 only claimed to disallow quantum computation when we expect the ground state to be a tensor product over the history; we never claim Feynman's highly-entangled history state constructions are the only possible alternative. But how do we find another feasible embedding?

In a similar context, there exists a complexity classification of interactions of local Hamiltonians, see [CM14]; [PM15]. In brief, the authors prove that \mathcal{S} is a fixed set of k -qubit interactions, then the LOCAL HAMILTONIAN problem with terms from \mathcal{S} is either in P, NP-complete, StoqMA-complete, or QMA-complete. This includes important results such as the quantum Ising model [BH14] being StoqMA-complete. Central to these classifications are perturbation gadgets, i.e. high-weight terms in the Hamiltonian that creates an effective ground state interaction which can have higher degree than its constituents. This

idea is discussed and used ubiquitously in Hamiltonian complexity theory, see [JF08]; [OT05]; [AB09]; [Bran]; [BDLII], yet one major problem is that it either requires coupling strengths that grow with the system size, or growing number of the interactions [CN15].

Both ideas are, in a sense, unphysical, as they describe systems which would not appear in nature; so one open question is: can we build perturbation gadgets with $O(1)$ interaction strengths, and without growing the number of terms in the Hamiltonian? My hope is that novel insights into these topics will lead to a better understanding of how well we can simulate, approximate, or interpolate the behaviour and properties of many-body quantum systems, and that this dissertation has made a small contribution to this undertaking.

Bibliography

- [AAV13] Dorit Aharonov, Itai Arad, and Thomas Vidick. “Guest column: the quantum PCP conjecture”. In: *Acm sigact news* 44.2 (2013), pp. 47–79.
- [AB09] Sanjeev Arora and Boaz Barak. *Computational Complexity*. Cambridge: Cambridge University Press, 2009. ISBN: 9780511804090. DOI: 10.1017/CB09780511804090.
- [Aha+08] Dorit Aharonov et al. “Adiabatic quantum computation is equivalent to standard quantum computation”. In: *SIAM review* 50.4 (2008), pp. 755–787.
- [Aha+09a] Dorit Aharonov et al. “The detectability lemma and quantum gap amplification”. In: *Proceedings of the 41st annual ACM symposium on Symposium on theory of computing - STOC '09*. New York, New York, USA: ACM Press, 2009, p. 417. ISBN: 9781605585062. DOI: 10.1145/1536414.1536472.
- [Aha+09b] Dorit Aharonov et al. “The power of quantum systems on a line”. In: *Communications in Mathematical Physics* 287.1 (May 2009), pp. 41–65. DOI: 10.1007/s00220-008-0710-3. arXiv: 0705.4077.
- [Aid16] Monika Aidelsburger. *Artificial Gauge Fields with Ultracold Atoms in Optical Lattices*. Springer Theses. Cham: Springer International Publishing, 2016. ISBN: 978-3-319-25827-0. DOI: 10.1007/978-3-319-25829-4.
- [AKSo4] Manindra Agrawal, Neeraj Kayal, and Nitin Saxena. “PRIMES is in P”. In: *Annals of Mathematics* 160.2 (Sept. 2004), pp. 781–793. ISSN: 0003-486X. DOI: 10.4007/annals.2004.160.781.
- [AL81] Michael Aizenman and Elliott H. Lieb. “The third law of thermodynamics and the degeneracy of the ground state for lattice systems”. In: *Journal of Statistical Physics* 24.1 (Jan. 1981), pp. 279–297. ISSN: 0022-4715. DOI: 10.1007/BF01007649.
- [AN02] Dorit Aharonov and Tomer Naveh. “Quantum NP-a survey”. In: *arXiv preprint quant-ph/0210077* (2002).
- [Ara+15] Itai Arad et al. “Linear time algorithm for quantum 2SAT”. In: *Qip '16* 27 (Aug. 2015), pp. 1–20. arXiv: 1508.06340.

- [Bab15] László Babai. “Graph Isomorphism in Quasipolynomial Time”. In: (Dec. 2015). arXiv: 1512.03547.
- [Bab17] László Babai. *quasipolynomial claim restored*. 2017.
- [Bar68] Erwin H. Bareiss. “Sylvester’s identity and multistep integer-preserving Gaussian elimination”. In: *Mathematics of Computation* 22.103 (Sept. 1968), pp. 565–578. ISSN: 0025-5718. DOI: 10.1090/S0025-5718-1968-0226829-0.
- [Bar83] M. N. Barber. “Finite-Size Scaling”. In: *Phase transitions and critical phenomena Vol. 8*. Ed. by Cyril Domb and J. Lebowitz. New York: Academic Press, 1983. ISBN: 9780122203084.
- [Bas11] Richard F. Bass. *Stochastic Processes*. 1st ed. Cambridge University Press, 2011, p. 404. ISBN: 110700800X.
- [Bau+16] Johannes Bausch et al. *Size-Driven Quantum Phase Transitions*. Dec. 2016. arXiv: 1512.05687.
- [BB76] Ph. Buffat and J-P. Borel. “Size effect on the melting temperature of gold particles”. In: *Physical Review A* 13.6 (June 1976), pp. 2287–2298. DOI: 10.1103/physreva.13.2287.
- [BBT06] Sergey Bravyi, Arvid J Bessen, and Barbara M Terhal. “Merlin-Arthur games and stoquastic complexity”. In: *arXiv preprint quant-ph/0611021* (2006).
- [BC16a] Johannes Bausch and Elizabeth Crosson. *Increasing the quantum UNSAT penalty of the circuit-to-Hamiltonian construction*. Sept. 2016. arXiv: 1609.08571.
- [BC16b] Johannes Bausch and Toby Cubitt. “The complexity of divisibility”. In: *Linear Algebra and its Applications* 504 (Sept. 2016), pp. 64–107. ISSN: 00243795. DOI: 10.1016/j.laa.2016.03.041. arXiv: 1411.7380.
- [BCO17] Johannes Bausch, Toby Cubitt, and Maris Ozols. “The Complexity of Translationally-Invariant Spin Chains with Low Local Dimension”. In: *accepted for publication in Annales Henri Poincaré* (2017), p. 52. arXiv: 1605.01718.
- [BDL11] Sergey Bravyi, David P. DiVincenzo, and Daniel Loss. “Schrieffer–Wolff transformation for quantum many-body systems”. In: *Annals of Physics* 326.10 (Oct. 2011), pp. 2793–2826. ISSN: 00034916. DOI: 10.1016/j.aop.2011.06.004. arXiv: 1105.0675.
- [Bel10] Mihir Bellare. *Notes on Randomized Algorithms*. 2010.
- [Ben+97] Charles H. Bennett et al. “Strengths and Weaknesses of Quantum Computing”. In: *SIAM Journal on Computing* 26.5 (Oct. 1997), pp. 1510–1523. ISSN: 0097-5397. DOI: 10.1137/S0097539796300933.

- [Ben73] Charles H. Bennett. “Logical reversibility of computation”. In: *IBM Journal of Research and Development* 17.6 (Nov. 1973), pp. 525–532. ISSN: 0018-8646. DOI: 10.1147/rd.176.0525.
- [Ber66] R. Berger. *The Undecidability of the Domino Problem*. American Mathematical Society memoirs. American Mathematical Society, 1966.
- [BG15] Sergey Bravyi and David Gosset. “Gapped and gapless phases of frustration-free spin-1/2 chains”. In: *Journal of Mathematical Physics* 56.6 (June 2015), p. 061902. ISSN: 0022-2488. DOI: 10.1063/1.4922508. arXiv: 1503.04035v1.
- [BG16] Niel de Beaudrap and Sevag Gharibian. *A Linear Time Algorithm for Quantum 2-SAT*. 2016. DOI: 10.4230/LIPIcs.CCC.2016.27.
- [BH12] Thomas Barthel and Robert Hübener. “Solving Condensed-Matter Ground-State Problems by Semidefinite Relaxations”. In: *Physical Review Letters* 108.20 (May 2012), p. 200404. ISSN: 0031-9007. DOI: 10.1103/PhysRevLett.108.200404. arXiv: 1106.4966.
- [BH14] Sergey Bravyi and Matthew Hastings. “On complexity of the quantum Ising model”. In: (Oct. 2014). arXiv: 1410.0703.
- [BL08] Jacob D Biamonte and Peter J Love. “Realizable Hamiltonians for universal adiabatic quantum computers”. In: *Physical Review A* 78.1 (2008), p. 12352.
- [BP17] Johannes Bausch and Stephen Piddock. “The Complexity of Translationally-Invariant Low-Dimensional Spin Lattices in 3D”. In: *accepted for publication in JMP* (Feb. 2017). arXiv: 1702.08830.
- [BR97] Ola Bratteli and Derek W. Robinson. *Operator Algebras and Quantum Statistical Mechanics* 2. Springer Science + Business Media, 1997. DOI: 10.1007/978-3-662-03444-6.
- [Bran] Sergey Bravyi. “Efficient algorithm for a quantum analogue of 2-SAT”. In: Feb. 2011, pp. 33–48. DOI: 10.1090/conm/536/10552. arXiv: 0602108 [quant-ph].
- [Bra83] Allen H Brady. “The determination of the value of Rado’s noncomputable function $\Sigma(k)$ for four-state Turing machines”. In: *Mathematics of Computation* 40.162 (1983), pp. 647–665.
- [BT09] Sergey Bravyi and Barbara Terhal. “Complexity of stoquastic frustration-free Hamiltonians”. In: *Siam journal on computing* 39.4 (2009), pp. 1462–1485.
- [BT14a] Nikolas P. Breuckmann and Barbara M. Terhal. “Space-time circuit-to-Hamiltonian construction and its applications”. In: *Journal of Physics A: Mathematical and Theoretical* 47.19 (May 2014), p. 195304. ISSN: 1751-8113. DOI: 10.1088/1751-8113/47/19/195304. arXiv: 1311.6101.

- [BT14b] Nikolas P Breuckmann and Barbara M Terhal. “Space-time circuit-to-Hamiltonian construction and its applications”. In: *Journal of Physics A: Mathematical and Theoretical* 47.19 (2014), p. 195304.
- [BV97a] Ethan Bernstein and Umesh Vazirani. “Quantum Complexity Theory”. In: *SIAM Journal on Computing* 26.5 (Oct. 1997), pp. 1411–1473. ISSN: 0097-5397. DOI: 10.1137/S0097539796300921.
- [BV97b] Ethan Bernstein and Umesh Vazirani. “Quantum complexity theory”. In: *SIAM Journal on Computing* 26.5 (1997), pp. 1411–1473. DOI: 10.1137/S0097539796300921.
- [BŻ06] Ingemar Bengtsson and Karol Życzkowski. *Geometry of quantum states: an introduction to quantum entanglement*. Cambridge University Press, 2006. ISBN: 052189140X.
- [CEW12a] Toby S. Cubitt, Jens Eisert, and Michael M. Wolf. “Extracting Dynamical Equations from Experimental Data is NP Hard”. In: *Physical Review Letters* 108.12 (Mar. 2012), p. 120503. ISSN: 0031-9007. DOI: 10.1103/PhysRevLett.108.120503. arXiv: arXiv:1005.0005v2.
- [CEW12b] Toby S. Cubitt, Jens Eisert, and Michael M. Wolf. “The Complexity of Relating Quantum Channels to Master Equations”. In: *Communications in Mathematical Physics* 310.2 (Jan. 2012), pp. 383–418. ISSN: 0010-3616. DOI: 10.1007/s00220-011-1402-y.
- [Che11] Jianxin Chen et al. “No-go theorem for one-way quantum computing on naturally occurring two-level systems”. In: *Physical Review A* 83.5 (May 2011), p. 050301. ISSN: 1050-2947. DOI: 10.1103/PhysRevA.83.050301. arXiv: 1004.3787.
- [Che70] J Cheeger. “A lower bound for the smallest eigenvalue of the Laplacian, Problems in Analysis”. In: *Papers dedicated to Salomon Bochner (1969)*. Princeton Univ. Press, Princeton, N. J., 1970, pp. 195–199.
- [Chi10] Andrew M. Childs et al. “Characterization of universal two-qubit Hamiltonians”. In: (Apr. 2010). arXiv: 1004.1645.
- [Cho75] Man-Duen Choi. “Completely positive linear maps on complex matrices”. In: *Linear Algebra and its Applications* 10.3 (1975), pp. 285–290. ISSN: 00243795. DOI: 10.1016/0024-3795(75)90075-0.
- [CM14] Toby S. Cubitt and Ashley Montanaro. “Complexity Classification of Local Hamiltonian Problems”. In: *2014 IEEE 55th Annual Symposium on Foundations of Computer Science*. IEEE, Oct. 2014, pp. 120–129. ISBN: 978-1-4799-6517-5. DOI: 10.1109/FOCS.2014.21. arXiv: 1311.3161.
- [CM16] Toby Cubitt and Ashley Montanaro. “Complexity classification of local Hamiltonian problems”. In: *SIAM Journal on Computing* 45.2 (2016), pp. 268–316.

- [CN15] Yudong Cao and Daniel Nagaj. “Perturbative gadgets without strong interactions”. In: *Quantum Information & Computation* 15.13 (Aug. 2015), pp. 1197–1222. arXiv: 1408.5881.
- [CNN16] Matthew Cha, Pieter Naaijkens, and Bruno Nachtergaele. *The complete set of infinite volume ground states for Kitaev’s abelian quantum double models*. Aug. 2016. arXiv: 1608.04449v1 [math-ph].
- [Coo71] Stephen A. Cook. “The complexity of theorem-proving procedures”. In: *Proceedings of the third annual ACM symposium on Theory of computing - STOC ’71*. New York, New York, USA: ACM Press, 1971, pp. 151–158. DOI: 10.1145/800157.805047.
- [CPC17] Arturo Camacho-Guardian, Rosario Paredes, and Santiago F. Caballero-Benitez. “Quantum Simulation of Competing Orders with Fermions in Quantum Optical Lattices”. In: (June 2017). arXiv: 1706.02347.
- [CPW15a] Toby S. Cubitt, David Perez-Garcia, and Michael M. Wolf. “Undecidability of the spectral gap”. In: *Nature* 528.7581 (Dec. 2015), pp. 207–211. ISSN: 0028-0836. DOI: 10.1038/nature16059. arXiv: 1502.04573.
- [CPW15b] Toby S. Cubitt, David Perez-Garcia, and Michael M. Wolf. “Undecidability of the spectral gap”. In: *Nature* 528.7581 (Dec. 2015), pp. 207–211. ISSN: 0028-0836. DOI: 10.1038/nature16059.
- [Cra36] Harald Cramér. “Über eine Eigenschaft der normalen Verteilungsfunktion”. In: *Mathematische Zeitschrift* 41.1 (Dec. 1936), pp. 405–414. ISSN: 0025-5874. DOI: 10.1007/BF01180430.
- [CS13] Guan-Yu Chen and Laurent Saloff-Coste. “On the mixing time and spectral gap for birth and death chains”. In: *ALEA Lat. Am. J. Probab. Math. Stat.* 10.1 (2013), pp. 293–321. arXiv: 1304.4346.
- [Cub15] Toby Cubitt. *Lecture notes in Advanced Quantum Information Theory*. 2015.
- [CW34] W. G. Cochran and J. Wishart. “The distribution of quadratic forms in a normal system, with applications to the analysis of covariance”. English. In: *Mathematical Proceedings of the Cambridge Philosophical Society* 30.02 (Apr. 1934), p. 178. ISSN: 0305-0041. DOI: 10.1017/S0305004100016595.
- [CWG08] Theodore Charitos, Peter R de Waal, and Linda C van der Gaag. “Computing short-interval transition matrices of a discrete-time Markov chain from partially observed data”. In: *Statistics in Medicine* 27.6 (Mar. 2008), pp. 905–921. ISSN: 02776715. DOI: 10.1002/sim.2970.
- [Dav58] M Davis. *Computability & Unsolvability*. Dover Books on Computer Science Series. Dover, 1958, p. 248. ISBN: 9780486614717.

- [Dier0] Reinhard Diestel. *Graph Theory*. 4th ed. Springer Berlin Heidelberg, 2010, pp. XVIII, 410. ISBN: 3642142788.
- [DKP14] John Dengis, Robert König, and Fernando Pastawski. “An optimal dissipative encoder for the toric code”. In: *New Journal of Physics* 16.1 (Jan. 2014), p. 013023. DOI: 10.1088/1367-2630/16/1/013023.
- [ED79] John E. Hopcroft and Jeffrey D. Ullman. *Introduction to automata theory, languages, and computation*. Addison-Wesley, 1979. ISBN: 9780201029888.
- [Elf37] Gustav Elfving. “Zur Theorie der Markoffschen Ketten”. In: *Acta Soc. Sci. Fennicae n. Ser. A* 2.8 (1937), pp. 1–17.
- [ELNo4] Patricia D. Egleston, Terry D. Lenker, and Sivaram K. Narayan. “The nonnegative inverse eigenvalue problem”. In: *Linear Algebra and its Applications* 379 (Mar. 2004), pp. 475–490. ISSN: 00243795. DOI: 10.1016/j.laa.2003.10.019.
- [ERS75] R. Eisberg, R. Resnick, and Jeremiah D. Sullivan. “Quantum Physics of Atoms, Molecules, Solids, Nuclei and Particles”. In: *Physics Today* 28.12 (Dec. 1975), pp. 51–52. ISSN: 0031-9228. DOI: 10.1063/1.3069243.
- [Far+00] Edward Farhi et al. “Quantum Computation by Adiabatic Evolution”. In: (Jan. 2000). arXiv: 0001106 [quant-ph].
- [Far+11] Edward Farhi et al. “Unstructured Randomness, Small Gaps and Localization”. In: *Quantum Information & Computation* 11.9–10 (Sept. 2011). eprint: 1010.0009.
- [Fero8] Pablo Fernández. “Google’s pagerank and beyond: The science of search engine rankings”. In: *The Mathematical Intelligencer* 30.1 (Mar. 2008), pp. 68–69. ISSN: 0343-6993. DOI: 10.1007/BF02985759.
- [Fey86] Richard P. Feynman. “Quantum mechanical computers”. In: *Foundations of Physics* 16.6 (June 1986), pp. 507–531. ISSN: 0015-9018. DOI: 10.1007/BF01886518.
- [Fie73] Miroslav Fiedler. “Algebraic connectivity of graphs”. In: *Czechoslovak Mathematical Journal* 23.2 (1973), pp. 298–305. DOI: 10338.dmlcz/101168.
- [GEA92] A. N. Goldstein, C. M. Echer, and A. P. Alivisatos. “Melting in Semiconductor Nanocrystals”. In: *Science* 256.5062 (June 1992), pp. 1425–1427. DOI: 10.1126/science.256.5062.1425.
- [Gha+14] Sevag Gharibian et al. “Quantum Hamiltonian complexity”. In: *Foundations and Trends in Theoretical Computer Science* 10.3 (2014), pp. 159–282. ISSN: 1551-305X. DOI: 10.1561/04000000066. arXiv: 1401.3916.

- [GI09] Daniel Gottesman and Sandy Irani. “The Quantum and Classical Complexity of Translationally Invariant Tiling and Hamiltonian Problems”. In: *Theory of Computing* 9.1 (May 2009), pp. 31–116. ISSN: 1557-2862. DOI: 10 . 4086 / toc . 2013 . v009a002. arXiv: 0905 . 2419.
- [GI13] Daniel Gottesman and Sandy Irani. “The Quantum and Classical Complexity of Translationally Invariant Tiling and Hamiltonian Problems”. In: *Theory of Computing* 9.1 (May 2013), pp. 31–116. ISSN: 1557-2862. DOI: 10 . 4086 / toc . 2013 . v009a002. arXiv: 0905 . 2419.
- [GJ79] Michael R. Garey and David S. Johnson. *Computers and Intractability: A Guide to the Theory of NP-Completeness*. W. H. Freeman & Co., Jan. 1979. ISBN: 0716710447.
- [GN13] David Gosset and Daniel Nagaj. “Quantum 3-SAT Is QMA1-Complete”. In: *2013 IEEE 54th Annual Symposium on Foundations of Computer Science*. IEEE, Oct. 2013, pp. 756–765. ISBN: 978-0-7695-5135-7. DOI: 10 . 1109 / FOCS . 2013 . 86.
- [Gor76] Vittorio Gorini. “Completely positive dynamical semigroups of N-level systems”. In: *Journal of Mathematical Physics* 17.5 (Aug. 1976), p. 821. ISSN: 00222488. DOI: 10 . 1063 / 1 . 522979.
- [GR01] Chris Godsil and Gordon Royle. *Algebraic Graph Theory*. Vol. 207. Graduate Texts in Mathematics. New York, NY: Springer, 2001. ISBN: 978-0-387-95220-8. DOI: 10 . 1007 / 978 - 1 - 4613 - 0163 - 9.
- [GS13] Anand Ganti and Rolando Somma. “On the gap of Hamiltonians for the adiabatic simulation of quantum circuits”. In: *International Journal of Quantum Information* 11.07 (July 2013), p. 1350063. ISSN: 0219-7499. DOI: 10 . 1142 / S0219749913500639. arXiv: 1307 . 4993.
- [GTV15] David Gosset, Barbara M. Terhal, and Anna Vershynina. “Universal Adiabatic Quantum Computation via the Space-Time Circuit-to-Hamiltonian Construction”. In: *Physical Review Letters* 114.14 (Apr. 2015), p. 140501. ISSN: 0031-9007. DOI: 10 . 1103 / PhysRevLett . 114 . 140501. arXiv: 1409 . 7745.
- [Gui+15] T. Guidi et al. “Direct observation of finite size effects in chains of antiferromagnetically coupled spins”. In: *Nature Communications* 6 (2015), p. 7061.
- [Hal83] F. D. M. Haldane. “Nonlinear Field Theory of Large-Spin Heisenberg Antiferromagnets: Semiclassically Quantized Solitons of the One-Dimensional Easy-Axis Néel State”. In: *Physical Review Letters* 50.15 (Apr. 1983), pp. 1153–1156. ISSN: 0031-9007. DOI: 10 . 1103 / PhysRevLett . 50 . 1153.

- [Haso4] M. B. Hastings. “Lieb-Schultz-Mattis in higher dimensions”. In: *Physical Review B* 69.10 (Mar. 2004), p. 104431. ISSN: 1098-0121. DOI: 10.1103/PhysRevB.69.104431.
- [Haso7] Matthew B Hastings. “An area law for one-dimensional quantum systems”. In: *Journal of Statistical Mechanics: Theory and Experiment* 2007.08 (Aug. 2007), Po8024–Po8024. ISSN: 1742-5468. DOI: 10.1088/1742-5468/2007/08/P08024. arXiv: 0705.2024.
- [HG03] Qi-Ming He and Eldon Gunn. “A note on the stochastic roots of stochastic matrices”. In: *Journal of Systems Science and Systems Engineering* 12.2 (June 2003), pp. 210–223. ISSN: 1004-3756. DOI: 10.1007/s11518-006-0131-9.
- [HHN11] William Hart, Mark van Hoeij, and Andrew Novocin. “Practical polynomial factoring in polynomial time”. In: *Proceedings of the 36th international symposium on Symbolic and algebraic computation - ISSAC '11*. New York, New York, USA: ACM Press, June 2011, p. 163. ISBN: 9781450306751. DOI: 10.1145/1993886.1993914.
- [HHW67] R. Haag, N. M. Hugenholtz, and M. Winnink. “On the equilibrium states in quantum statistical mechanics”. In: *Communications in Mathematical Physics* 5.3 (June 1967), pp. 215–236. ISSN: 0010-3616. DOI: 10.1007/BF01646342.
- [Hig87] Nicholas J. Higham. “Computing real square roots of a real matrix”. In: *Linear Algebra and its Applications* 88-89.1987 (Apr. 1987), pp. 405–430. ISSN: 00243795. DOI: 10.1016/0024-3795(87)90118-2.
- [HL11] Nicholas J. Higham and Lijing Lin. “On p th roots of stochastic matrices”. In: *Linear Algebra and its Applications* 435.3 (Aug. 2011), pp. 448–463. ISSN: 00243795. DOI: 10.1016/j.laa.2010.04.007.
- [HNN13] Sean Hallgren, Daniel Nagaj, and Sandeep Narayanaswami. “The Local Hamiltonian problem on a line with eight states is QMA-complete”. In: *Quantum Information and Computation* 13.9&10 (Dec. 2013), p. 28. arXiv: 1312.1469.
- [Ira07] Sandy Irani. “The Complexity of Quantum Systems on a One-dimensional Chain”. In: (May 2007). arXiv: 0705.4067.
- [Jar97] R. A. Jarrow. “A Markov model for the term structure of credit risk spreads”. In: *Review of Financial Studies* 10.2 (Apr. 1997), pp. 481–523. ISSN: 14657368. DOI: 10.1093/rfs/10.2.481.
- [JFo8] Stephen P. Jordan and Edward Farhi. “Perturbative gadgets at arbitrary orders”. In: *Physical Review A* 77.6 (June 2008), p. 062329. ISSN: 1050-2947. DOI: 10.1103/PhysRevA.77.062329. arXiv: 0802.1874.

- [Jor75] Camille Jordan. “Essai sur la géométrie à n dimensions”. In: *Bulletin de la Société mathématique de France* 3 (1875), pp. 103–174. ISSN: 0037-9484.
- [Kat77] S. K. Katti. “Infinite divisibility of discrete distributions. III.” 1977.
- [Kin62] John Frank Charles Kingman. “The imbedding problem for finite Markov chains”. In: *Probability Theory and Related Fields* 1.1 (1962), pp. 14–24.
- [Kito03] A Yu Kitaev. “Fault-tolerant quantum computation by anyons”. In: *Annals of Physics* 303.1 (2003), pp. 2–30.
- [KKo8] Pok-Son Kim and Arne Kutzner. “Ratio Based Stable In-Place Merging”. In: *Theory and Applications of Models of Computation*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2008, pp. 246–257. DOI: 10.1007/978-3-540-79228-4_22.
- [KKRo6] Julia Kempe, Alexei Yu. Kitaev, and Oded Regev. “The Complexity of the Local Hamiltonian Problem”. In: *SIAM Journal on Computing* 35.5 (Jan. 2006), pp. 1070–1097. ISSN: 0097-5397. DOI: 10.1137/S0097539704445226. arXiv: 0406180 [quant-ph].
- [KO17] Fumitaka Kagawa and Hiroshi Oike. “Quenching of Charge and Spin Degrees of Freedom in Condensed Matter”. In: *Advanced Materials* 29.25 (July 2017), p. 1601979. ISSN: 09359648. DOI: 10.1002/adma.201601979.
- [Kog83] John B. Kogut. “The lattice gauge theory approach to quantum chromodynamics”. In: *Rev. Mod. Phys.* 55 (3 July 1983), pp. 775–836. DOI: 10.1103/RevModPhys.55.775.
- [Kro] Pavel Kropitz. *6-state 2-symbol #b*. http://www.drb.insel.de/~heiner/BB/simKro62_b.html. accessed: April 27th, 2015.
- [KSVo2] Alexei Yu. Kitaev, Alexander Shen, and Mikhail N. Vyalyi. “Classical and quantum computing”. In: *Quantum Information*. New York, NY: Springer New York, 2002, pp. 203–217. DOI: 10.1007/978-0-387-36944-0_13.
- [Lea03] A. E. Leanhardt. “Cooling Bose-Einstein Condensates Below 500 Picokelvin”. In: *Science* 301.5639 (Sept. 2003), pp. 1513–1515. ISSN: 0036-8075. DOI: 10.1126/science.1088827.
- [Li+16] Dehui Li et al. “Size-dependent phase transition in methylammonium lead iodide perovskite microplate crystals”. In: *Nature Communications* 7 (Apr. 2016), p. 11330. DOI: 10.1038/ncomms11330.
- [Lin11] Lijing Lin. “Roots of Stochastic Matrices and Fractional Matrix Powers”. PhD thesis. University of Manchester, 2011.
- [Lin76] Göran Lindblad. “On the generators of quantum dynamical semigroups”. In: *Communications in Mathematical Physics* 48.2 (1976), pp. 119–130. ISSN: 1432-0916.

- [Liu+01] C S Liu et al. “Cooling rate dependence of structural properties of aluminium during rapid solidification”. In: *Journal of Physics: Condensed Matter* 13.9 (Mar. 2001), pp. 1873–1890. ISSN: 0953-8984. DOI: 10.1088/0953-8984/13/9/311.
- [Lju87] Lennart Ljung. *System identification: theory for the user*. Prentice-Hall, 1987, p. 519. ISBN: 0138816409.
- [LL80] L D Landau and E.M. Lifshitz. *Statistical Physics, Third Edition, Part 1: Volume 5 (Course of Theoretical Physics, Volume 5)*. Butterworth-Heinemann, 1980. ISBN: 0750633727.
- [LLS01] D.P. Landau, S.P. Lewis, and H.B. Schüttler. *Computer Simulation Studies in Condensed-Matter Physics XIII: Proceedings of the Thirteenth Workshop Athens, Ga, Usa, February 21-25, 2000*. Computer Simulation Studies in Condensed-matter Physics. Springer, 2001. ISBN: 9783540411901.
- [Lor78] A. E. Loring. *A Hand-book of the Electromagnetic Telegraph*. D. Van Nostrand, 1878, p. 98.
- [Lou+08] S. Lounis et al. “Magnetism of nanowires driven by novel even-odd effects”. In: *Phys. Rev. Lett.* 101 (2008), p. 107204.
- [LPW09] David Asher Levin, Yuval Peres, and Elizabeth Lee Wilmer. *Markov chains and mixing times*. American Mathematical Soc., 2009.
- [LR65] Shen Lin and Tibor Rado. “Computer studies of Turing machine problems”. In: *Journal of the ACM (JACM)* 12.2 (1965), pp. 196–212.
- [LSM61] Elliott Lieb, Theodore Schultz, and Daniel Mattis. “Two soluble models of an antiferromagnetic chain”. In: *Annals of Physics* 16.3 (Dec. 1961), pp. 407–466. ISSN: 00034916. DOI: 10.1016/0003-4916(61)90115-4.
- [LVV15] Zeph Landau, Umesh Vazirani, and Thomas Vidick. “A polynomial time algorithm for the ground state of one-dimensional gapped local Hamiltonians”. In: *Nature Physics* 11.7 (June 2015), pp. 566–569. ISSN: 1745-2473. DOI: 10.1038/nphys3345. arXiv: 1307.5143.
- [Mar+90] Heiner Marxen et al. “Attacking the busy beaver 5”. In: *Bull EATCS*. <http://www.drb.insel.de/~heiner/BB/mabu90.html>, accessed: April 27th, 2015. 1990.
- [Mar92] Norman H. March. *Electron Density Theory of Atoms and Molecules*. London: Academic Press, 1992. ISBN: 9780124705258.
- [McG14] Catherine C. McGeoch. “Adiabatic Quantum Computation and Quantum Annealing: Theory and Practice”. In: *Synthesis Lectures on Quantum Computing* 5.2 (July 2014), pp. 1–93. ISSN: 1945-9726. DOI: 10.2200/S00585ED1V01Y201407QMC008.

- [McH97] J. M. McHale. “Surface Energies and Thermodynamic Phase Stability in Nanocrystalline Aluminas”. In: *Science* 277.5327 (Aug. 1997), pp. 788–791. DOI: 10.1126/science.277.5327.788.
- [MD11] D C McKay and B DeMarco. “Cooling in strongly correlated optical lattices: prospects and challenges”. In: *Reports on Progress in Physics* 74.5 (May 2011), p. 054401. ISSN: 0034-4885. DOI: 10.1088/0034-4885/74/5/054401. arXiv: 1010.0198.
- [Min88] Henryk Minc. *Nonnegative Matrices*. Wiley, 1988, p. 206. ISBN: 0471839663.
- [Moro8] Kenichi Morita. “Reversible computing and cellular automata—A survey”. In: *Theoretical Computer Science* 395.1 (2008), pp. 101–131. ISSN: 0304-3975. DOI: 10.1016/j.tcs.2008.01.041.
- [MRW15] Alexander Muller-Hermes, David Reeb, and Michael M. Wolf. “Quantum Subdivision Capacities and Continuous-Time Quantum Coding”. In: *IEEE Transactions on Information Theory* 61.1 (Jan. 2015), pp. 565–581. ISSN: 0018-9448. DOI: 10.1109/TIT.2014.2366456. arXiv: 1310.2856.
- [MS14] Ramis Movassagh and Peter W. Shor. “Power law violation of the area law in quantum spin chains”. In: *arXiv preprint arXiv:1408.1657* (Aug. 2014), p. 37. arXiv: 1408.1657.
- [Mül87] Norbert Th. Müller. “Uniform computational complexity of Taylor series”. In: *Automata, Languages and Programming. 14th International Colloquium, Proceedings*. Ed. by Thomas Ottmann. Vol. 267. Lecture Notes in Computer Science. Berlin, Heidelberg: Springer Berlin Heidelberg, Jan. 1987, pp. 435–444. ISBN: 978-3-540-18088-3. DOI: 10.1007/3-540-18088-5.
- [Nago8] Daniel Nagaj. “Local Hamiltonians in Quantum Computation”. In: (Aug. 2008). arXiv: 0808.2117.
- [Nag12] Daniel Nagaj. “Universal two-body-Hamiltonian quantum computing”. In: *Physical Review A* 85.3 (Mar. 2012), p. 032330. ISSN: 1050-2947. DOI: 10.1103/PhysRevA.85.032330. arXiv: 1002.0420.
- [Nag14] Daniel Nagaj. *Tick-tock Goes the Clock*. Lecture at the Simons Institute. Feb. 2014.
- [NC10] Michael A. Nielsen and Isaac L. Chuang. *Quantum Computation and Quantum Information*. Cambridge: Cambridge University Press, 2010, p. 676. ISBN: 9780511976667. DOI: 10.1017/CB09780511976667.
- [NKL98] Michael A. Nielsen, E Knill, and R. Laflamme. “Complete quantum teleportation using nuclear magnetic resonance”. In: *Nature* 396.6706 (Nov. 1998), p. 15. ISSN: 0028-0836. DOI: 10.1038/23891. arXiv: 9811020 [quant-ph].

- [NW08] Daniel Nagaj and Pawel Wocjan. “Hamiltonian quantum cellular automata in one dimension”. In: *Physical Review A* 78.3 (Sept. 2008), p. 032311. ISSN: 1050-2947. DOI: 10.1103/PhysRevA.78.032311.
- [Oli+14] K.A. Olive et al. “Review of Particle Physics”. In: *Chin.Phys.* C38 (2014), p. 090001. DOI: 10.1088/1674-1137/38/9/090001.
- [OT05] Roberto I. Oliveira and Barbara M. Terhal. “The complexity of quantum spin systems on a two-dimensional square lattice”. In: *Quantum Information & Computation* (Apr. 2005), pp. 1–23. arXiv: 0504050 [quant-ph].
- [Ozo16] Maris Ozols. “Private Communication”. 2016.
- [Pac12] Jiannis K Pachos. *Introduction to topological quantum computation*. Cambridge University Press, 2012.
- [Pat14] Matthew J. Patitz. “An introduction to tile-based self-assembly and a survey of recent results”. In: *Natural Computing* 13.2 (June 2014), pp. 195–224. ISSN: 1567-7818. DOI: 10.1007/s11047-013-9379-4.
- [Pir+12] B. Pirvu et al. “Matrix product states for critical spin chains: Finite-size versus finite-entanglement scaling”. In: *Physical Review B* 86.7 (Aug. 2012). DOI: 10.1103/physrevb.86.075117.
- [PM15] Stephen Piddock and Ashley Montanaro. “The complexity of antiferromagnetic interactions and 2D lattices”. In: (June 2015), p. 35. arXiv: 1506.04014.
- [Pre15] John Preskill. “Lecture notes for Quantum Information”. In: *Lectured at California Institute of Technology, as course Ph219/CS219*. 2015.
- [Rad62] Tibor Rado. “On Non-Computable Functions”. In: *Bell System Technical Journal* 41.3 (1962), pp. 877–884.
- [Riv+11] Jessy B. Rivest et al. “Size Dependence of a Temperature-Induced Solid–Solid Phase Transition in Copper(I) Sulfide”. In: *The Journal of Physical Chemistry Letters* 2.19 (Oct. 2011), pp. 2402–2406. DOI: 10.1021/jz2010144.
- [Rob71a] Raphael M Robinson. “Undecidability and nonperiodicity for tilings of the plane”. In: *Inventiones mathematicae* 12.3 (1971), pp. 177–209.
- [Rob71b] Raphael M. Robinson. “Undecidability and nonperiodicity for tilings of the plane”. In: *Inventiones mathematicae* 12.3 (Sept. 1971), pp. 177–209. ISSN: 0020-9910. DOI: 10.1007/BF01418780.

- [Ryl+16] R. E. Ryltsev et al. “Cooling rate dependence of simulated Cu_{64.5}Zr_{35.5} metallic glass structure”. In: *The Journal of Chemical Physics* 145.3 (July 2016), p. 034506. ISSN: 0021-9606. DOI: 10.1063/1.4958631.
- [SP13] Kalyan S. Perumalla. *Introduction to reversible computing*. Chapman & Hall/CRC Computational Science Series. CRC Press, 2013. ISBN: 9781439873403.
- [Sal07] David Salomon. *Data Compression*. 4th ed. London: Springer London, 2007, pp. XXVIII, 1092. ISBN: 978-1-84628-602-5. DOI: 10.1007/978-1-84628-603-2.
- [Sax14] Nitin Saxena. “Progress on Polynomial Identity Testing-II”. In: *Perspectives in Computational Complexity*. Cham: Springer International Publishing, Jan. 2014, pp. 131–146. DOI: 10.1007/978-3-319-05446-9_7. arXiv: 1401.0976.
- [ŠB13] L. Šamaj and Z. Bajnok. *Introduction to the Statistical Physics of Integrable Many-body Systems*. Cambridge University Press, 2013. ISBN: 9781107030435.
- [Sha92] Adi Shamir. “IP = PSPACE”. In: *Journal of the ACM* 39.4 (Oct. 1992), pp. 869–877. ISSN: 00045411. DOI: 10.1145/146585.146609.
- [SJ89] Alistair Sinclair and Mark Jerrum. “Approximate counting, uniform generation and rapidly mixing Markov chains”. In: *Information and Computation* 82.1 (1989), pp. 93–133.
- [SK79] Fred W Steutel and JT Kent. “Infinite divisibility in theory and practice”. In: *Scandinavian Journal of Statistics* 6.2 (1979), pp. 57–64.
- [Sti55] W. Forrest Stinespring. “Positive Functions on C*-Algebras”. In: *Proceedings of the American Mathematical Society* 6.2 (Apr. 1955), p. 211. ISSN: 00029939. DOI: 10.2307/2032342.
- [TA94] S. H. Tolbert and A. P. Alivisatos. “Size Dependence of a First Order Solid-Solid Phase Transition: The Wurtzite to Rock Salt Transformation in CdSe Nanocrystals”. In: *Science* 265.5170 (July 1994), pp. 373–376. DOI: 10.1126/science.265.5170.373.
- [Tag+08] L. Tagliacozzo et al. “Scaling of entanglement support for matrix product states”. In: *Physical Review B* 78.2 (July 2008). DOI: 10.1103/physrevb.78.024410.
- [Tho10] Wolfgang Thomas. “„When nobody else dreamed of these things“ – Axel Thue und die Termersetzung”. In: *Informatik-Spektrum* 33.5 (Oct. 2010), pp. 504–508. ISSN: 0170-6012. DOI: 10.1007/s00287-010-0468-9.
- [Tho77a] Olof Thorin. “On the infinite divisibility of the Pareto distribution”. In: *Scandinavian Actuarial Journal* 1977.1 (Jan. 1977), pp. 31–40. ISSN: 0346-1238. DOI: 10.1080/03461238.1977.10405623.

- [Tho77b] Olof Thorin. “On the infinite divisibility of the lognormal distribution”. In: *Scandinavian Actuarial Journal* 1977.3 (Mar. 1977), pp. 121–148. ISSN: 0346-1238. DOI: 10.1080/03461238.1977.10405635.
- [Tru13] Richard J. Trudeau. *Introduction to graph theory*. Dover Books on Mathematics. Dover Publications, 2013. ISBN: 9780486318660.
- [VW16] Thomas Vidick and John Watrous. “Quantum Proofs”. In: *Foundations and Trends in Theoretical Computer Science* 11.1–2 (2016), pp. 1–215. ISSN: 1551-305X. DOI: 10.1561/04000000068.
- [WA67] Frederick V. Waugh and Martin E. Abel. “On Fractional Powers of a Matrix”. In: *Journal of the American Statistical Association* 62.319 (Sept. 1967), pp. 1018–1021. ISSN: 0162-1459. DOI: 10.1080/01621459.1967.10500913.
- [Wat12a] John Watrous. “Quantum Computational Complexity”. In: *Computational Complexity*. Ed. by Robert A. Meyers. New York, NY: Springer New York, 2012, pp. 2361–2387. ISBN: 978-0-387-75888-6. DOI: 10.1007/978-1-4614-1800-9_147. arXiv: 0804.3401.
- [Wat12b] John Watrous. “Quantum computational complexity”. In: *Computational Complexity: Theory, Techniques, and Applications*. Ed. by Robert A. Meyers. Springer, 2012, pp. 2361–2387. ISBN: 978-1-4614-1799-6. DOI: 10.1007/978-1-4614-1800-9_147. arXiv: 0804.3401.
- [WCo8] Michael M. Wolf and J. Ignacio Cirac. “Dividing Quantum Channels”. In: *Communications in Mathematical Physics* 279.1 (Feb. 2008), pp. 147–168. ISSN: 0010-3616. DOI: 10.1007/s00220-008-0411-y.
- [Wen13] Xiao-Gang Wen. “Topological Order: From Long-Range Entangled Quantum Matter to a Unified Origin of Light and Electrons”. In: *ISRN Condensed Matter Physics* 2013 (2013), pp. 1–20. DOI: 10.1155/2013/198710.
- [Whi72] Hassler Whitney. *Complex analytic varieties (Addison-Wesley Series in Mathematics)*. Addison-Wesley, 1972, p. 399. ISBN: 0-201-08653-0.
- [WL15] Tzu-Chieh Wei and John C. Liang. “Hamiltonian quantum computer in one dimension”. In: *Physical Review A* 92.6 (Dec. 2015), p. 062334. ISSN: 1050-2947. DOI: 10.1103/PhysRevA.92.062334. arXiv: 1512.06775.
- [Wol02] Stephen Wolfram. *A New Kind of Science*. 1st ed. Vol. 33. 4. Wolfram Media Inc, 2002, p. 1197. ISBN: 1-57955-008-8.
- [Wol08] Michael M. Wolf. *Lecture notes on Quantum Channels and Operations*. 2008.
- [Xi+15] Xiaoxiang Xi et al. “Strongly enhanced charge-density-wave order in monolayer NbSe₂”. In: *Nature Nanotechnology* 10.9 (July 2015), pp. 765–769. DOI: 10.1038/nnano.2015.143.

- [Yep01] Jeffrey Yepez. “Quantum lattice-gas model for computational fluid dynamics”. In: *Phys. Rev. E* 63 (4 Mar. 2001), p. 046702. DOI: 10.1103/PhysRevE.63.046702.
- [YSNo2] Alexei Yu. Kitaev, Alexander Shen, and Mikhail N. Vyalyi. *Classical and quantum computing*. Graduate Studies in Mathematics 47. American Mathematical Society, 2002. ISBN: 9780821832295.
- [Yu+15] Yijun Yu et al. “Gate-tunable phase transitions in thin flakes of 1T-TaS₂”. In: *Nature Nanotechnology* 10.3 (Jan. 2015), pp. 270–276. DOI: 10.1038/nnano.2014.323.