

# Two Bases Suffice for $\text{QMA}_1$ -Completeness

Henry Ma  

CSAIL, Massachusetts Institute of Technology, Cambridge, MA, USA

Anand Natarajan  

CSAIL, Massachusetts Institute of Technology, Cambridge, MA, USA

---

## Abstract

We introduce a *basis-restricted* variant of the QUANTUM- $k$ -SAT problem, in which each term in the input Hamiltonian is required to be diagonal in either the standard or Hadamard basis. Our main result is that the QUANTUM-6-SAT problem with this basis restriction is already  $\text{QMA}_1$ -complete, defined with respect to a natural gateset. Our construction is based on the Feynman-Kitaev circuit-to-Hamiltonian construction, with a modified clock encoding that interleaves two clocks in the standard and Hadamard bases. In light of the central role played by CSS codes and the uncertainty principle in the proof of the NLTS theorem of Anshu, Breuckmann, and Nirkhe (STOC '23), we hope that the CSS-like structure of our Hamiltonians will make them useful for progress towards a quantum PCP theorem.

**2012 ACM Subject Classification** Theory of computation  $\rightarrow$  Complexity classes; Theory of computation  $\rightarrow$  Problems, reductions and completeness; Theory of computation  $\rightarrow$  Quantum complexity theory

**Keywords and phrases** quantum complexity theory, Hamiltonian complexity, Quantum Merlin Arthur (QMA),  $\text{QMA}_1$ , quantum satisfiability problem

**Digital Object Identifier** 10.4230/LIPIcs.ITCS.2026.101

**Related Version** *Full Version:* <https://arxiv.org/abs/2509.24390>

**Funding** *Henry Ma:* HM was supported by the NSF Graduate Research Fellowship Program.

*Anand Natarajan:* AN was supported by NSF CAREER grant number 2339948.

**Acknowledgements** Part of this work was done while both authors were visiting the Simons Institute for the Theory of Computing as part of the 2025 Summer Cluster on Quantum Computing, and the Challenge Institute for Quantum Computation at UC Berkeley. We thank Chinmay Nirkhe for several helpful discussions.

## 1 Introduction

Local Hamiltonians play a central role in quantum complexity theory, tying the subject both to applications in condensed matter physics and quantum chemistry, and to the classical theory of NP-completeness, constraint satisfaction problems, and combinatorial optimization. Much of the complexity theory of local Hamiltonians has developed in analogy to the theory of NP-completeness, with the class QMA playing the role of NP, and Kitaev's result showing that the *local Hamiltonian problem* (estimating the ground energy of a local Hamiltonian) is QMA-complete playing the role of the Cook-Levin theorem (NP-completeness of 3SAT, and local constraint satisfaction problems more generally). Despite these analogies, there are considerable differences between the classical and quantum settings. One important difference relates to the issue of “perfect completeness.” In the case of classical constraint satisfaction problems, it is typically just as hard to solve the “SAT” problem of deciding whether all constraints are simultaneously satisfiable, as to solve the “decisional MAX-SAT” problem of deciding whether at least a certain fraction of the constraints are satisfiable: both problems are NP-complete. In contrast, the equivalent SAT and MAX-SAT problems for



© Henry Ma and Anand Natarajan;

licensed under Creative Commons License CC-BY 4.0

17th Innovations in Theoretical Computer Science Conference (ITCS 2026).

Editor: Shubhangi Saraf; Article No. 101; pp. 101:1–101:22

Leibniz International Proceedings in Informatics



LIPICs Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

local Hamiltonians appear to have different complexities: all of our known QMA-completeness results are MAX-SAT type results, while SAT-type problems appear to capture the complexity class  $\text{QMA}_1$  of QMA proof systems with *perfect completeness* (i.e. the verifier accepts good proofs with certainty). By definition,  $\text{QMA}_1 \subseteq \text{QMA}$ , and the two classes are believed to be qualitatively similar in power, but the exact relation between them is murky.

Due to the nature of our tools for showing computational hardness of local Hamiltonian problems, our understanding of the MAX-SAT setting is better than our understanding of the SAT setting: in the MAX-SAT world, we have dichotomy theorems [10] for 2-local Hamiltonians, and a wide variety of families of Hamiltonians are known to be QMA-complete, including very simple and natural models like the  $XZ$ -model [6]. In the SAT setting, while we know that QUANTUM-3-SAT (the SAT problem for 3-local Hamiltonians) is  $\text{QMA}_1$ -complete [13], and that various other cases of QUANTUM- $k$ -SAT are easier (e.g. for  $k = 2$  the problem is in P [7], and for so-called “stoquastic” Hamiltonians the problem is in MA [8]), we don’t have dichotomy theorems, nor do we have hardness results as clean as the  $XZ$ -model hardness mentioned above. This is because tools like *gadget reductions* used to reduce between different types of Hamiltonians do not preserve perfect completeness.

At the same time, the class  $\text{QMA}_1$  is scientifically of great interest as a testbed to better understand Hamiltonian complexity. Arguably the single most important open problem in Hamiltonian complexity is the quantum PCP conjecture: that approximating the ground energy of a local Hamiltonian even up to a constant fraction of the total norm of the Hamiltonian remains QMA-hard. A natural weakening of this conjecture would be to show  $\text{QMA}_1$ -hardness for this problem, and in fact, the biggest partial milestone we have towards a quantum PCP theorem comes from the SAT setting. Specifically, the NLTS theorem of Anshu, Breuckmann, and Nirkhe [5] shows the existence of local Hamiltonians whose low-energy states satisfy a certain structural property that we expect should hold for quantum PCP Hamiltonians. These Hamiltonians are parent Hamiltonians of quantum stabilizer error correcting codes, and thus are exactly satisfiable (all valid code states are 0-energy ground states) – and this characterization of the ground space is used explicitly in the proof of the NLTS theorem. However, for the same reason, these Hamiltonians contain no computational hardness. The most natural next step beyond the NLTS theorem would be to show that the NLTS structural property holds for a family of Hamiltonians where the local Hamiltonian problem is *computationally hard*. In order to achieve such a result, can we show a  $\text{QMA}_1$ -hardness result for a class of local Hamiltonians that is sufficiently structured to use the techniques of Anshu, Breuckmann and Nirkhe’s proof?

In this work, we focus on one structural property of the code Hamiltonians of [5]: every local term in the Hamiltonian is a projector and is diagonal in either the standard basis (the “Pauli  $Z$ -basis”), or the Hadamard basis (the “Pauli  $X$ -basis”). Stabilizer codes with this property are called *CSS codes*, and the CSS structure of the code is crucial to the proof of the NLTS theorem in two ways. Firstly, to prove the NLTS property of low-energy states, Anshu et al, following the approach of Eldar and Harrow [12], reduce the problem to proving that low-energy states of their Hamiltonian, when measured in either the standard or Hadamard bases, give rise to “well spread” probability distributions over the Boolean hypercube. They are able to prove this property in turn using a Heisenberg uncertainty principle relating these two bases, together with properties of the codes they consider. Secondly, these code properties in turn are formulated and proved by viewing CSS codes as consisting of a pair of *classical* error correcting codes, one in each basis. In fact, the view of CSS codes as chain complexes, which is at the foundation of the vast literature on constructing and analyzing such codes, relies on this two-basis structure.

Constraints	SAT	MAX-SAT
$Z$ -basis, linear	In P (by Gaussian elimination)	NP-hard (by MAX-CUT)
$Z$ -basis, general	NP-hard (Cook-Levin)	NP-hard (by $\leftarrow$ )
$Z$ and $X$ bases, linear	In P (by Gaussian elimination)	QMA-hard [6]
$Z$ and $X$ bases, general	QMA <sub>1</sub> -hard (this work)	QMA-hard (by $\uparrow$ )

The main result of this work is to show that the SAT problem for two-basis local Hamiltonians is QMA<sub>1</sub>-complete. More precisely, our main theorem is the following.

► **Theorem 1.1.** *Let XZ-QUANTUM-6-SAT be the following problem: given a local Hamiltonian  $H = \sum_i H_i$  on  $n$  qubits, where each local term  $H_i$  is a projector acting on 6 qubits and is diagonal in either the standard or Hadamard basis, determine whether  $\lambda_{\min}(H) = 0$  or  $\lambda_{\min}(H) \geq b(n)$  where  $b(n) = \Omega(1/\text{poly}(n))$ , promised that one of the two is the case. Then for an appropriately chosen function  $b(n)$ , it holds that XZ-QUANTUM-6-SAT is complete for the class QMA<sub>1</sub><sup>G<sub>2</sub></sup> of quantum Merlin-Arthur proof systems with perfect completeness where the verifier consists of a circuit made up of gates from the gate set  $\mathcal{G}_2$ .*

There is a technical subtlety in the theorem statement, arising from the dependence of QMA<sub>1</sub> on the choice of gate set for the verifier’s circuit. We show completeness with respect to the gate set  $\mathcal{G}_2$  as defined by Rudolph [18], consisting of NOT, controlled-NOT, and Toffoli gates ( $X, CX, CCX$ ) and the tensor product  $\hat{H} \otimes \hat{H}$  of two Hadamard gates. This gate set is a slight variant of the commonly used Hadamard-plus-Toffoli gate set [1], which is universal for quantum computing.

To better understand the class of Hamiltonians for which we show hardness, it is perhaps useful to draw a classical analogy. A classical *linear* code is associated with a collection of linear constraints over  $\mathbb{F}_2$ . By the Gaussian elimination algorithm, the SAT problem for linear constraints can always be solved in polynomial time, but once the constraints are allowed to be nonlinear, the complexity of the problem jumps up to NP. In the quantum world, the parent Hamiltonian of a CSS code consists of a pair of linear classical constraint satisfaction problems (CSPs), one for each basis: a valid code state is one that, when measured in the  $Z$ -basis, yields a satisfying string for the  $Z$ -basis linear constraints, and when measured in the  $X$ -basis, yields a satisfying string for the  $X$ -basis linear constraints. By standard stabilizer techniques, Gaussian elimination is sufficient to solve the SAT problem for these Hamiltonians in polynomial time. The Hamiltonians we consider also consist of a pair of classical constraint satisfaction problems in the two bases, except with nonlinear<sup>1</sup> constraints now being allowed, and our result shows that, as in the classical case, allowing for nonlinearity causes the complexity to jump from P to the maximal possible level of hardness (QMA<sub>1</sub> in this case). This classical analogy also helps illustrate the relation between our result and the MAX-SAT case. In the MAX-SAT case, classically, even two-local linear constraints become NP-hard (this is by the NP-hardness of the MAX-CUT problem), and similarly, in the quantum case, the result of Biamonte and Love for the XZ-model shows that the MAX-SAT problem for “linear” constraints is QMA-hard.

<sup>1</sup> To be clear, the constraints are nonlinear in their action on the binary string measurement outcomes viewed as elements of  $\mathbb{F}_2^n$  – the Hamiltonian is still a linear operator over the Hilbert space.

### The two basis paradigm

Viewing matters more subjectively and at a higher level, it is a striking fact that many interesting phenomena in quantum computing rest on the interplay between the standard and Hadamard bases. Examples of this “two basis paradigm” in action include the BB84 protocol and Wiesner’s quantum money scheme, Simon’s algorithm, the magic square game (and, arguably, the proof that  $\text{MIP}^* = \text{RE}$ ), the forrelation problem (some versions of which are BQP-complete), Aaronson and Christiano’s subspace quantum money scheme, and Mahadev’s measurement protocol, besides the previously mentioned examples of CSS codes and Biamonte and Love’s QMA-hardness results. We see our result as another instance of this paradigm.

### 1.1 Technical overview

Our proof of Theorem 1.1 consists of two parts. To show  $\text{QMA}_1^{\mathcal{G}_2}$ -completeness of our problem, we must show that it is both contained in  $\text{QMA}_1^{\mathcal{G}_2}$  and that it is  $\text{QMA}_1^{\mathcal{G}_2}$ -hard. The containment follows along standard lines: we construct a verifier for XZ-QUANTUM-6-SAT instances that samples a random term from the Hamiltonian and coherently measures it on the witness state. The only nontrivial step in this construction is showing that the verifier’s gate set  $\mathcal{G}_2$  can exactly simulate a controlled Hadamard gate, since this gate is needed to perform the measurement of the witness in the appropriate basis.

The bulk of the proof is concerned with the  $\text{QMA}_1^{\mathcal{G}_2}$ -hardness. We build on the Kitaev circuit-to-Hamiltonian construction [15], which is a quantum analog of the fundamental construction used by Cook and Levin [9, 16] to show that SAT is NP-complete. Let us consider a circuit consisting of gates  $U_1, U_2, \dots, U_T$ , acting on a Hilbert space  $\mathcal{H}$ . Kitaev’s construction associates this circuit to a Hamiltonian  $H$  acting on a space  $\mathcal{H} \otimes \mathcal{H}_{\text{clock}}$ , where  $\mathcal{H}_{\text{clock}}$  is the Hilbert space of an ancillary “clock” register. The space  $\mathcal{H}_{\text{clock}}$  contains at least  $T + 1$  orthonormal states  $|\hat{0}\rangle, |\hat{1}\rangle, \dots, |\hat{T}\rangle$ , but typically has a much larger dimension –  $2^T$  in Kitaev’s original construction. The full Hamiltonian  $H$  can be written as a sum of four components:

$$H = H_{\text{prop}} + H_{\text{format}} + H_{\text{in}} + H_{\text{out}},$$

where each component is a sum of local projectors. Moreover, Kitaev’s analysis shows that, when applied to a  $\text{QMA}_1$  circuit  $U_1, \dots, U_T$  that accepts some witness state with certainty, the Hamiltonian  $H$  has a 0-energy ground state, and otherwise, if applied to a circuit that rejects all witnesses with high probability, then the minimum energy of  $H$  is at least  $1/\text{poly}(n)$ .

Our approach is to use Kitaev’s construction, but modify the terms  $H_{\text{prop}}$  and  $H_{\text{format}}$ , as well as the encoding of the clock states  $|\hat{0}\rangle, \dots, |\hat{T}\rangle$ , so that the resulting Hamiltonian  $H$  satisfies our two-basis constraint. Since we will not modify them significantly, for now let us ignore the terms  $H_{\text{in}}$  and  $H_{\text{out}}$ , and study the ground space of the remaining two terms. These are designed in Kitaev’s construction so that they *always* have a nonempty 0-energy ground space, which consists of *history states* of the form

$$|\psi_{\text{history}}\rangle = \frac{1}{\sqrt{T+1}} \sum_{t=0}^T U_t \dots U_1 |\psi_0\rangle \otimes |\hat{t}\rangle.$$

This is achieved in two ways:

- The “format” Hamiltonian  $H_{\text{format}}$  forces the state to be supported only on the “good” clock register states  $|\hat{0}\rangle, \dots, |\hat{T}\rangle$ . This is done in Kitaev’s construction by taking  $\mathcal{H}_{\text{clock}}$  to be a space of  $T$  qubits, and taking the good clock states to be encodings of  $0, \dots, T$  in unary in the standard basis, so that

$$|\hat{t}\rangle = |1\rangle_1 \dots |1\rangle_{t-1} |1\rangle_t |0\rangle_{t+1} \dots |0\rangle_T.$$

To force the state to be supported only on these states,  $H_{\text{format}}$  consist of projectors that enforce the constraint on each pair of adjacent qubits  $(t, t+1)$  that a 0 can never be followed by a 1. These constraints are already diagonal in the  $Z$ -basis.

- The “propagation” Hamiltonian  $H_{\text{prop}}$  consists of a sum of  $T$  constraints, each of which forces the  $|\widehat{t-1}\rangle$  and  $|\hat{t}\rangle$  components of the state to be related by an application of the unitary  $U_t$ . In Kitaev’s construction, these terms take the form

$$\begin{aligned} H_{\text{prop},t} = & \frac{1}{2} I \otimes (|110\rangle\langle 110| + |100\rangle\langle 100|)_{t-1,t,t+1} \\ & - \frac{1}{2} U_t \otimes (|110\rangle\langle 100|)_{t-1,t,t+1} - \frac{1}{2} U_t^\dagger \otimes (|100\rangle\langle 110|)_{t-1,t,t+1}, \end{aligned}$$

where the second tensor factor acts on the specified qubits of the clock register. It can be checked that for general choices of  $U_t$ , these terms are not diagonal in either the  $X$ -basis or the  $Z$ -basis.

To make progress, we must modify the propagation terms. To do this, we make the following key observation: the terms  $H_{\text{prop},t}$  are “almost” diagonal in the  $X$ -basis whenever the unitary  $U_t$  itself is diagonal in the  $X$ -basis and is Hermitian (i.e.  $U_t = U_t^\dagger$ ). Specifically, in this case,  $H_{\text{prop},t}$  can be written as

$$H_{\text{prop},t} = \frac{1}{2} (I \otimes I_t - U_t \otimes X_t) \otimes (|10\rangle\langle 10|)_{t-1,t+1},$$

which is diagonal when all qubits are placed in the  $X$ -basis except for qubits  $t-1, t+1$  of the clock register, which remain in the  $Z$ -basis. Moreover, the same property would hold, with  $X$  and  $Z$  interchanged, if the good clock states were  $X$ -basis states, and  $U_t$  were Hermitian and diagonal in the  $Z$ -basis.

This suggests the following modifications:

- Encode good clock states by putting alternating qubits in the  $X$ - and  $Z$ -bases. This means that valid clock states would look like

$$|\hat{0}\rangle = |0+0+0+\dots\rangle, |\hat{1}\rangle = |1+0+0+\dots\rangle, |\hat{2}\rangle = |1-0+0+\dots\rangle, |\hat{3}\rangle = |1-1+0+\dots\rangle, \dots$$

- Choose a universal gate set where each gate is Hermitian and diagonal in either the  $X$ - or  $Z$ -basis. We construct such a gate set and show that it is computationally equivalent to  $\mathcal{G}_2$ .
- Padding the circuit so that each  $X$ -basis gate falls on an odd timestep and each  $Z$ -basis timestep falls on an even timestep. This makes the propagation terms fully diagonal in the  $X$ -basis for odd times, and the  $Z$ -basis for even times: e.g. for odd  $t$ ,

$$H_{\text{prop},t} = \frac{1}{2} (I \otimes I_t - U_t \otimes X_t) \otimes |-\rangle\langle -|_{t-1,t+1}.$$

With these changes, every term in  $H_{\text{prop},t}$  is diagonal in either the  $X$ - or  $Z$ -basis. However, now that we have changed the encoding of the clock, we must also change  $H_{\text{format}}$ , and it is not hard to see that any Hamiltonian that restricts us only to valid clock states will not be diagonal in either basis. The solution to this is somewhat surprising.

## 101:6 Two Bases Suffice for $\text{QMA}_1$ -Completeness

- We modify  $H_{\text{format}}$  to include checks that only check consistency *separately* on the  $X$ - and  $Z$ -basis parts of the clock state. Specifically, our new  $H_{\text{format}}$  checks that the clock register in a state  $|t_Z, t_X\rangle$  consisting of an “interleaving” of some valid unary-encoding of an integer  $t_Z$  in the  $Z$ -basis in odd positions, and an integer  $t_X$  in the  $X$ -basis in even positions. It does *not* check that these two interleaved times are synchronized with each other, meaning that this Hamiltonian accepts invalid states such as

$$|0-0-0+\dots 0+0\rangle,$$

which do not correspond to valid clock states. We call these states “fake” states.

- Surprisingly, we show that, in fact,  $H_{\text{prop}} + H_{\text{format}}$  *together* force the ground state to be supported only on valid clock states. This is because every fake clock state is coupled by some term in  $H_{\text{prop}}$  to a “bad” clock state, which incurs an energy penalty from  $H_{\text{format}}$ . We analyze this quantitatively by relating the Hamiltonian to a graph Laplacian.

The majority of the technical work in our analysis consists in (a) proving universality of our new gate set, and (b) quantitatively bounding the minimum energy of  $H_{\text{format}} + H_{\text{prop}}$  on the subspace of invalid clock strings.

### 1.2 Open questions

We see several possibilities for future work with a quantum PCP flavor.

1. Our construction, since it is based on the Feynman-Kitaev clock, cannot achieve the NLTS property: indeed, the product state  $|0\rangle \otimes |\hat{0}\rangle$  violates only a single propagation term of the Hamiltonian. To get to NLTS while maintaining the two-basis structure, can we apply our ideas to the tensor-network-based circuit-to-Hamiltonian construction of [4]?
2. What do random instances of our two-basis Hamiltonians look like? Can we find a distribution such that Hamiltonians sampled from this distribution have a zero-energy ground state with high probability, but where finding the ground state is computationally hard? It would be especially interesting to relate this to work on the combinatorial NLTS property for tensor network Hamiltonians constructed from random SAT instances [3].
3. One candidate approach to proving the quantum PCP conjecture is to design a “locality-preserving gap amplification” procedure, that operates on instances of local Hamiltonians by increasing the minimum energy of unsatisfiable Hamiltonians (amplifying the “promise gap”), while preserving the locality of the Hamiltonian. To build towards such a procedure, can we apply classical “gap-amplification” or “locality-reduction” transformations separately to one of the two bases and make some kind of meaningful progress towards gap amplification? Moreover, is there an analog of “distance balancing” for CSS codes [20], whereby a code with high distance in one basis but low distance in the other can be converted into a code with decent distance in both bases?
4. In this paper we have taken the point of view that  $\text{QMA}_1^{\mathcal{G}_2}$  is likely to be qualitatively similar to  $\text{QMA}$  in power. But if  $\text{QMA}_1^{\mathcal{G}_2}$  is in fact much weaker, could our completeness result give us a route to showing this by putting the problem XZ-QUANTUM-6-SAT into a smaller class like QCMA? One very speculative route to doing this is to give an efficient classical description of ground states of such Hamiltonians, perhaps using tools from additive combinatorics, since Hamiltonian terms in the  $X$ -basis can be viewed as additive constraints on the support of the ground state in the standard basis.

## 2 Preliminaries

### 2.1 Quantum computation

We briefly set up the basic formalism of quantum computation. For a more detailed exposition, we refer the reader to [17].

An  $n$ -qubit *quantum state*  $|\psi\rangle$  is a unit vector in a complex Hilbert space  $(\mathbb{C}^2)^{\otimes n}$ . A  $k$ -qubit *quantum operation* is a unitary operator  $U$  acting on a  $k$ -qubit space  $(\mathbb{C}^2)^{\otimes k}$ . We follow the conventions of Dirac notation. We use  $U^\dagger$  to denote the adjoint of an operator. Let  $[n] = \{1, \dots, n\}$ . We can extend  $U$  to an operation  $U \otimes I_{[n]\setminus S}$  on  $n$ -qubit space, where  $S \subseteq [n]$  is the set of  $k$  qubits which  $U$  acts on, and  $I_{[n]\setminus S}$  is the identity operator on the remaining qubits.

We now fix some notation for relevant states and operations. As usual,  $|0\rangle$  and  $|1\rangle$  refer to the single-qubit standard basis states, while  $|+\rangle$  and  $|-\rangle$  are the Hadamard basis states. We will often omit tensor products, e.g.  $|00\rangle = |0\rangle \otimes |0\rangle$ . Let  $X$  and  $Z$  be the corresponding single-qubit Pauli operations. We will also refer to the standard and Hadamard bases as the  $Z$  basis and  $X$  basis respectively. Let  $\hat{H}$  be the Hadamard gate (we put the hat to avoid confusion with a Hamiltonian  $H$ ). Let  $CX$ ,  $CZ$ , and  $\text{SWAP}$  denote the two-qubit CNOT, controlled- $Z$  and swap operations respectively. Let  $CCX$  denote the Toffoli gate and  $CCZ$  the controlled-controlled- $Z$  gate.

A *quantum circuit* on  $n$  qubits is an  $n$ -qubit operation  $U$  which can be decomposed into a sequence of gates  $U = G_m \cdots G_1$ , where each  $G_i$  is a quantum operation from some fixed *gate set*; typically, this gate set contains operations which each acts nontrivially on just a small number of qubits. In our protocols, we will consider projective measurements of the first qubit in the  $Z$  basis, which is given by the orthogonal projectors  $\{|0\rangle\langle 0|_1 \otimes I, |1\rangle\langle 1|_1 \otimes I\}$ , where the subscript indicates the first qubit register, and the identity operation acts on the remaining qubits.

### 2.2 Quantum complexity theory

We now define the model of quantum verification of interest in this work. A *promise problem* is a pair  $(L_{\text{yes}}, L_{\text{no}})$  of subsets of bitstrings  $L_{\text{yes}}, L_{\text{no}} \subseteq \{0, 1\}^*$  which satisfies  $L_{\text{yes}} \cap L_{\text{no}} = \emptyset$ . Promise problems generalize languages, a notion from classical complexity theory: a language is just a promise problem which additionally satisfies  $L_{\text{yes}} \cup L_{\text{no}} = \{0, 1\}^*$ .

A *quantum verifier* is a uniform family of quantum circuits  $\{V_x\}$  whose goal is to determine whether a given string  $x$  is in  $L_{\text{yes}}$  or  $L_{\text{no}}$ , promised that one is the case. On an input  $x$  of length  $n$ , the verifier is given access to a quantum proof state  $|\psi\rangle$  and a register of ancilla qubits all initialized to  $|0\rangle$ . The circuit is efficient, in the sense that both the proof and ancilla registers have  $\text{poly}(n)$  qubits, and there are  $\text{poly}(n)$  gates in the circuit.

After the circuit is applied, the final step of the verification is to measure the first qubit in the  $Z$  basis. We say the verifier accepts if the measurement outcome is 1 and rejects if the outcome is 0. The probability that the verifier  $V_x$  accepts is then

$$\|(|1\rangle\langle 1|_1 \otimes I)V_x|\psi\rangle|0 \cdots 0\rangle_{\text{anc}}\|^2. \quad (1)$$

We now define  $\text{QMA}_1$  and  $\text{QMA}$ ; we will mainly be interested in  $\text{QMA}_1$  in this work.

► **Definition 2.1.** A promise problem  $(L_{\text{yes}}, L_{\text{no}})$  is in  $\text{QMA}_1$  if there is a uniform family of efficient quantum circuits  $\{V_x\}$  such that for any  $x \in \{0, 1\}^n$ ,

1. If  $x \in L_{\text{yes}}$ , then there is some proof state for which  $V_x$  accepts with probability 1.
2. If  $x \in L_{\text{no}}$ , then for any proof state,  $V_x$  accepts with probability  $\leq 1/3$ .

## 101:8 Two Bases Suffice for QMA<sub>1</sub>-Completeness

We refer to the first condition as *completeness* and the second condition as *soundness*. Importantly, the definition of QMA<sub>1</sub> is not known to be independent of the choice of gate set for the circuits. In this work, we choose the gate set

$$\mathcal{G}_2 = \{X, CX, CCX, \hat{H} \otimes \hat{H}\}$$

for QMA<sub>1</sub>. When needing to explicitly refer to QMA<sub>1</sub> with this gate set, we use the notation QMA<sub>1</sub> <sup>$\mathcal{G}_2$</sup> . We now make some observations about the chosen gate set.

► **Remark 2.2.** First,  $\mathcal{G}_2$  is computationally universal, i.e. any quantum circuit can be simulated by a circuit which only uses gates from  $\mathcal{G}_2$ . This is because gates in  $\mathcal{G}_2$  allow us to implement  $CCX$  and  $\hat{H}$  (which can be implemented by applying  $\hat{H} \otimes \hat{H}$  on the desired qubit and an otherwise unused ancilla qubit); these two gates already form a computationally universal gate set [19, 1].

Second,  $\mathcal{G}_2$  is studied in a work of Rudolph [18] which makes progress on the interesting open problem of whether QMA<sub>1</sub> has a universal gate set. They show that the problem of Gapped Clique Homology on weighted graphs (introduced in [14]) is QMA<sub>1</sub> <sup>$\mathcal{G}_2$</sup> -complete. This result gives a surprising connection between an important problem in computational topology and QMA<sub>1</sub> <sup>$\mathcal{G}_2$</sup> , motivating our study of other properties of QMA<sub>1</sub> <sup>$\mathcal{G}_2$</sup> .

QMA is defined in the same way as QMA<sub>1</sub>, except the completeness parameter is  $2/3$ , i.e. for  $x \in L_{\text{yes}}$ , the verifier must accept with probability at least  $2/3$ . In contrast to QMA<sub>1</sub>, the definition of QMA is independent of the choice of gate set, since the Solovay-Kitaev theorem [11] shows that any quantum gate can be efficiently approximated using gates from a universal gate set. This argument does not straightforwardly extend to QMA<sub>1</sub>, since approximating a gate may not preserve the perfect completeness required of a QMA<sub>1</sub> protocol.

The classical theory of NP-completeness generalizes to QMA<sub>1</sub> and QMA in a straightforward way. A promise problem  $K = (K_{\text{yes}}, K_{\text{no}})$  has an *efficient reduction* to a promise problem  $L = (L_{\text{yes}}, L_{\text{no}})$  if there is an efficient deterministic algorithm  $A$  which maps a bitstring  $x$  to a bitstring  $A(x)$  such that  $A(x) \in L_{\text{yes}}$  if  $x \in K_{\text{yes}}$ , and  $A(x) \in L_{\text{no}}$  if  $x \in K_{\text{no}}$ . A promise problem  $L$  is QMA<sub>1</sub>-hard if for every  $K \in \text{QMA}_1$ , there is an efficient reduction from  $K$  to  $L$ . A promise problem  $L$  is QMA<sub>1</sub>-complete if  $L$  is in QMA<sub>1</sub> and  $L$  is QMA<sub>1</sub>-hard. The definitions for QMA are analogous.

### 2.3 Hamiltonian complexity

In this section, we define the QUANTUM- $k$ -SAT problem and introduce its basis-restricted variant. An  $n$ -qubit *Hamiltonian* is a Hermitian operator acting on  $n$  qubits. A  $k$ -local operator on  $n$  qubits is an operator which acts nontrivially on at most  $k$  of the qubits. The *energy* of a state  $|\psi\rangle$  with respect to Hamiltonian  $H$  is  $\langle\psi|H|\psi\rangle$ . Note that this quantity is always real and non-negative, since  $H$  is Hermitian.

► **Definition 2.3.** Let  $k \geq 1$  and let  $\mathcal{S}$  be a set of Hermitian  $k$ -local projectors. The problem QUANTUM- $k$ -SAT is a promise problem whose input is a classical description of a Hamiltonian  $H = \sum_{i=1}^m h_i$ , where each term  $h_i$  acts on  $n$  qubits and belongs to  $\mathcal{S}$ .  $H$  is a “yes” instance if there is an  $n$ -qubit state  $|\psi\rangle$  such that  $\langle\psi|H|\psi\rangle = 0$ .  $H$  is a “no” instance if for all  $n$ -qubit states  $|\psi\rangle$ ,  $\langle\psi|H|\psi\rangle \geq 1/\text{poly}(n)$ .

The NP-complete problem  $k$ -SAT can be viewed as a special case of QUANTUM- $k$ -SAT with a highly restricted Hamiltonian: each projector term is diagonal in the same, predetermined basis (namely, the  $Z$  basis) [2]. Let us refer to this as a *classical Hamiltonian*. This brings

into view one of the guiding questions of this work: if we relax the restrictions placed on a classical Hamiltonian, at what point does the Hamiltonian become “quantum”? This question has been previously studied for the relaxation in which the Hamiltonian is no longer required to be classical, but instead the Hamiltonian terms must pairwise commute. In this work, we introduce a new way of generalizing classical Hamiltonians which we call *basis restriction*.

► **Definition 2.4.** For  $n \geq 1$ , let  $\mathcal{B}_n$  be a set of bases of  $n$ -qubit space. Let  $\mathcal{B} = \cup_{n \geq 1} \mathcal{B}_n$ . Define  $\mathcal{B}$ -QUANTUM- $k$ -SAT to be the problem QUANTUM- $k$ -SAT where we choose  $\mathcal{S}$  (the set of allowed projectors) to be the set of Hermitian  $k$ -local projectors which are diagonal in some basis in  $\mathcal{B}$ .

As an example, if for all  $n$  we let  $\mathcal{B}_n$  include just the  $Z$  basis on  $n$  qubits, then an instance of  $\mathcal{B}$ -QUANTUM- $k$ -SAT is a Hamiltonian  $H = \sum_i h_i$  for which each  $h_i$  is diagonal in the  $Z$  basis. Then  $H$  is a classical Hamiltonian and  $\mathcal{B}$ -QUANTUM- $k$ -SAT is in NP.

We are interested in the complexity of basis-restricted QUANTUM- $k$ -SAT when we allow for more than one “type” of basis. In particular, is there some setting in which the problem already becomes QMA-hard, even though the number of basis types is small? We answer this question in the following sections by considering the problem XZ-QUANTUM- $k$ -SAT, which we define as basis-restricted QUANTUM- $k$ -SAT problem where each  $h_i$  is diagonal in either the  $Z$  or the  $X$  basis.

### 3 XZ-Quantum-6-Sat is in QMA<sub>1</sub>

In this section, we begin the proof of Theorem 1.1 by showing that XZ-QUANTUM-6-SAT is in QMA<sup>G<sub>2</sub></sup>. We start with a useful property of the gate set  $\mathcal{G}_2$ . For gate  $U$ , let  $\Gamma(U)$  denote the controlled- $U$  gate.

► **Lemma 3.1.** *For every  $U \in \mathcal{G}_2$ ,  $\Gamma(U)$  can be exactly implemented using gates in  $\mathcal{G}_2$  (using an ancilla qubit, which can be in any state and will be left unchanged after the simulation).*

**Proof.**  $\Gamma(X)$  and  $\Gamma(CX)$  are already included in  $\mathcal{G}_2$ . Let  $a$  denote an ancilla qubit.  $\Gamma(CCX)$  on control qubits  $i, j, k$  and target qubit  $l$  can be decomposed as

$$\Gamma(CCX)_{i,j,k,l} \otimes I_a = CCX_{i,j,a} CCX_{k,a,l} CCX_{i,j,a} CCX_{k,a,l}.$$

For  $\Gamma(\hat{H} \otimes \hat{H})$  with control qubit  $i$  and target qubits  $j, k$ , we first give a decomposition using the  $\Gamma(\text{SWAP})$  gate.

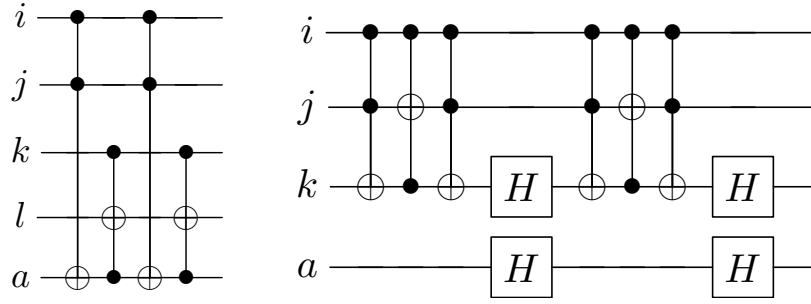
$$\Gamma(\hat{H} \otimes \hat{H})_{i,j,k} \otimes I_a = \Gamma(\text{SWAP})_{i,j,k} (\hat{H} \otimes \hat{H})_{k,a} \Gamma(\text{SWAP})_{i,j,k} (\hat{H} \otimes \hat{H})_{k,a}.$$

Then, since  $\Gamma(AB) = \Gamma(A)\Gamma(B)$  for any gates  $A$  and  $B$ , we may obtain a decomposition of  $\Gamma(\text{SWAP})$  in terms of  $CCX$ 's by applying the identity  $\text{SWAP}_{i,j} = CX_{i,j} CX_{j,i} CX_{i,j}$ . This completes the proof. We illustrate these constructions in Figure 1. ◀

► **Theorem 3.2.** *XZ-QUANTUM-6-SAT is in QMA<sub>1</sub><sup>G<sub>2</sub></sup>.*

**Proof.** First, without loss of generality, we can build a QMA<sub>1</sub><sup>G<sub>2</sub></sup> verification circuit which allows intermediate  $Z$  basis measurements and quantum operations conditioned on the measurement results. This is because a conditionally applied gate  $U$  can be replaced by the controlled operation  $\Gamma(U)$ , removing the intermediate measurement. Lemma 3.1 then says that  $\Gamma(U)$  can then be decomposed back into the gate set  $\mathcal{G}$ . Thus, a circuit with intermediate

## 101:10 Two Bases Suffice for QMA<sub>1</sub>-Completeness



■ **Figure 1** Decomposition of  $\Gamma(CCX)$  and  $\Gamma(\hat{H} \otimes \hat{H})$  into gates from  $\mathcal{G}_2$ .

measurements can be rewritten as a circuit with just a single measurement at the end of the computation, as in our definition of QMA<sub>1</sub>. Also, classical processing can be performed freely, since our gate set includes  $CCX$ , which is universal for classical computation.

We use the verification procedure of Gosset and Nagaj [13], which is correct as long as the following condition holds:

- (\*) There is an efficient quantum algorithm using gates in  $\mathcal{G}$  which, given a projector  $\Pi \in \mathcal{S}$ , exactly performs the projective measurement  $\{I - \Pi, \Pi\}$  on a given state  $|\psi\rangle$ .

Recall that  $\mathcal{S}$  is the set of Hermitian 6-local projectors which are diagonal in either the  $Z$  basis or the  $X$  basis.

We briefly sketch how the QMA<sub>1</sub> protocol works assuming (\*) holds. First, choose a random projector term  $\Pi$  in the Hamiltonian  $H$  (random bits can be obtained by applying  $X$  to a  $|0\rangle$  ancilla then measuring). Then, perform the projective measurement in (\*) on the proof state  $|\psi\rangle$ . Accept when the outcome is  $I - \Pi$ , and reject otherwise. The probability of accepting is  $1 - \langle \psi | \Pi | \psi \rangle$ . When  $H$  has a zero-energy state, the protocol accepts this state with probability 1, showing perfect completeness. If instead  $\langle \psi | H | \psi \rangle \geq 1$  for all  $|\psi\rangle$ , the protocol will reject with probability at least  $1/\text{poly}(n)$ , which can be amplified to give the desired soundness; the full analysis is given in [13].

We now prove that (\*) holds. A projector  $\Pi \in \mathcal{S}$  can be specified by a set of six qubits  $Q$  on which it acts nontrivially, a change of basis matrix  $V$  (which is either  $I$  or a tensor product of  $\hat{H}$ 's), and a set of supported strings  $S \subseteq \{0, 1\}^6$ , such that

$$\Pi = \sum_{z \in S} V|z\rangle\langle z|V.$$

The algorithm first applies  $V$  on  $|\psi\rangle$ , by applying  $\hat{H}$  on each of the specified qubits (using the  $\hat{H} \otimes \hat{H}$  gate, with the second  $\hat{H}$  acting on an otherwise unused ancilla qubit). It then measures the qubits in  $Q$  in the  $Z$  basis. Let  $z^*$  be the measurement outcome. The algorithm returns outcome  $\Pi$  if  $z^* \in S$ , and otherwise returns outcome  $I - \Pi$ . The probability of getting outcome  $\Pi$  is

$$\sum_{z \in S} \|\langle z | \psi \rangle V | \psi \rangle\|^2 = \langle \psi | \Pi | \psi \rangle,$$

as desired. This completes the proof. ◀

## 4 Circuit-to-XZ-Hamiltonian reduction

In this section, we discuss some preliminaries for the  $\text{QMA}_1$ -hardness proof of Section 5.

### 4.1 Two gate sets for $\text{QMA}_1$

Recall that our gate set for  $\text{QMA}_1$  is  $\mathcal{G}_2 = \{X, CX, CCX, \hat{H} \otimes \hat{H}\}$ . Define a new gate set

$$\mathcal{G}_{XZ} = \{X, CZ, CCZ, \hat{G}\},$$

where  $\hat{G}$  is defined as the two-qubit operation equivalent to

$$\hat{G} = (\hat{H} \otimes \hat{H})CZ(\hat{H} \otimes \hat{H}).$$

We show that  $\mathcal{G}_{XZ}$  and  $\mathcal{G}_2$  are interchangeable gate sets, in the following sense:

► **Lemma 4.1.**  $\text{QMA}_1^{\mathcal{G}_2} = \text{QMA}_1^{\mathcal{G}_{XZ}}$ .

**Proof.** We start by showing that  $\text{QMA}_1^{\mathcal{G}_{XZ}} \subseteq \text{QMA}_1^{\mathcal{G}_2}$ . It suffices to show that each gate in  $\mathcal{G}_{XZ}$  can be exactly written as a (constant length) sequence of gates from  $\mathcal{G}_2$ . Such a simulation can freely use the  $\text{QMA}_1$  verifier's ancilla register. In the sequences we construct, the ancilla qubits can be in any state and will be left unchanged after the gate simulation. Also note that *exact* simulation is critical: merely approximating a gate using another gate set would not necessarily preserve the  $\text{QMA}_1$  protocol's perfect completeness.

To set some notation, for a gate  $U$  and a set of gates  $S$ , we say that  $U \in \overline{S}$  if  $U$  can be exactly written as a sequence of gates from  $S$ . The following closure property is immediate: if  $S' \subseteq \overline{S}$  and  $U \in \overline{S'}$ , then  $U \in \overline{S}$ . In particular, if  $V \in \overline{S}$  and  $U \in \overline{S \cup \{V\}}$ , then  $U \in \overline{S}$ .

We now consider the gates in  $\mathcal{G}_{XZ}$ . The  $X$  gate is already in  $\mathcal{G}_2$ . We can simulate the  $CZ$  gate on qubits  $i, j$  with an ancilla qubit  $a$ , using the circuit identity

$$CZ_{i,j} \otimes I_a = (\hat{H} \otimes \hat{H})_{j,a} CX_{i,j} (\hat{H} \otimes \hat{H})_{j,a}.$$

This puts  $CZ \in \overline{\mathcal{G}_2}$ . A similar identity

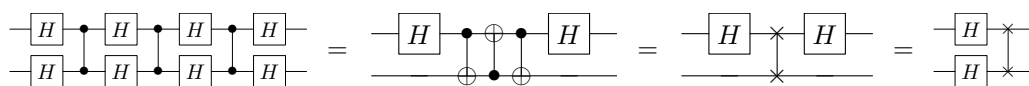
$$CCZ_{i,j,k} \otimes I_a = (\hat{H} \otimes \hat{H})_{k,a} CCX_{i,j,k} (\hat{H} \otimes \hat{H})_{k,a}$$

puts  $CCZ \in \overline{\mathcal{G}_2}$ . Finally,  $\hat{G} \in \overline{\mathcal{G}_2 \cup \{CZ\}}$  by definition, so by the closure property,  $\hat{G} \in \overline{\mathcal{G}_2}$ .

We now show that  $\text{QMA}_1^{\mathcal{G}_2} \subseteq \text{QMA}_1^{\mathcal{G}_{XZ}}$  using the same approach. We introduce a useful intermediate gate: define  $\hat{F}$  to be the two-qubit operation equivalent to

$$\hat{F} = \text{SWAP}(\hat{H} \otimes \hat{H}).$$

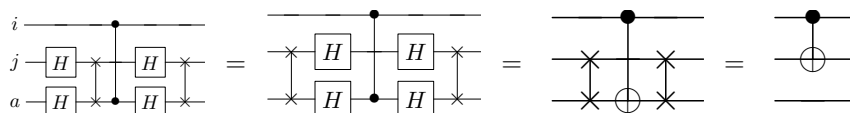
We observe that  $\hat{F} = \hat{G} \cdot CZ \cdot \hat{G}$  using the circuit



where we used that  $\hat{H}_j CZ_{i,j} \hat{H}_j = CX_{i,j}$  and  $\text{SWAP}_{i,j} = CX_{i,j} CX_{j,i} CX_{i,j}$ . This puts  $\hat{F} \in \overline{\mathcal{G}_{XZ}}$ . Next, using that  $\hat{H} \otimes \hat{H}$  and SWAP commute, we have that

$$CX_{i,j} \otimes I_a = \hat{F}_{j,a} CZ_{i,a} \hat{F}_{j,a},$$

since



## 101:12 Two Bases Suffice for QMA<sub>1</sub>-Completeness

Thus  $CX \in \overline{\mathcal{G}_{XZ} \cup \{\hat{F}\}}$ , implying that  $CX \in \overline{\mathcal{G}_{XZ}}$  by the closure property. The identity

$$CCX_{i,j,k} \otimes I_a = \hat{F}_{k,a} CCZ_{i,j,a} \hat{F}_{k,a}$$

similarly puts  $CCX \in \overline{\mathcal{G}_{XZ}}$ . Finally, note that  $\hat{H} \otimes \hat{H} = \text{SWAP} \cdot \hat{F}$ . Writing SWAP as three  $CX$ 's, we have that  $\hat{H} \otimes \hat{H} \in \overline{\{\hat{F}, CX\}}$ , so  $\hat{H} \otimes \hat{H} \in \overline{\mathcal{G}_{XZ}}$  by the closure property. ◀

We find it useful to change to the gate set  $\mathcal{G}_{XZ}$  since each gate in  $\mathcal{G}_{XZ}$  has the following nice properties, which are straightforward to check:

► **Lemma 4.2.** *Each gate in  $\mathcal{G}_{XZ}$  is Hermitian, 3-local, and diagonal in either the  $Z$  basis or the  $X$  basis.*

These properties are used in the proof that XZ-QUANTUM-6-SAT is QMA<sub>1</sub>-hard, when we show that the Hamiltonian term  $H_{\text{prop}}$  is an instance of XZ-QUANTUM-6-SAT (Lemma 4.4).

### 4.2 The Kitaev circuit-to-Hamiltonian construction

We now sketch the original Kitaev construction and note which parts of the construction carry through to the proof of Theorem 1.1 unchanged. Let  $(L_{\text{yes}}, L_{\text{no}})$  be a promise problem in QMA<sub>1</sub>, and let  $V_x$  be the corresponding circuit which verifies a length  $n$  bitstring  $x$ , given a proof state  $|\psi\rangle$ . Recall that the verification starts from state  $|\text{init}\rangle = |\psi\rangle|0 \cdots 0\rangle_{\text{anc}}$ . Let  $\mathcal{H}$  denote the Hilbert space in which the initial state lives. Let  $Q_{\text{proof}}$  and  $Q_{\text{anc}}$  denote the indices of the qubits in the proof and ancilla registers. Write the circuit as  $V_x = U_T \cdots U_1$ , where each  $U_i$  is a gate from the chosen gate set for QMA<sub>1</sub>.

The Kitaev construction is an efficient mapping from the circuit  $V_x$  to a Hamiltonian

$$H = H_{\text{in}} + H_{\text{out}} + H_{\text{prop}} + H_{\text{format}}.$$

$H$  acts on Hilbert space  $\mathcal{H} \otimes \mathcal{H}_{\text{clock}}$ , where  $\mathcal{H}_{\text{clock}}$  is an additional  $T$ -qubit *clock register*. The *Kitaev clock states*  $|\hat{t}\rangle \in \mathcal{H}_{\text{clock}}$ ,  $t \in \{0, \dots, T\}$  are defined as  $|\hat{t}\rangle = |\mathbf{1}^t \mathbf{0}^{T-t}\rangle$  (the definition of clock states will differ in our construction). This reduction has the following properties:

**Completeness** If  $x \in L_{\text{yes}}$ , then the *history state*

$$|\text{hist}\rangle = \frac{1}{\sqrt{T+1}} \sum_{t=0}^T U_t \cdots U_1 |\text{init}\rangle \otimes |\hat{t}\rangle \quad (2)$$

is a zero-energy state of  $H$ .

**Soundness** If  $x \in L_{\text{no}}$ , then any state  $|\varphi\rangle$  has at least  $\frac{1}{\text{poly}(n)}$  energy with respect to  $H$ .

**Instance of Quantum-Sat**  $H$  is a sum of projectors. Moreover, if each gate in the chosen gate is  $k$ -local, then  $H$  is  $(k+3)$ -local.

Each term in the Hamiltonian represents a constraint on the proof state (which is claimed to be the history state); an energy penalty is given if the constraint is not met.

### 4.3 $H_{\text{prop}}$ and $H_{\text{format}}$

Our construction differs from the Kitaev construction in the definition of  $H_{\text{prop}}$  and  $H_{\text{format}}$ , which we define in this section. First, we make a simple observation that the verification circuit  $V_x$  can be assumed to be in a standard form.

► **Lemma 4.3.** *Without loss of generality,  $V_x = U_T \cdots U_1$  (where  $U_i \in \mathcal{G}_{XZ}$ ) satisfies the following properties: (i)  $T$  is odd, (ii)  $U_i$  is diagonal in the  $X$  basis when  $i$  is odd, and (iii)  $U_i$  is diagonal in the  $Z$  basis when  $i$  is even.*

**Proof.** Let  $W_\tau \cdots W_1$  be any QMA<sub>1</sub> verification circuit with gates  $W_j$  from the gate set  $\mathcal{G}_{XZ}$ . Recall that by Lemma 4.2, each  $W_j$  is diagonal in either the  $X$  basis or the  $Z$  basis. By padding the circuit with identity operations, we can construct an equivalent circuit  $U_T \cdots U_1$  of  $T = 2\tau + 1$  gates which has the desired form. Concretely, for each  $j \in [\tau]$ , if  $W_j$  is diagonal in the  $X$  basis, set  $U_{2j-1} = W_j$ , and otherwise set  $U_{2j} = W_j$ . The gates  $U_i$  which remain undefined by this procedure are set to be the identity operation. ◀

We now define the clock states in our construction, which differ from those for the original Kitaev Hamiltonian. Let

$$|\hat{0}\rangle = |0+0+\cdots+0\rangle, |\hat{1}\rangle = |1+0+\cdots+0\rangle, |\hat{2}\rangle = |1-0+\cdots+0\rangle, \dots, |\hat{T}\rangle = |1-1-\cdots-1\rangle.$$

To remove ambiguity, we call these states *good clock states*, and refer to the clock states of the Kitaev Hamiltonian as *Kitaev clock states*. Unlike a Kitaev clock state, which has  $Z$  basis states on all qubits, a good clock state consists of alternating  $Z$  and  $X$  basis states. On the  $i$ -th “tick” of the clock, the  $|0\rangle$  in the  $i$ -th position changes to a  $|1\rangle$  if  $i$  is odd and changes from  $|+\rangle$  to  $|-\rangle$  if  $i$  is even.

Our new propagation term has the form  $H_{\text{prop}} = \sum_{t=1}^T H_{\text{prop},t}$ , with

$$H_{\text{prop},t} = \begin{cases} \frac{1}{2}[I_{\mathcal{H}} \otimes (|0+\rangle\langle 0+|_{1,2} + |1+\rangle\langle 1+|_{1,2}) \\ \quad - U_1 \otimes |1+\rangle\langle 0+|_{1,2} - U_1^\dagger \otimes |0+\rangle\langle 1+|] & \text{if } t = 1, \\ \frac{1}{2}[I_{\mathcal{H}} \otimes (|1+0\rangle\langle 1+0|_{t-1,t,t+1} + |1-0\rangle\langle 1-0|_{t-1,t,t+1}) \\ \quad - U_t \otimes |1-0\rangle\langle 1+0|_{t-1,t,t+1} - U_t^\dagger \otimes |1+0\rangle\langle 1-0|_{t-1,t,t+1}] & \text{if } t \text{ is even,} \\ \frac{1}{2}[I_{\mathcal{H}} \otimes (|-0+\rangle\langle -0+|_{t-1,t,t+1} + |-1+\rangle\langle -1+|_{t-1,t,t+1}) \\ \quad - U_t \otimes |-1+\rangle\langle -0+|_{t-1,t,t+1} - U_t^\dagger \otimes |-0+\rangle\langle -1+|] & \text{if } t \text{ is odd, } t \neq 1, \\ & t \neq T, \\ \frac{1}{2}[I_{\mathcal{H}} \otimes (|-0\rangle\langle -0|_{T-1,T} + |-1\rangle\langle -1|_{T-1,T}) \\ \quad - U_T \otimes |-1\rangle\langle -0|_{T-1,T} - U_T^\dagger \otimes |-0\rangle\langle -1|_{T-1,T}] & \text{if } t = T. \end{cases}$$

Importantly, when the gates  $U_t$  come from  $\mathcal{G}_{XZ}$ ,  $H_{\text{prop},t}$  satisfies the conditions imposed on instances of XZ-QUANTUM-6-SAT.

► **Lemma 4.4.** *For gates  $U_t \in \mathcal{G}_{XZ}$ ,  $H_{\text{prop},t}$  is a 6-local projector which is diagonal in either the  $Z$  basis or the  $X$  basis.*

**Proof.** It is straightforward to check that  $H_{\text{prop},t}$  is a projector, i.e.  $H_{\text{prop},t}^2 = H_{\text{prop},t}$ . First, consider when  $t$  is odd and  $t \neq 1, T$ . By Lemma 4.2,  $U_t = U_t^\dagger$ , so by factoring out qubits  $t-1$  and  $t+1$  in the clock register, we have

$$\begin{aligned} H_{\text{prop},t} &= |-+\rangle\langle -+|_{t-1,t+1} \otimes \frac{1}{2} [I_{\mathcal{H}} \otimes (|0\rangle\langle 0|_t + |1\rangle\langle 1|_t) - U_t \otimes (|1\rangle\langle 0|_t + |0\rangle\langle 1|_t)] \\ &= |-+\rangle\langle -+|_{t-1,t+1} \otimes \frac{1}{2} [I_{\mathcal{H}} \otimes I_t - U_t \otimes X_t], \end{aligned}$$

where  $I_t$  and  $X_t$  are the identity and Pauli  $X$  operators on clock qubit  $t$ . By Lemma 4.3,  $U_t$  is diagonal in the  $X$  basis, since  $t$  is odd. In this form, it is clear what basis diagonalizes  $H_{\text{prop},t}$ : the change-of-basis matrix applies  $\hat{H}$  on all qubits. Thus,  $H_{\text{prop},t}$  is diagonal in the  $X$  basis. Moreover, by Lemma 4.2,  $U_t$  is 3-local, so  $H_{\text{prop},t}$  is 6-local.

By a similar argument,  $H_{\text{prop},t}$  is diagonal in the  $Z$  basis for odd  $t$  (including  $t = 1$  and  $t = T$ ), noting that  $|-\rangle\langle +|_t + |-\rangle\langle +|_t = Z_t$ . ◀

## 101:14 Two Bases Suffice for QMA<sub>1</sub>-Completeness

We now define the formatting Hamiltonian term  $H_{\text{format}} = H_{\text{format},X} + H_{\text{format},Z}$ , where

$$\begin{aligned} H_{\text{format},X} &= I_{\mathcal{H}} \otimes (|+-\rangle\langle+-|_{2,4} + |+-\rangle\langle+-|_{4,6} + \cdots + |+-\rangle\langle+-|_{T-3,T-1}), \\ H_{\text{format},Z} &= I_{\mathcal{H}} \otimes (|01\rangle\langle 01|_{1,3} + |01\rangle\langle 01|_{3,5} + \cdots + |01\rangle\langle 01|_{T-2,T}). \end{aligned}$$

The  $Z$  format terms give an energy penalty unless the odd clock qubits form a Kitaev clock state. Likewise, the  $X$  format terms only allow Kitaev clock states on the even clock qubits (under the relabeling  $|0\rangle \rightarrow |+\rangle$  and  $|1\rangle \rightarrow |-\rangle$ ).

It is clear that every good clock state is a zero-energy state of  $H_{\text{format}}$ . However, note that there are some zero-energy states of  $H_{\text{format}}$  which are not good clock states! We refer to these states as *fake clock states*. Intuitively, these are states where the  $X$  and  $Z$  clocks are each valid Kitaev clock states, but they are not properly synchronized with each other. In the following section, we will show the ground space of  $H$  has no support on fake clock states, even though they are zero-energy states of  $H_{\text{format}}$ . We also use the term *bad clock states* to refer to non-zero energy states of  $H_{\text{format}}$ .

### 4.4 $H_{\text{in}}$ and $H_{\text{out}}$

The terms  $H_{\text{in}}$  and  $H_{\text{out}}$  of the Kitaev construction remain unchanged in our construction, and we state them here:

$$H_{\text{in}} = \sum_{i \in Q_{\text{anc}}} |1\rangle\langle 1|_i \otimes |0\rangle\langle 0|_1, \quad H_{\text{out}} = |0\rangle\langle 0|_1 \otimes |1\rangle\langle 1|_T. \quad (3)$$

Within the subspace of clock states, the local check  $|0\rangle\langle 0|_1$  projects onto the subspace spanned by  $|\hat{0}\rangle$ . Thus,  $H_{\text{in}}$  assigns an energy penalty to states whose  $|\hat{0}\rangle$  clock component contains ancilla qubits which are not initialized to  $|0\rangle$ . Likewise,  $H_{\text{out}}$  assigns an energy penalty when the  $|\hat{T}\rangle$  clock component of the state does not have a  $|1\rangle$  on the first qubit of  $\mathcal{H}$  (recall that this is qubit measured at the end of the QMA<sub>1</sub> verification). Note that  $\langle \text{hist} | H_{\text{in}} | \text{hist} \rangle = \langle \text{hist} | H_{\text{out}} | \text{hist} \rangle = 0$ .

## 5 XZ-Quantum-6-Sat is QMA<sub>1</sub>-hard

In this section, we show that XZ-QUANTUM-6-SAT is QMA<sub>1</sub><sup>G<sub>XZ</sub></sup>-hard, which suffices to prove Theorem 1.1 by Lemma 4.1. Completeness is immediate: it is straightforward to check that in the “yes” case, the history state given by Equation (2) (where  $|\hat{t}\rangle$  now refers to a good clock state, not a Kitaev clock state, and  $|\psi\rangle$  is the accepting proof of the QMA<sub>1</sub> verifier) is a zero-energy state of our Hamiltonian  $H$ .

In the remainder of this section, we prove soundness: in the “no” case, any state in  $\mathcal{H} \otimes \mathcal{H}_{\text{clock}}$  has energy at least  $1/\text{poly}(n)$  with respect to  $H$ , i.e. the smallest eigenvalue  $\lambda_{\min}(H)$  is at least  $1/\text{poly}(n)$ . Recall that  $T$  is odd by Lemma 4.3, and let  $T = 2\tau + 1$ .

### Setup and notation

We first partition  $\mathcal{H} \otimes \mathcal{H}_{\text{clock}}$  into subspaces corresponding to good, fake, and bad clock states. For strings  $a = a_1 \cdots a_{j+1}$  and  $b = b_1 \cdots b_j$ , we use  $a \bowtie b = a_1 b_1 a_2 b_2 \cdots a_j b_j a_{j+1}$  to denote the string formed by interleaving  $a$  and  $b$ . Define the set of strings  $S = \{a \bowtie b : a \in \{0, 1\}^{\tau+1}, b \in \{+, -\}^{\tau}\}$ . The states  $\{|s\rangle : s \in S\}$  form a basis for  $\mathcal{H}_{\text{clock}}$ . We partition  $S$  into subsets  $S_{\text{good}}$ ,  $S_{\text{fake}}$ , and  $S_{\text{bad}}$ . First, define

$$S_{\text{format}} = \{1^{t_Z} 0^{\tau+1-t_Z} \bowtie -^{t_X} +^{\tau-t_X} : t_Z \in \{0, \dots, \tau+1\}, t_X \in \{0, \dots, \tau\}\}.$$

Let  $S_{\text{good}} \subseteq S_{\text{format}}$  contain the strings of the above form which satisfy either  $t_Z = t_X$  or  $t_Z = t_X + 1$ . Let  $S_{\text{fake}} = S_{\text{format}} \setminus S_{\text{good}}$  and  $S_{\text{bad}} = S \setminus S_{\text{format}}$ . It is clear that  $S_{\text{good}}$  and  $S_{\text{fake}}$  correspond to good and fake clock states respectively, while  $S_{\text{bad}}$  corresponds to non-zero-energy states of  $H_{\text{format}}$ . Let  $\mathcal{H}_{\text{good}}$  be the subspace of  $\mathcal{H}_{\text{clock}}$  spanned by  $\{|s\rangle : s \in S_{\text{good}}\} = \{|\hat{0}\rangle, |\hat{1}\rangle, \dots, |\hat{T}\rangle\}$ . Note that  $\mathcal{H}_{\text{good}}^\perp$  (i.e. the orthogonal complement of  $\mathcal{H}_{\text{good}}$  in  $\mathcal{H}_{\text{clock}}$ ) is the span of  $\{|s\rangle : s \in S_{\text{fake}} \cup S_{\text{bad}}\}$ .

### Factoring out the good subspace

We show that  $\mathcal{H} \otimes \mathcal{H}_{\text{good}}$  is an invariant subspace of  $H$ . This is clear for  $H_{\text{in}}$ ,  $H_{\text{out}}$ , and  $H_{\text{format}}$  since each term acts as identity or zero on  $\mathcal{H}_{\text{clock}}$ . Then, for  $t \in [T]$  and  $|\varphi\rangle \in \mathcal{H}$ ,

$$H_{\text{prop},t}|\varphi\rangle|\widehat{t-1}\rangle = |\varphi\rangle|\widehat{t-1}\rangle + U_t|\varphi\rangle|\hat{t}\rangle, \quad (4)$$

$$H_{\text{prop},t}|\varphi\rangle|\hat{t}\rangle = U_t^\dagger|\varphi\rangle|\widehat{t-1}\rangle + |\varphi\rangle|\hat{t}\rangle, \quad (5)$$

$$H_{\text{prop},t}|\varphi\rangle|\hat{k}\rangle = 0 \text{ for } k \in \{0, \dots, T\} \setminus \{t-1, t\}, \quad (6)$$

as desired. We can now write

$$\mathcal{H} \otimes \mathcal{H}_{\text{clock}} = \mathcal{H} \otimes (\mathcal{H}_{\text{good}} \oplus \mathcal{H}_{\text{good}}^\perp) = (\mathcal{H} \otimes \mathcal{H}_{\text{good}}) \oplus (\mathcal{H} \otimes \mathcal{H}_{\text{good}}^\perp). \quad (7)$$

In fact,  $\mathcal{H} \otimes \mathcal{H}_{\text{good}}$  and  $\mathcal{H} \otimes \mathcal{H}_{\text{good}}^\perp$  are orthogonal, by the orthogonality of  $\mathcal{H}_{\text{good}}$  and  $\mathcal{H}_{\text{good}}^\perp$ . It follows from Equation (7) that  $\mathcal{H} \otimes \mathcal{H}_{\text{good}}^\perp = (\mathcal{H} \otimes \mathcal{H}_{\text{good}})^\perp$  (the orthogonal complement of  $\mathcal{H} \otimes \mathcal{H}_{\text{good}}$  in  $\mathcal{H} \otimes \mathcal{H}_{\text{clock}}$ ).

We now apply the following linear algebra fact:

► **Lemma 5.1.** *If  $W$  be a  $O$ -invariant subspace of Hilbert space  $V$ , then  $W^\perp$  is  $O^\dagger$ -invariant.*

**Proof.** Let  $z \in W^\perp$ . Denoting the inner product by  $(\cdot, \cdot)$ , we have for any  $w \in W$  that  $(w, O^\dagger z) = (Ow, z) = 0$ , since  $Ow \in W$  and  $z \in W^\perp$ . Thus  $O^\dagger z \in W^\perp$ , as desired. ◀

We conclude that  $\mathcal{H} \otimes \mathcal{H}_{\text{good}}^\perp$  is  $H$ -invariant as well, using Lemma 5.1 and the fact that  $H$  is Hermitian.  $H$  can then be block-diagonalized, simplifying our analysis. We get the eigenvalues of  $H \upharpoonright_{\mathcal{H} \otimes \mathcal{H}_{\text{clock}}}$  (i.e.  $H$  as an operator on  $\mathcal{H} \otimes \mathcal{H}_{\text{clock}}$ ) by taking the union of the eigenvalues of  $H \upharpoonright_{\mathcal{H} \otimes \mathcal{H}_{\text{good}}}$  and those of  $H \upharpoonright_{\mathcal{H} \otimes \mathcal{H}_{\text{good}}^\perp}$ . In particular,

$$\lambda_{\min}(H \upharpoonright_{\mathcal{H} \otimes \mathcal{H}_{\text{clock}}}) = \min(\lambda_{\min}(H \upharpoonright_{\mathcal{H} \otimes \mathcal{H}_{\text{good}}}), \lambda_{\min}(H \upharpoonright_{\mathcal{H} \otimes \mathcal{H}_{\text{good}}^\perp})). \quad (8)$$

We first lower bound  $\lambda_{\min}(H \upharpoonright_{\mathcal{H} \otimes \mathcal{H}_{\text{good}}})$ . Observe that  $H_{\text{format}}$  has zero energy on good states, and the remaining part  $H_{\text{in}} + H_{\text{out}} + H_{\text{prop}}$  is identical to the Hamiltonian in Kitaev's original reduction [15] after doing a change of basis which applies  $\hat{H}$  on all even clock qubits. Since the eigenvalues are invariant under change of basis, the lower bound from the analysis in [15] carries over directly: we have that

$$\lambda_{\min}(H \upharpoonright_{\mathcal{H} \otimes \mathcal{H}_{\text{good}}}) \geq 1/\text{poly}(n). \quad (9)$$

In the next part of this section, we find a lower bound when  $H$  acts on  $\mathcal{H} \otimes \mathcal{H}_{\text{good}}^\perp$ .

### Bounding the energy outside the good subspace

First, note that for any  $|\varphi\rangle \in \mathcal{H} \otimes \mathcal{H}_{\text{good}}^\perp$ ,

$$\langle \varphi | (H_{\text{in}} + H_{\text{out}} + H_{\text{prop}} + H_{\text{format}}) | \varphi \rangle \geq \langle \varphi | (H_{\text{prop}} + H_{\text{format}}) | \varphi \rangle,$$

## 101:16 Two Bases Suffice for QMA<sub>1</sub>-Completeness

since  $H_{\text{in}} + H_{\text{out}}$  is positive definite. It thus suffices to argue that  $H_{\text{prop}} + H_{\text{format}}$  has energy at least  $1/\text{poly}(n)$  on  $\mathcal{H} \otimes \mathcal{H}_{\text{good}}^\perp$  (in fact, we will show that the energy is  $\Omega(1)$ ).

Define  $\mathcal{H}_{\text{fake}}$  and  $\mathcal{H}_{\text{bad}}$  in terms of  $S_{\text{fake}}$  and  $S_{\text{bad}}$  analogously to  $\mathcal{H}_{\text{good}}$ , so that  $\mathcal{H}_{\text{good}}^\perp = \mathcal{H}_{\text{fake}} \oplus \mathcal{H}_{\text{bad}}$ . Suppose we have a state  $|\psi\rangle$  supported only on  $\mathcal{H}_{\text{good}}^\perp$ . We write this state as

$$|\psi\rangle = \sum_i |\psi_i\rangle \otimes |i\rangle_{\text{clock}}, \quad (10)$$

where the vectors  $|i\rangle$  are clock basis states, and the vectors  $|\psi_i\rangle$  are subnormalized. We are now going to bound  $H_{\text{prop}} + H_{\text{format}}$  by decomposing it in terms of the components  $|\psi_i\rangle$ . Firstly,  $H_{\text{format}}$  is simple to bound:

$$\langle \psi | H_{\text{format}} | \psi \rangle \geq \sum_{i \in S_{\text{bad}}} \|\psi_i\|^2. \quad (11)$$

This follows because  $H_{\text{format}}$  assigns an energy penalty of at least 1 to every bad string. Next, let us look at a single propagation term  $H_{\text{prop},t}$ . We will assume that  $t$  is odd and  $1 < t < T$  for simplicity, and also use the fact that all of our gates are self-adjoint to simplify expressions. We will write this propagation term as a sum of terms that look like graph Laplacians over the graph whose vertices consist of clock basis strings, and whose edges correspond to pairs of strings that are paired by  $H_{\text{prop},t}$ .

$$\langle \psi | H_{\text{prop},t} | \psi \rangle = \langle \psi | \left[ \frac{1}{2} (I_{\mathcal{H}} \otimes I_t - (U_t)_{\mathcal{H}} \otimes X_t) \otimes |-\rangle\langle -|_{t-1,t+1} \otimes I_{\text{restofclock}} \right] | \psi \rangle \quad (12)$$

$$= \frac{1}{2} \sum_{i,j} \langle \psi_i | \otimes \langle i |_{\text{clock}} \left[ (I_{\mathcal{H}} \otimes I_t - (U_t)_{\mathcal{H}} \otimes X_t) \otimes |-\rangle\langle -|_{t-1,t+1} \otimes I_{\text{restofclock}} \right] | \psi_j \rangle \otimes | j \rangle_{\text{clock}} \quad (13)$$

$$= \frac{1}{2} \sum_{(i,j) \in E_t} (\langle \psi_i | \psi_i \rangle + \langle \psi_j | \psi_j \rangle - \langle \psi_i | U_t | \psi_j \rangle - \langle \psi_j | U_t | \psi_i \rangle) \quad (14)$$

$$= \frac{1}{2} \sum_{(i,j) \in E_t} \|\psi_i - U_t | \psi_j \rangle\|^2, \quad (15)$$

where  $E_t$  consists of all unordered pairs  $(i, j)$  of strings in  $S_{\text{fake}} \cup S_{\text{bad}}$  such that  $i$  and  $j$  agree at all positions except position  $t$ , *disagree* at position  $t$ , and both strings have  $-$  in position  $t-1$  and  $+$  in position  $t+1$ . (For other values of  $t$  other than the ones we have considered here,  $E_t$  may be defined analogously: we give a full definition encapsulating all edge cases in Definition 6.1.) To simplify notation for what will follow, let us replace  $U_t$  in the last line with  $U_{ij}$ : this is well-defined because it is easy to see that the sets  $E_t$  are disjoint for different values of  $t$ . Thus, we have

$$\langle \psi | H_{\text{prop},t} | \psi \rangle = \frac{1}{2} \sum_{(i,j) \in E_t} \|\psi_i - U_{ij} | \psi_j \rangle\|^2. \quad (16)$$

Putting these together, the total energy of  $H_{\text{prop}} + H_{\text{format}}$  is

$$\langle \psi | (H_{\text{prop}} + H_{\text{format}}) | \psi \rangle \geq \sum_{(i,j) \in E} \frac{1}{2} \|\psi_i - U_{ij} | \psi_j \rangle\|^2 + \sum_{i \in S_{\text{bad}}} \|\psi_i\|^2, \quad (17)$$

where  $E = \bigcup_t E_t$  is the union of all the edges associated with all the terms of  $H_{\text{prop}}$ .

To lower-bound this quantity, we need to show that the fake strings are well-connected to the bad strings, thus forcing a large amount of weight onto the second summation. To do this, we will use two facts shown in the next section to study the connected component structure of the graph  $G$  whose vertices are  $S_{\text{fake}} \cup S_{\text{bad}}$  and whose edges are  $E$ . By Lemma 6.2, every fake string is connected to at least one bad string. This means that the connected components of  $G$  consist either exclusively bad strings, or a mixture of fake and bad strings: there are no components of  $G$  containing only fake strings. Furthermore, by Lemma 6.4, each connected component contains at most one fake string. Thus, the components either consist entirely of bad strings, or exactly one fake string and at least one bad string.

Moreover, we can freely drop edges from the summation in Equation (17), and only lower the energy. We will try dropping all the “bad to bad” edges, so that the only remaining edges are between fake and bad strings. Let us introduce the notation  $N(i)$  to refer to the set of all neighbors in the graph  $G$  of a string  $i$ . Observe that by the facts about the component structure of  $G$  mentioned in the previous paragraph, it holds that for distinct  $i, j \in S_{\text{fake}}$ , the sets  $N(i), N(j)$  are disjoint, so every string  $x \in S_{\text{bad}}$  is contained in at most one  $N(i)$  for  $i \in S_{\text{fake}}$ . Using this characterization, we write Equation (17) after the bad-to-bad edges have been omitted as follows:

$$\begin{aligned} \langle \psi | (H_{\text{prop},t} + H_{\text{format}}) | \psi \rangle &\geq \sum_{i \in S_{\text{fake}}} \left[ \sum_{j \in N(i)} \left( \frac{1}{2} \|\psi_i\rangle - U_{ij} |\psi_j\rangle \right)^2 + \|\psi_j\rangle^2 \right] \\ &\quad + \sum_{i \in S_{\text{bad-leftover}}} \|\psi_i\rangle^2, \end{aligned} \tag{18}$$

where  $S_{\text{bad-leftover}} = S_{\text{bad}} \setminus \bigcup_{i \in S_{\text{fake}}} N(i)$ .

Now, because the neighborhoods  $N(i)$  are disjoint for distinct  $i \in S_{\text{fake}}$ , we see that the entire RHS of Equation (18) can be written as a sum

$$\sum_{i \in S_{\text{fake}} \cup S_{\text{bad-leftover}}} \langle \psi | M_i | \psi \rangle,$$

where the matrices  $M_i$  are Hermitian matrices that act on different factors of the clock space, so that

$$\langle \psi | M_i | \psi \rangle = \begin{cases} \sum_{j \in N(i)} \left( \frac{1}{2} \|\psi_i\rangle - U_{ij} |\psi_j\rangle \right)^2 + \|\psi_j\rangle^2 & \text{if } i \in S_{\text{fake}}, \\ \|\psi_i\rangle^2 & \text{if } i \in S_{\text{bad-leftover}}. \end{cases} \tag{19}$$

From this, it is easy to see that the minimum value of the RHS of Equation (18) is simply the minimum eigenvalue of the matrices  $M_i$ , restricted to their associated factors. For  $i \in S_{\text{bad-leftover}}$ ,  $M_i$  acts as the identity matrix on its corresponding clock sector, so its minimum eigenvalue is 1. It thus remains to lower bound the minimum eigenvalue of  $M_i$  when  $i \in S_{\text{fake}}$ . By Corollary 6.6, this is at least  $1/4$ . So overall, we conclude that for  $|\psi\rangle$  supported on  $\mathcal{H}_{\text{good}}^\perp$ ,

$$\langle \psi | (H_{\text{prop}} + H_{\text{format}}) | \psi \rangle \geq 1/4,$$

and hence

$$\lambda_{\min}(H \upharpoonright_{\mathcal{H} \otimes \mathcal{H}_{\text{good}}^\perp}) \geq 1/4. \tag{20}$$

Combining Equation (8) with Equation (9) and Equation (20), we conclude that

$$\lambda_{\min}(H) \geq 1/\text{poly}(n)$$

as desired, establishing the soundness property.

## 6 Soundness: technical lemmas

In this section we prove the technical lemmas that were used in the soundness proof in the previous section.

### The structure of $H_{\text{prop}}$

We start with some useful facts about how strings in  $S_{\text{fake}} \cup S_{\text{bad}}$  are coupled together by  $H_{\text{prop}}$ . Our first step is to more formally define the graph  $G$  capturing this coupling, which was introduced in the previous section.

► **Definition 6.1.** Define a graph  $G$  with vertex set  $S_{\text{fake}} \cup S_{\text{bad}}$  and edge set  $E = \bigcup_{i=1}^T E_i$ , where

- $(u, v) \in E_1$  iff  $u_1 \neq v_1$ ,  $u_2 = v_2 = +$ , and  $u_i = v_i$  for all  $i \geq 3$ .
- For  $t \in \{2, 4, \dots, T-3, T-1\}$ ,  $(u, v) \in E_t$  iff  $u_t \neq v_t$ ,  $u_{t-1} = v_{t-1} = \mathbf{1}$ ,  $u_{t+1} = v_{t+1} = \mathbf{0}$ , and  $u_i = v_i$  for all  $i \neq t-1, t, t+1$ .
- For  $t \in \{3, 5, \dots, T-4, T-2\}$ ,  $(u, v) \in E_t$  iff  $u_t \neq v_t$ ,  $u_{t-1} = v_{t-1} = -$ ,  $u_{t+1} = v_{t+1} = +$ , and  $u_i = v_i$  for all  $i \neq t-1, t, t+1$ .
- $(u, v) \in E_T$  iff  $u_T \neq v_T$ ,  $u_{T-1} = v_{T-1} = -$ , and  $u_i = v_i$  for all  $i \leq T-2$ .

For vertex  $v$  in  $G$ , let  $\mathcal{C}(v)$  be the connected component of  $G$  which contains  $v$ .

It is clear by the definition of  $H_{\text{prop}}$  that edges in  $G$  correspond to states which are connected under the action of  $H_{\text{prop}}$  (in the sense of Equation (4)). We now prove important lemmas about the connectivity of  $G$  (Lemma 6.2, Lemma 6.4), which allow us to understand  $H_{\text{prop}}$ .

► **Lemma 6.2.** *For every  $f \in S_{\text{fake}}$ , there is some  $b \in S_{\text{bad}}$  such that  $(f, b) \in E$ .*

**Proof.** Let  $f \in S_{\text{fake}}$ . Then  $f = \mathbf{1}^{t_Z} \mathbf{0}^{\tau+1-t_Z} \bowtie -^{t_X} +^{\tau-t_X}$  for some  $t_Z \in \{0, \dots, \tau+1\}$ ,  $t_X \in \{0, \dots, \tau\}$  satisfying  $t_Z \neq t_X$  and  $t_Z \neq t_X + 1$ . We case on  $t_X$ . First, suppose  $t_X = 0$ . Then  $t_Z \geq 2$ , so  $f$  starts with  $\mathbf{1+1}$ . Define  $b$  to be the same string as  $f$ , but with the first position changed to a  $\mathbf{0}$ . Then  $(f, b) \in E_1$ , by the definition of  $E_1$ . Note that there is a  $\mathbf{0}$  followed by a  $\mathbf{1}$  on the odd terms of  $b$ , which puts  $b \in S_{\text{bad}}$ , as desired.

Suppose instead that  $t_X \in [\tau-1]$ . We can subdivide into two cases:  $t_Z$  satisfies either  $t_Z \leq t_X - 1$  or  $t_Z \geq t_X + 2$ . If  $t_Z \leq t_X - 1$ , then  $f$  looks like

$$\dots * \underline{*} \underline{-} \underline{0} \underline{0} \underline{0} \underline{0} \dots,$$

where we have underlined position  $2t_X$  and offset the odd and even terms for clarity, with  $*$  denoting either a  $\mathbf{0}$  or a  $\mathbf{1}$ . Let  $b$  be the same as  $f$  but with position  $2t_X + 1$  changed to a  $\mathbf{1}$ . Then  $(f, b) \in E_{2t_X+1}$ , and  $b \in S_{\text{bad}}$  since  $b_{2t_X-1} = \mathbf{0}$  while  $b_{2t_X+1} = \mathbf{1}$ .

If instead  $t_Z \geq t_X + 2$ , then  $f$  looks like

$$\dots \underline{\mathbf{1}} \underline{\mathbf{1}} \underline{\mathbf{1}} \underline{\mathbf{1}} \underline{*} \underline{*} \dots,$$

where we have underlined position  $2t_X$ . Define  $b$  to be  $f$  but with position  $2t_X + 1$  changed to a  $\mathbf{0}$ . Once again, we have  $(f, b) \in E_{2t_X+1}$ , and  $b \in S_{\text{bad}}$  since  $b_{2t_X+1} = \mathbf{0}$  while  $b_{2t_X+3} = \mathbf{1}$ .

Finally, suppose  $t_X = \tau$ . Then  $t_Z \leq \tau - 1$ , so  $f$  ends in  $\mathbf{0-0}$ . Let  $b$  be the same as  $f$  but with a  $\mathbf{1}$  in the final position. Then  $(f, b) \in E_T$  and  $b \in S_{\text{bad}}$ , completing the proof. ◀

We now prove an intermediate “locking” lemma.

► **Lemma 6.3.** *Let  $s \in S_{\text{fake}} \cup S_{\text{bad}}$ .*

1. *If  $s$  contains  $+1$  as a substring (say, in positions  $k$  and  $k+1$ ), then any string in  $\mathcal{C}(s)$  contains  $+1$  in positions  $k$  and  $k+1$ .*
2. *If  $s$  contains  $0-$  in positions  $k$  and  $k+1$ , then any string in  $\mathcal{C}(s)$  contains  $0-$  in positions  $k$  and  $k+1$ .*

**Proof.** It suffices to consider the four (two-way) rewrite rules

$$[0+ \Leftrightarrow [1+ \quad 1+0 \Leftrightarrow 1-0 \quad -0+ \Leftrightarrow -1+ \quad -0] \Leftrightarrow -1]$$

on string  $s$ , where  $[$  marks the beginning of a string and  $]$  marks the end of a string. We first show (1). By the second rewrite rule, position  $k+1$  must be  $0$  in order for position  $k$  to change under a rewrite starting from  $s$ . Likewise, by the remaining rewrite rules, position  $k$  must be  $-$  in order for position  $k+1$  to change under a rewrite. Together, this implies that positions  $k$  and  $k+1$  are invariant under rewrites from  $s$ , as desired. The analysis is analogous to show (2). ◀

► **Lemma 6.4.** *For any  $f \in S_{\text{fake}}$ ,  $f$  is the only element of  $S_{\text{fake}}$  which is in  $\mathcal{C}(f)$ .*

**Proof.** Fix distinct  $f, g \in S_{\text{fake}}$ . We want to show that  $f$  and  $g$  are in different components of  $G$ . We case on the the first position  $k \in [T]$  at which  $f_k \neq g_k$ . Suppose  $k = 1$  and without loss of generality take  $f_1 = 0$  and  $g_1 = 1$ . Since  $f \in S_{\text{fake}}$ , we must have  $f_i = 0$  for all odd  $i$ .

We now argue that  $f_2 = -$ . If instead  $f_2 = +$ , then  $f_i = +$  for all even  $i$ , since  $f \in S_{\text{fake}}$ . But then  $f = 0^{\tau+1} \bowtie +^\tau \in S_{\text{good}}$ , which is a contradiction.

Applying Lemma 6.3 to  $f_1 f_2 = 0-$ , every string  $h \in \mathcal{C}(f)$  satisfies  $h_1 h_2 = 0-$ . In particular, since  $g_1 = 1$ ,  $g_1 \notin \mathcal{C}(f)$ , as desired. We depict the above deductions as

$$\begin{aligned} f &= \underline{0}0*0\cdots \\ g &= 1****\cdots, \end{aligned}$$

where the underlined positions are those “locked” by Lemma 6.3, and  $*$  indicates a position which is unconstrained (or otherwise unimportant in our argument).

Now suppose  $k = T$ . Without loss of generality, take  $f_T = 0$  and  $g_T = 1$ . We show that

$$\begin{aligned} f &= \cdots****0 \\ g &= \cdots 1*1\underline{1}. \end{aligned}$$

Since  $g \in S_{\text{fake}}$ ,  $g_i = 1$  for all odd  $i$ . Now, note that  $g_{T-1} = +$ : if instead  $g_{T-1} = -$ , then  $g_i = -$  for all even  $i$ , which puts  $g = 1^{\tau+1} \bowtie -^\tau \in S_{\text{good}}$ , a contradiction. Then, applying Lemma 6.3 to  $g_{T-1} g_T = +1$ , every string in  $\mathcal{C}(g)$  ends in  $+1$ , and thus  $f \notin \mathcal{C}(g)$  since  $f_T = 0$ .

In the third case, suppose  $k$  is odd and  $1 < k < T$ . As usual, take  $f_k = 0$  and  $g_k = 1$ . We consider two subcases corresponding to the possible values of  $g_{k-1}$ . The case  $g_{k-1} = +$  is immediate: Lemma 6.3 locks positions  $k-1$  and  $k$  of any string in  $\mathcal{C}(g)$  to  $+1$ , so  $f \notin \mathcal{C}(g)$ . In the other case  $g_{k-1} = -$ , we will deduce that

$$\begin{aligned} f &= 1-1-\cdots 1-\check{0}-0*0\cdots \\ g &= 1-1-\cdots 1-1****\cdots, \end{aligned}$$

where we have accented the  $k$ -th position of  $f$ . Since  $g \in S_{\text{fake}}$ ,  $g_i = 1$  for all odd  $i \leq k$ , and  $g_i = -$  for all even  $i \leq k-1$ . By definition,  $k$  is the first position in which  $f$  and  $g$  differ, so  $f_i = 1$  for all odd  $i < k$ , and  $f_i = -$  for all even  $i \leq k-1$ . Moreover, since  $f \in S_{\text{fake}}$ ,  $f_i = 0$

## 101:20 Two Bases Suffice for QMA<sub>1</sub>-Completeness

for all odd  $i \geq k$ . Together, this implies that  $f_{k+1} = -$ ; if instead  $f_{k+1} = +$ , then  $f_i = +$  for all even  $i \geq k+1$  and thus  $f \in S_{\text{good}}$ , which is a contradiction. Lemma 6.3 then locks positions  $k$  and  $k+1$  of any string in  $\mathcal{C}(f)$  to  $0-$ , so  $g \notin \mathcal{C}(f)$ , as desired.

Finally, suppose  $k$  is even and take  $f_k = +$ ,  $g_k = -$ . This case is analogous to the previous one. First, if  $g_{k-1} = 0$  then Lemma 6.3 locks positions  $k-1$  and  $k$  of any string in  $\mathcal{C}(g)$  to  $0-$ , so  $f \notin \mathcal{C}(g)$ . If instead  $g_{k-1} = 1$ , then it is easy to check that  $f$  and  $g$  have the form

$$\begin{aligned} f &= 1-1-\cdots-1\underline{+}1+**\cdots \\ g &= 1-1-\cdots-1-****\cdots \end{aligned}$$

using the reasoning of the previous case. Thus  $g \notin \mathcal{C}(f)$ , completing the proof.  $\blacktriangleleft$

### Optimizing the energy over one component

We now bound the minimum of the quadratic form appearing in Equation (18) for a component containing a fake string, and show that it is lower-bounded by a constant independent of the number of bad strings in the component. We first show the desired bound in a special case where the computational Hilbert space  $\mathcal{H}$  is one-dimensional, from which the general case will follow as an easy corollary.

► **Lemma 6.5.** *Let  $k \geq 1$  be an integer and let  $\psi = (\psi_0, \dots, \psi_k) \in \mathbb{C}^{k+1}$  be a unit vector. Then*

$$f(\psi) := \sum_{i=1}^k \left( \frac{1}{2} |\psi_0 - \psi_i|^2 + |\psi_i|^2 \right) \geq 1/4. \quad (21)$$

**Proof.** We would like to minimize

$$f(\psi) = \sum_{i=1}^k \left( \frac{1}{2} |\psi_0 - \psi_i|^2 + |\psi_i|^2 \right) \quad (22)$$

subject to  $\psi$  being a unit vector in  $\mathbb{C}^{k+1}$ . Our first observation is to see that we can without loss of generality take all coordinates of  $\psi$  to be real and nonnegative. Next, observe that

$$\sum_{i=1}^k |\psi_0 - \psi_i|^2 = k\psi_0^2 + (1 - \psi_0^2) - 2\psi_0 \sum_{i=1}^k \psi_i \quad (23)$$

$$\geq k\psi_0^2 + (1 - \psi_0^2) - 2\psi_0 \sqrt{k} \cdot \sqrt{\sum_{i=1}^k \psi_i^2} \quad (24)$$

$$= (\sqrt{k}\psi_0 - \sqrt{1 - \psi_0^2})^2, \quad (25)$$

where we have used the Cauchy-Schwarz inequality in passing to (24), and then applied the normalization condition. So overall the quantity we want to minimize is

$$f(\psi) \geq g(\psi_0) := \frac{1}{2} (\sqrt{k}\psi_0 - \sqrt{1 - \psi_0^2})^2 + (1 - \psi_0^2) \quad (26)$$

$$= \begin{pmatrix} \psi_0 & \sqrt{1 - \psi_0^2} \end{pmatrix} \cdot \underbrace{\begin{pmatrix} k/2 & -\sqrt{k}/2 \\ -\sqrt{k}/2 & 3/2 \end{pmatrix}}_G \cdot \begin{pmatrix} \psi_0 \\ \sqrt{1 - \psi_0^2} \end{pmatrix}. \quad (27)$$

Thus, we have reduced the problem to finding the minimum eigenvalue of the  $2 \times 2$  matrix  $G$ . Through explicit computation, we can find the eigenvalues of this matrix.

$$\lambda = \frac{3+k}{4} \cdot (1 \pm \sqrt{1 - 8k/(3+k)^2}). \quad (28)$$

It is not hard to show that for all integer  $k \geq 1$ ,  $\lambda \geq 1/4$  (for details, see the full version). ◀

► **Corollary 6.6.** *Let  $k, d \geq 1$  be integers and let  $|\alpha_0\rangle, \dots, |\alpha_k\rangle$  be (not necessarily unit) vectors over  $\mathbb{C}^d$  such that  $\sum_{i=0}^k \|\alpha_i\|^2 = 1$ . Moreover, let  $U_1, \dots, U_k$  be  $d \times d$  unitary matrices. Then*

$$\sum_{i=1}^k \left( \frac{1}{2} \|\alpha_0 - U_i \alpha_i\|^2 + \|\alpha_i\|^2 \right) \geq 1/4. \quad (29)$$

**Proof.** Observe that for any two vectors  $|\alpha\rangle, |\beta\rangle$  and for any unitary  $U$ , it holds that

$$\|\alpha - U\beta\|^2 = \|\alpha\|^2 + \|\beta\|^2 - \langle \alpha | U | \beta \rangle - \langle \beta | U | \alpha \rangle \quad (30)$$

$$\geq \|\alpha\|^2 + \|\beta\|^2 - 2\|\alpha\| \cdot \|\beta\| \quad (31)$$

$$= \|\alpha\| - \|\beta\|. \quad (32)$$

Thus, if we let  $\psi_i = \|\alpha_i\|$ , then the vector  $(\psi_0, \dots, \psi_k)$  satisfies the conditions of Lemma 6.5, and by Equation (32), the quantity we wish to bound is at least  $f(\psi) \geq 1/4$  by Lemma 6.5. ◀

---

## References

- 1 Dorit Aharonov. A Simple Proof that Toffoli and Hadamard are Quantum Universal, 2003. doi:10.48550/arXiv.quant-ph/0301040.
- 2 Dorit Aharonov, Itai Arad, and Thomas Vidick. Guest column: The quantum PCP conjecture. *SIGACT News*, 44(2):47–79, 2013. doi:10.1145/2491533.2491549.
- 3 Eric R. Anschuetz, David Gamarnik, and Bobak Kiani. Combinatorial NLTS from the overlap gap property. *Quantum*, 8:1527, 2024. doi:10.22331/Q-2024-11-19-1527.
- 4 Anurag Anshu, Nikolas P Breuckmann, and Quynh T Nguyen. Circuit-to-hamiltonian from tensor networks and fault tolerance. In *Proceedings of the 56th Annual ACM Symposium on Theory of Computing*, pages 585–595, 2024. doi:10.1145/3618260.3649690.
- 5 Anurag Anshu, Nikolas P. Breuckmann, and Chinmay Nirkhe. NLTS Hamiltonians from good quantum codes. In *Proceedings of the 55th Annual ACM Symposium on Theory of Computing*, STOC 2023, pages 1090–1096, New York, NY, USA, 2023. Association for Computing Machinery. doi:10.1145/3564246.3585114.
- 6 Jacob D. Biamonte and Peter J. Love. Realizable Hamiltonians for universal adiabatic quantum computers. *Physical Review A*, 78(1):012352, 2008. doi:10.1103/PhysRevA.78.012352.
- 7 Sergey Bravyi. Efficient algorithm for a quantum analogue of 2-SAT, 2006. doi:10.48550/arXiv.quant-ph/0602108.
- 8 Sergey Bravyi and Barbara Terhal. Complexity of stoquastic frustration-free hamiltonians. *Siam journal on computing*, 39(4):1462–1485, 2010. doi:10.1137/08072689X.
- 9 Stephen A. Cook. The complexity of theorem-proving procedures. In *Proceedings of the Third Annual ACM Symposium on Theory of Computing*, STOC '71, pages 151–158, New York, NY, USA, 1971. Association for Computing Machinery. doi:10.1145/800157.805047.
- 10 Toby Cubitt and Ashley Montanaro. Complexity Classification of Local Hamiltonian Problems. *SIAM Journal on Computing*, 45(2):268–316, 2016. doi:10.1137/140998287.
- 11 Christopher M. Dawson and Michael A. Nielsen. The Solovay-Kitaev algorithm. *Quantum Info. Comput.*, 6(1):81–95, 2006. doi:10.26421/QIC6.1-6.

## 101:22 Two Bases Suffice for $\text{QMA}_1$ -Completeness

- 12 Lior Eldar and Aram W. Harrow. Local Hamiltonians Whose Ground States Are Hard to Approximate. In *2017 IEEE 58th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 427–438, 2017. doi:10.1109/FOCS.2017.46.
- 13 David Gosset and Daniel Nagaj. Quantum 3-SAT Is  $\text{QMA}_1$ -Complete. *SIAM Journal on Computing*, 45(3):1080–1128, 2016. doi:10.1137/140957056.
- 14 Robbie King and Tamara Kohler. Gapped Clique Homology on Weighted Graphs is  $\text{QMA}_1$ -Hard and Contained in QMA. In *2024 IEEE 65th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 493–504, 2024. doi:10.1109/FOCS61266.2024.00039.
- 15 A. Yu. Kitaev, A. Shen, and M. N. Vyalyi. *Classical and Quantum Computation*. American Mathematical Society, USA, June 2002.
- 16 L. A. Levin. Universal sequential search problems. *Problems of Information Transmission*, 9(3):265–266, 1973.
- 17 Michael A. Nielsen and Isaac L. Chuang. *Quantum Computation and Quantum Information: 10th Anniversary Edition*, 2010. doi:10.1017/CB09780511976667.
- 18 Dorian Rudolph. Towards a universal gateset for  $\text{QMA}_1$ , 2025. doi:10.48550/arXiv.2411.02681.
- 19 Yaoyun Shi. Both Toffoli and controlled-NOT need little help to do universal quantum computing. *Quantum Info. Comput.*, 3(1):84–92, 2003. doi:10.26421/QIC3.1-7.
- 20 Adam Wills, Ting-Chun Lin, and Min-Hsiu Hsieh. Local testability of distance-balanced quantum codes. *npj Quantum Information*, 10(1):120, 2024. doi:10.1038/s41534-024-00908-8.