Circuit-to-Hamiltonian from tensor networks and fault tolerance

Anurag Anshu* Nikolas P. Breuckmann† Quynh T. Nguyen‡

Abstract

We define a map from an arbitrary quantum circuit to a local Hamiltonian whose ground state encodes the quantum computation. All previous maps relied on the Feynman-Kitaev construction, which introduces an ancillary 'clock register' to track the computational steps. Our construction, on the other hand, relies on injective tensor networks with associated parent Hamiltonians, avoiding the introduction of a clock register. This comes at the cost of the ground state containing only a noisy version of the quantum computation, with independent stochastic noise. We can remedy this - making our construction robust - by using quantum fault tolerance. In addition to the stochastic noise, we show that any state with energy density exponentially small in the circuit depth encodes a noisy version of the quantum computation with adversarial noise. We also show that any 'combinatorial state' with energy density polynomially small in depth encodes the quantum computation with adversarial noise. This serves as evidence that any state with energy density polynomially small in depth has a similar property.

As an application, we show that contracting injective tensor networks to additive error is BQP-hard. We also discuss the implication of our construction to the quantum PCP conjecture, combining with an observation that QMA verification can be done in logarithmic depth.

1 Introduction

The Feynman-Kitaev 'clock based' mapping [1] from quantum circuits to local Hamiltonians is the central tool bridging quantum complexity theory and quantum many-body physics. The mapping and its variants have been used to justify the hardness of computing the ground energy of natural local Hamiltonians [2, 3, 4, 5, 6, 7]. It has been used to construct explicit local Hamiltonians with 'complex' ground states - in terms of large entanglement entropy [8, 9] or circuit depth [10]. Other important applications include the equivalence of adiabatic and circuit models [11], delegation of quantum computing [12] etc. However, a well known limitation of the Feynman-Kitaev mapping is the soundness. While quantum computations that output 'accept' with probability (near) 1 get mapped to (near) frustration-free local Hamiltonians, the quantum computations that output 'reject' with high probability get mapped to local Hamiltonians with ground energy density $1/\operatorname{poly}(\text{number of gates})$. This serves as the main bottleneck to the quantum PCP conjecture [13, 14], which seeks a constant energy density in the rejecting case.

An alternative mapping of quantum computation to many-body systems was laid out in [15] by using measurement-based quantum computing (MBQC). It was shown that running MBQC and post-selecting on 'no correction' led to a tensor network which encoded the result of the quantum computation. However, this technique does not yield a desired circuit-to-Hamiltonian mapping due to two issues. First, the encoding tensor network may not be the ground state of any local Hamiltonian. Second, the tensor networks also capture quantum computation with post-selection, which leads to a class much larger than QMA.

Our starting point is the observation that both the issues no longer exist if we consider the class of injective tensor networks. Injectivity prevents us from post-selecting on events of very small probability. The injective tensor networks also have a natural parent Hamiltonian. The price we pay is that the tensor

*School of Engineering and Applied Sciences, Harvard University
†School of Mathematics, University of Bristol
‡School of Engineering and Applied Sciences and Harvard Quantum Initiative, Harvard University

Feynman-Kitaev construction [1]	**Present construction**
Ground state: superposition over partial computations of W	Ground state: tensor network encoding a noisy version of W with i.i.d noise per wire
States with energy density $\frac{O(1)}{\|W\|^3}$ encode W	Combinatorial states with $\frac{O(1)}{D}$ fraction violations encode a noisy version of W with adversarial noise (Theorem 4.4). • States with energy density $e^{-\Omega(D \log D)}$ (for $D = o(\log \|W\|)$) encode a noisy version of W with adversarial noise (Theorem 4.3).
There exists a combinatorial state with $\frac{O(1)}{\|W\|}$ fraction of violations containing no information about W (see a Note in the proof of Claim 5.1).	There exists a combinatorial state with $\frac{O(1)}{D}$ fraction of violations contain no information about W.

Table 1: A comparison between Feynman-Kitaev and our construction for a QMA verification circuit W. Above, D is the depth of the circuit. Our main open question is that any state with energy density $\frac{1}{\text{poly}(D)}$ encode noisy version of W with adversarial noise. Since we can choose $D = O(\log |W|)$ in QMA protocols (Section 5), this serves as a link between polylog weaker quantum PCP and adversarial quantum fault tolerance.

network now represents a noisy version of the quantum circuit. This is handled by considering a fault-tolerant version of the circuit.

The details of the construction appear in Section 2, where we use standard teleportation instead of measurement-based quantum computing. A high level overview is as follows, using a simple circuit $U_2U_1 |0\rangle$ involving 1 qubit gates on $|0\rangle$. Introduce 5 qubits in the state $|0\rangle \otimes (\mathrm{I} \otimes U_1) |\Phi_\mathrm{I}\rangle \otimes (\mathrm{I} \otimes U_2) |\Phi_\mathrm{I}\rangle$, where $|\Phi_\mathrm{I}\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$. Projecting qubits 1,2 and 3,4 with $|\Phi_\mathrm{I}\rangle\langle\Phi_\mathrm{I}|$ would lead to the desired state $U_2U_1 |0\rangle$ on qubit 5. However, this is not an injective tensor network as the map $|\Phi_\mathrm{I}\rangle\langle\Phi_\mathrm{I}|$ is not injective. Instead, we project with the map $|\Phi_\mathrm{I}\rangle\langle\Phi_\mathrm{I}| + \delta(\mathrm{I} - |\Phi_\mathrm{I}\rangle\langle\Phi_\mathrm{I}|)$. It can be verified that the last qubit is now a noisy version of the original circuit (with depolarizing noise of strength $O(\delta^2)$) and qubits 1,2 and 3,4 record the Pauli errors.

This scheme applies to general quantum circuits. Our main technical contribution is a characterization of low energy states of the parent Hamiltonian of the above tensor network. We exhibit the following properties for a quantum circuit W of depth D (that may be, for examples, a QMA verification circuit or a BQP circuit).

- Any state with energy density $e^{-\Omega(D \log D)}$ and $D = o(\log |W|)$ can be viewed as the output of the circuit with $O(\delta^2)$ fraction of adversarial noise. See Section 4.3 for the proof idea as well as the detailed proof.
- Any combinatorial state with energy density (equal to the fraction of violated constraints) $\frac{1}{\text{poly}(D)}$ can be viewed as the output of the circuit with $O(\delta^2)$ fraction of adversarial noise. See Section 4.2 for the proof idea as well as the detailed proof.
- The ground state is a noisy version with stochastic iid noise with strength $O(\delta^2)$ per wire. See Section 2.3.

We choose $\delta = \frac{1}{\text{poly}(D)}$ to keep the fraction of adversarial noise $\frac{1}{\text{poly}(D)}$, which keeps the error budget low enough that the adversary can not stop the whole computation. The main open question is that any $1/\text{poly}(D)$ energy state can be viewed as the output of adversarial noisy version. The second result above (on combinatorial states) is evidence in its favor.

Application to quantum PCP conjecture: The quantum PCP conjecture [13, 14] states that it is QMA-hard to decide if the ground energy density of a local Hamiltonian problem is less than a given number a or

more than $a+\Delta$ for a constant Δ. A 'polylog weaker' version of this conjecture - QMA hardness of deciding that ground energy density is $\leq a$ or $> a + \frac{1}{\text{polylog}(n)}$ - is also open.[1]

Our attempt in this work is to link adversarial quantum fault tolerance with the above 'polylog weaker' quantum PCP. At a high level, we expect such a connection due to the correspondence between Hamiltonians and quantum circuits [1] and the view that quantum PCP conjecture is about adversarial violations of local Hamiltonian terms. An issue with this is that quantum PCP conjecture expects soundness against constant fraction of violations, but in a depth D quantum circuit we can at most expect $O(\frac{1}{D})$ adversarial errors. However, as shown in Section 5, QMA verification can be achieved in logarithmic depth ($D = O(\log n)$; n is the number of qubits in the QMA verifier circuit). Thus, if we seek the 'polylog weaker' quantum PCP, connection with adversarial quantum fault tolerance can be more transparent.

Our result takes a step towards this connection by showing that combinatorial states with $\frac{1}{\text{poly}(D)}$ fraction violations encode a circuit with adversarial errors. Suppose $\frac{1}{\text{poly}(D)}$ energy density states in our construction also encode a circuit with adversarial error, which is our main open question. And suppose any $O(\log(n))$-depth QMA verifier can be transformed into a polylog(n)-depth QMA verifier that is sound against $\frac{1}{\text{polylog}(n)}$ fraction of adversarial errors in the circuit.[2] Then the 'polylog weaker' version of quantum PCP holds.

Classical analogue of this line of argument is similar, which we discuss in Appendix A.

Complexity of injective tensor networks: Injective tensor networks constitute a more physical family of quantum states and have been shown to be efficiently preparable on a quantum computer [17, 18] or contractable in classical quasi-polynomial time [19] under assumptions on the parent Hamiltonian spectral gap. However, the lack of the postselection ability makes it less clear how to characterize injective TN from a complexity-theoretic point of view.

Combining our construction with existing quantum fault-tolerance schemes for local stochastic noise [20], we conclude that preparing injective TN states on a quantum computer is BQP-hard. This can be seen as a complement to prior works [17, 18], that showed preparing injective TN states under spectral gap assumptions is in BQP. Compared with the PostBQP-hardness shown in [15], the BQP-hardness naturally reflects the non-postselecting nature of injective TN. Regarding the classical complexity of injective TN, our construction also implies that evaluating local observable expectation values on injective-TN states is BQP-hard to $O(1)$-additive error. In addition, we show the same task for a non-local observable is #P-hard to $O(1)$-multiplicative error.

2 The Model

Let us first outline the general idea behind the construction: given a quantum circuit W, we consider a tensor network associated with the implementation of W (Section 2.2). We make the tensor network injective by perturbing each of its projectors P_i by some small amount δ, so that we can associate it with a parent Hamiltonian (Section 2.3). However, these local perturbations are unwanted. Crucially, we observe that they can be interpreted as Pauli-errors occurring during the execution of W. Hence, we have to consider a fault-tolerant version of W, which requires us to implement a quantum error correction protocol within the model itself.

2.1 Notations

Let the EPR states be $|\Phi_I\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$, $|\Phi_X\rangle = (\mathrm{I}\otimes X)|\Phi_I\rangle, |\Phi_{XZ}\rangle = (\mathrm{I}\otimes XZ)|\Phi_I\rangle, |\Phi_Z\rangle = (\mathrm{I}\otimes Z)|\Phi_I\rangle$. Denote $\mathcal{P} = \{I, X, XZ, Z\}$. For an operator A in a Hilbert space with tensor product structure $\mathcal{H} = (\mathbb{C}^d)^{\otimes n}$, we denote by $\text{supp}(A)$ the span of eigenvectors of A with nonzero eigenvalues and by $\text{loc}(A)$ the set of subsystems on which A acts nontrivially.

[1] It is QMA hard to decide that ground energy density is $\leq a$ or $> a + \frac{1}{n^c}$ for any constant $c > 0$, by a simple modification of the Feynman-Kitaev clock [16].

[2] Note that we also need soundness against a superposition over adversarial errors - see Section 4.

2.2 Quantum circuit to tensor network

We will now discuss the definition of the tensor network T associated to the circuit W. Let n be the total number of qubits on which W operates and D its depth. For simplicity and without loss of generality, let us assume that W consists of 2-qubit gates arranged in a brickwork layout, see Figure 1. The generalization to arbitrary circuits is straightforward. Consider a 2-qubit gate $U^{(\ell)}_{p,q}$ acting one qubits p and q at layer ℓ, we

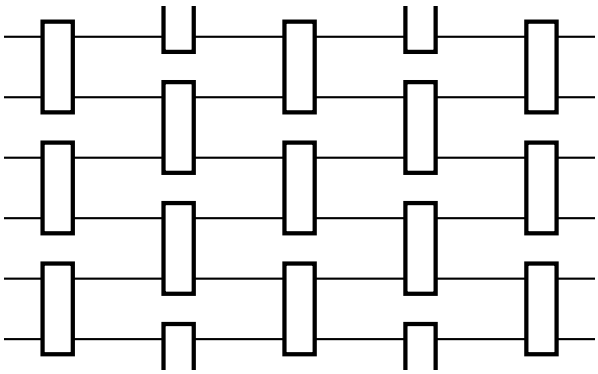


Figure 1: The circuit W consisting of a collection of gates (black boxes). This layout suffices to implement an arbitrary quantum circuit. However, our construction applies to general circuit layouts.

assign a 4-qubit state $|\Phi_U\rangle$ encoding the gate as follows

$$\left|\Phi^{(\ell)}_{p,q}\right\rangle = [I_{1,2} \otimes (U^{(\ell)}_{p,q})_{3,4}] \left(|00\rangle_{1,3} + |11\rangle_{1,3}\right) \otimes \left(|00\rangle_{2,4} + |11\rangle_{2,4}\right)/2. \tag{1}$$

See Figure 2 for a diagrammatic representation of this state.

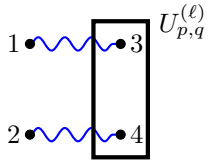


Figure 2: Representation of the state $\left|\Phi^{(\ell)}_{p,q}\right\rangle$. Qubits 1 and 3, as well as 2 and 4, are in the Bell state $|\Phi_I\rangle$, which is indicated by the blue wavy lines. The unitary $U^{(\ell)}_{p,q}$ is applied to qubits 3 and 4 (black box).

Our starting point is that the state $\left|\Phi^{(\ell)}_{p,q}\right\rangle$ allows implementing the gate $U^{(\ell)}_{p,q}$ via teleportation and postselection. For example, the application of gate $U^{(1)}_{p,q}$ on an input state $|\xi\rangle$ is simulated by projecting the joint system $|\xi_{p,q}\rangle \otimes \left|\Phi^{(1)}_{p,q}\right\rangle_{1,2}$ onto the EPR state $|\Phi_I\rangle = (|00\rangle + |11\rangle) \otimes (|00\rangle + |11\rangle)/2$. More generally, a gate $U^{(t)}_{p,q}$ can be effected by applying the projector

$$P = |\Phi_I\rangle \langle\Phi_I| \tag{2}$$

onto qubits p, q of the input state to the gate and qubits 1, 2 of $\left|\Phi^{(\ell)}_{p,q}\right\rangle$, see Figure 3.

For brevity we will often denote $|\Phi_U\rangle = (\mathrm{I} \otimes U) |\Phi_I\rangle^{\otimes 2}$, where U is a two-qubit gate acting on the second qubits of $|\Phi_I\rangle$, leaving the location in spacetime of U implicit.

Extending the previous idea to the entire circuit, we can encode any n-qubit quantum circuit W into a tensor network. In particular, we have n qubits in the first column of the tensor network storing the input

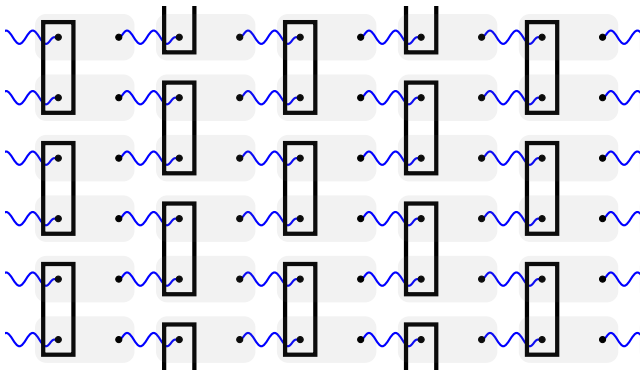


Figure 3: The circuit W (Figure 1) converted into a tensor network. We introduce a Bell pair for every position in the circuit (black dots connected by a wavy line) and apply the unitary operation corresponding to the location in the circuit (cf. Figure 2). We then apply projectors on pairs of qubits (gray boxes).

state, and the other columns storing the EPR encoding of the gates. The total number of qubits of the tensor network is $(2D+1)n$. Define the $(2D+1)n$-qubit product state

$$|\Phi_{W,\xi}\rangle = |\xi\rangle \otimes \bigotimes_{\ell,p,q} \left|\Phi_{p,q}^{(\ell)}\right\rangle, \tag{3}$$

where $|\xi\rangle$ is the n-qubit input state of the circuit W. For example, let $|\xi\rangle = |0\rangle^{\otimes n}$. Then applying the EPR projector $\Pi_W \triangleq \bigotimes_{\ell,p,q} P_{p,q}^{(\ell)}$ on $|\Phi_{W,\xi}\rangle$ results in the output state in the last column

$$\Pi_W |\Phi_{W,\xi}\rangle = |\Phi_I\rangle^{\otimes nD} \otimes \left(W |0\rangle^{\otimes n}\right). \tag{4}$$

Tensor networks of this form are in general termed projected entangled pair states (PEPS).

2.3 Making the tensor network injective

We say that a tensor network is δ-*injective* when its local maps have singular values lower bounded by δ. The tensor network defined in the previous section is non-injective since the projectors are singular. To make the tensor network injective, we follow the procedure in [21] and replace the projectors P by a δ-perturbation

$$Q = |\Phi_I\rangle\langle\Phi_I| + \delta \sum_{P\in\{X,XZ,Z\}} |\Phi_P\rangle\langle\Phi_P|. \tag{5}$$

Applying the invertible map Q on every other pair of row-adjacent qubits in $|\Phi_{W,\xi}\rangle = |\xi\rangle \bigotimes_{\ell\in[D],p,q} \left|\Phi_{p,q}^{(\ell)}\right\rangle$ we obtain the injective PEPS state

$$|\Psi_{W,\xi}\rangle \triangleq Q^{\otimes nD} |\Phi_{W,\xi}\rangle. \tag{6}$$

We introduce several notations. Let $T = nD$ be the number of gates, let $\left|\Phi_{\vec{P}}\right\rangle = \bigotimes_{i=1}^{T} |\Phi_{P_i}\rangle$ for $\vec{P} \in \mathcal{P}^{\otimes T}$ and let $|\vec{P}|$ denote the number of nontrivial operators in $\vec{P}$. Let $W_\ell = \otimes_{i\in \text{ layer } \ell} U_i$ and $\tilde{P}_\ell = \otimes_{i\in \text{ layer } \ell} P_i$ be the unitaries and the errors in the ℓ-th layer of W. Abusing notation, we sometimes denote $U_i \in W_\ell$ and $P_i \in \tilde{P}_\ell$ to mean that the unitaries and Pauli errors are in layer ℓ.

The key observation is that the injective tensor network represents a noisy version of the quantum computation.

Claim 2.1. *The state* $|\Psi_{W,\xi}\rangle$ *can be expanded as* $|\Psi_{W,\xi}\rangle = \sum_{\vec{P}\in\mathcal{P}^{\otimes T}} \delta^{|\vec{P}|} \left|\Phi_{\vec{P}}\right\rangle \otimes (U_T P_T \dots U_1 P_1 |0^a\rangle |\xi\rangle)$.

Proof. Expanding Q we have

$$|\Psi_{W,\xi}\rangle = \sum_{\vec{P}\in\mathcal{P}^{\otimes T}} \delta^{|\vec{P}|} \left|\Phi_{\vec{P}}\right\rangle \left\langle\Phi_{\vec{P}}\middle|\Phi_{W,\xi}\right\rangle. \tag{7}$$

Performing teleportation for each term in the summand, we find $\left\langle\Phi_{\vec{P}}\middle|\Phi_{W,\xi}\right\rangle = U_T P_T \dots U_1 P_1 |0^a\rangle |\xi\rangle$ as the state in the rightmost column of the tensor network. □

In other words, $|\Psi_{W,\xi}\rangle$, up to normalization, contains a noisy quantum computation with purified local depolarizing channels. The local i.i.d. depolarizing noise rate is $p = \delta^2/(1+3\delta^2)$. Here 'purified' means that the EPR states in the bulk of the tensor network record the occurred errors.

Claim 2.1 can be alternatively written as $|\Psi_{W,\xi}\rangle \propto \sum_{\vec{P}\in\mathcal{P}^{\otimes T}} \delta^{|\vec{P}|} \left|\Phi_{\vec{P}}\right\rangle \otimes (W_D \tilde{P}_D \dots W_1 \tilde{P}_1 |\xi\rangle)$.

We can define a unitary

$$V = \sum_{\vec{P}\in\mathcal{P}^{\otimes nD}} \left|\Phi_{\vec{P}}\right\rangle \left\langle\Phi_{\vec{P}}\right| \otimes (W_D \tilde{P}_D \dots W_1 \tilde{P}_1), \tag{8}$$

such that

$$V^\dagger |\Psi_{W,\xi}\rangle = \sum_{\vec{P}\in\mathcal{P}^{\otimes T}} \delta^{|\vec{P}|} \left|\Phi_{\vec{P}}\right\rangle \otimes |\xi\rangle.$$

Note that the state $\sum_{\vec{P}\in\mathcal{P}^{\otimes T}} \delta^{|\vec{P}|} \left|\Phi_{\vec{P}}\right\rangle$ is a tensor product of T i.i.d pure states. Thus, when $|\xi\rangle = |0\rangle^{\otimes n}$ (which arises for computations in BQP), the state $|\Psi_{W,\xi}\rangle$ can be prepared by a quantum circuit. This is similar to the history state [1], which can be prepared efficiently for quantum computations in BQP.

2.4 The parent Hamiltonian

The nice property of the injective tensor network state $|\Psi_{W,\xi}\rangle$ is that it is the unique ground state of a local Hamiltonian. In particular, we consider the $n(2D+1)$-qubit Hilbert space containing the PEPS state $|\Psi_{W,\xi}\rangle$ corresponding to a circuit W.

Let $\Lambda = \delta |\Phi_I\rangle\langle\Phi_I| + \sum_{p\in\{X,XZ,Z\}} |\Phi_p\rangle\langle\Phi_p|$, such that $Q \propto \Lambda^{-1}$.

Definition 2.2 (Parent Hamiltonian)**.** *Associate for each gate two-qubit gate U in the circuit an 8-qubit Hamiltonian term $h_U = \Lambda^{\otimes 4}(\mathrm{I} - |\Phi_U\rangle\langle\Phi_U|)\Lambda^{\otimes 4}$. Furthermore, suppose the initial state $|\xi\rangle$ is the unique ground state of a frustration-free local Hamiltonian $H_\xi = \sum_j g_j$. Then the unnormalized state $\Phi_{W,\xi}$ is the unique ground state of the frustration-free Hamiltonian $H_{\text{parent}} = \sum_j \Lambda^{\otimes N(j)} g_j \Lambda^{\otimes N(j)} + \sum_{U\in W} h_U$, where $N(j)$ is the set of EPRs that have intersecting support with g_j. We refer to the first term as H_{in} and the second term as H_{prop}.*

An example of H_ξ is $H_{|0^n\rangle} = \sum_{i=1}^n |1\rangle\langle 1|_i$ which ensures the input state is $|\xi\rangle = |0\rangle^{\otimes n}$.

In a QMA protocol, we relax the condition that the initial state is unique. Instead, the initial state is of the form $|0^a\rangle|\xi\rangle$, where $|\xi\rangle$ is any $(n-a)$-qubit witness coming from the prover. So $H_{\text{in}} = \Lambda^{\otimes n}(\sum_{j=1}^a |1\rangle\langle 1|_j^{\text{in}})\Lambda^{\otimes n}$ has a ground space of degeneracy 2^{n-a} and so does H_{parent}. See Figure 4 for an example of an injective PEPS and its parent Hamiltonian. Later when we work with states of this form, we will continue referring to the ground states as $|\Psi_{W,\xi}\rangle$, leaving the ancillas $|0^a\rangle$ in the initial state implicit. As long as δ is chosen such that the noise rate is smaller than the fault tolerance threshold, see Theorem 3.9, the ground states of H_{parent} contain the desired quantum computation when W is replaced by its fault-tolerant version W_{FT} in the basic noise model (e.g., using the scheme in [20]).

Finally, similar to the Feynman-Kitaev construction, we can use an output check term $H_{\text{out}} \triangleq |1\rangle\langle 1|_j^{\text{out}}$ to verify qubit j in the output.

2.5 Connection with prior works

A scheme related to ours is that of Ref. [22] in which the authors give a construction of quantum error-correcting subsystem codes with almost linear distance. Their construction can be understood as a map from fault-tolerant Clifford circuits that facilitate check measurements to a set of non-commuting Pauli-check operators. More concretely, each location in the circuit is associated with a qubit and each Clifford gate is associated with a Pauli operator that stabilizes the gate. For example, the idling gate (wire) is stabilized by XX and ZZ operating on the in- and out-locations. The main difference with our setting is that we do not need to assume Clifford circuit. Furthermore, our Hamiltonian remains frustration-free, whereas the Hamiltonian in Ref [22] is frustrated. Another difference is that we associate two qubits per circuit location that are projected onto an EPR state, cf. Figure 3.

In Ref. [23] Bartlett and Rudolph show using PEPS that a fault-tolerant cluster state, which is a universal resource state for MBQC, can be robustly encoded into the ground state of a Hamiltonian consisting of planar, 2-local interaction terms. They also note that the approximation error can be interpreted as stochastic Pauli-noise and that the energy gap of their construction is independent of system size. The difference to our approach is that Bartlett and Rudolph use tensor networks to obtain a resource state that can be used for quantum computation via MBQC, whereas our scheme encodes a quantum computation into a tensor network.

In [24] Aharonov and Irani consider a mapping of classical computation into a CSP, which we may think of as a classical local Hamiltonian. More concretely, they consider a two-dimensional $L \times L$ grid with translation invariant constraints and show that approximating the ground state energy to an additive $\Theta(\sqrt[4]{L})$ is NEXP-complete. They do so by encoding a computation into a tiling problem. The computation is fault tolerant by running the same computation several times in parallel to enforce a large cost for an incorrect computation. In contrast, our model is fully quantum and thus requires the quantum fault tolerance theorem of Ref. [20].

3 Background

3.1 Hamiltonian complexity

Here, we give a brief introduction to the complexity class QMA and main lemmas used in this work.

Definition 3.1. *The class* $\mathsf{QMA}_w[c,s]$ *is the class of promise problems* $A = (A_{\text{yes}}, A_{\text{no}})$ *with the property that, for every instance* x*, there exists a uniformly generated verifier quantum circuit* V_x *with the following properties:* V_x *is of size* $\text{poly}(|x|)$ *and acts on an input state* $|0^{\otimes m}\rangle$ *together with a witness state* $|\xi\rangle$ *of size* w *supplied by an all-powerful prover, with both* $m, w = \text{poly}(|x|)$*. Upon measuring the decision qubit* o*, the verifier accepts if* $o = 1$*, and rejects otherwise. If* $x \in A_{\text{yes}}$*, then* $\exists\, |\xi\rangle$ *such that* $\Pr[o = 1] \geq c$ *(completeness). If* $x \in A_{\text{no}}$*, then* $\forall\, |\xi\rangle$*,* $\Pr[o = 1] \leq s$ *(soundness), such that* $c - s \geq 1/\text{poly}(|x|)$*.*

It is well-known that the parameters c, s can be amplified, even without increasing the witness size.

Lemma 3.2 (Weak QMA amplification [1])**.** *For any* $r = \text{poly}(|x|)$*,* $\mathsf{QMA}_w[2/3, 1/3] = \mathsf{QMA}_{w'}[1 - 2^{-r}, 2^{-r}]$ *where* $w' = \text{poly}(w)$*.*

Lemma 3.3 (Strong QMA amplification [25])**.** *For any* $r = \text{poly}(|x|)$*,* $\mathsf{QMA}_w[2/3, 1/3] = \mathsf{QMA}_w[1-2^{-r}, 2^{-r}]$*.*

Definition 3.4 (k-Local Hamiltonian problem)**.** ***Input****:* $H_1, H_2, \ldots, H_T$ *set of* $T = \text{poly}(n)$ *Hermitian matrices with bounded spectral norm* $\|H_i\| \leq 1$ *acting on the Hilbert space of* n *qubits. In addition, each term acts nontrivially on at most* k *qubits and is described by* $\text{poly}(n)$ *bits. Furthermore, we are given two real numbers* a, b *(described by* $\text{poly}(n)$ *bits) such that* $b - a > 1/\text{poly}(n)$*.* ***Output****: Promised either the smallest eigenvalue of* $H = H_1 + H_2 + \ldots H_T$ *is smaller than* a *or all eigenvalues are larger than* b*, decide which case it is. We denote this problem by* k-$\mathsf{LH}[a, b]$*.*

The k-LH is in QMA for any $k = O(\log n)$ (see e.g., Theorem 1 in [13]). Furthermore, Kitaev showed in his seminal work [1] that 5-LH is QMA-complete.

Theorem 3.5 (Kitaev [1]). *Any* $\mathsf{QMA}_w[c,s]$ *protocol involving an* n*-qubit verifier circuit with* $T = \mathrm{poly}(n)$ *gates can be turned into a* 5-$\mathsf{LH}[a,b]$ *on* $\mathrm{poly}(n)$ *qubits with* $a = O((1-c)/T)$ *and* $b = \Omega((1-\sqrt{s})/T^3)$.

We will often simply write QMA, LH when the parameters are unimportant or clear from context. Next, we need the following lemmas in this work.

Lemma 3.6 (Detectability lemma [26]). *Let* $\{Q_1, \ldots, Q_m\}$ *be a set of projectors and* $H = \sum_{i=1}^m Q_i$. *Assume that each* Q_i *commutes with all but* g *others. Given a state* $|\psi\rangle$, *define* $|\phi\rangle := \prod_{i=1}^m (\mathrm{I} - Q_i)|\psi\rangle$, *where the product is taken in any order, and let* $e_\phi = \langle\phi| H |\phi\rangle / \|\phi\|^2$. *Then*

$$\|\phi\|^2 \leq \frac{1}{e_\phi/g^2 + 1}.$$

Lemma 3.7 (Quantum union bound [27]). *Consider the same setting as in Lemma 3.6, but this time we do* not *require each* Q_i *to commute with at most* g *others. It holds that*

$$\|\phi\|^2 \geq 1 - 4\langle\psi| H |\psi\rangle .$$

Lemma 3.8 (Jordan's lemma [28]). *Given two projectors* Π_1, Π_2 *acting on a* d*-dimensional complex vector space* $\mathcal{H}$, *there exists a change of basis such that* $\mathcal{H}$ *is decomposed as a direct sum of one- or two-dimensional mutually orthogonal subspaces* $\mathcal{H} = \bigoplus_i \mathcal{H}_i$, *such that both the projectors leave the subspaces invariant. In other words, we can write* $\Pi_1 = \sum_i a_i |u_i\rangle\langle u_i|$ *and* $\Pi_2 = \sum_i b_i |v_i\rangle\langle v_i|$, *with* $|u_i\rangle, |v_i\rangle \in \mathcal{H}_i$ *and* $a_i, b_i \in \{0,1\}$.

3.2 Fault tolerance

When defining our model in Section 2, we introduced perturbations to make the tensor network injective. This ensures the existence of an associated Hamiltonian and avoids the model from becoming too powerful [15]. Remarkably, the perturbations can be interpreted as undesired Pauli errors, see Claim 2.1. These 'errors' disturb our computation, leading to a degradation of the output, just as they would in a physical device. We can remedy this problem by substituting the circuit W with a fault-tolerant version of itself $\tilde{W}$, thereby guaranteeing robustness against the errors. In this section we will briefly summarize some results of fault-tolerant quantum computing that we require for our construction.

3.2.1 Quantum error correcting codes

Quantum error correcting codes are subspaces of the full Hilbert space of n bits. Each quantum code has three parameters: The number of logical qubits k tells us that the code protects a state vector of k qubits. The number of physical qubits n refers to the number of qubits into which the k logical qubits are being encoded. Finally, the distance d refers to the minimum number of single-qubit Pauli errors that are needed to map one encoded state onto another. In particular, a quantum code of distance d can correct any error acting on less than $d/2$ of the physical qubits. We will not review constructions and error correction procedures of different quantum codes, as we do not explicitly use them, and refer to Ref. [29] for details.

We note that the existence of quantum codes does not guarantee that quantum computing can be made robust against noise. Manipulating the encoded states via an error prone process leads to errors spreading and it is this spread of errors that needs to be controlled.

3.2.2 Quantum fault tolerance

In a seminal result, Shor showed that when any component of a quantum circuit, such as state preparation, gates and measurements, is replaced by a fault-tolerant version, it is possible to reduce errors under the assumption that the error rate per time step is polylogarithmically small in the length of the computation. Aharonov and Ben-Or [20] and Knill, Laflamme and Zurek [30] extended Shor's approach with a concatenation scheme. The main idea is as follows: At the top level each qubit is encoded into a quantum code $\mathcal{C}_1$ using n_1 physical qubits and with distance d_1. Next, each physical qubit of $\mathcal{C}_1$ is encoded further into a second code $\mathcal{C}_2$ using n_2 physical qubits and with distance d_2. This way, we have effectively a new code using $n_1 n_2$ physical qubits and which has distance $d_1 d_2$. Assuming that $\mathcal{C}_1$ and $\mathcal{C}_2$ come with a fault-tolerant set of circuit components, so does the concatenated code. Crucially, taking $\mathcal{C}_1 = \mathcal{C}_2$ the failure probability of any circuit component in the top level is now bounded by $c\left(cp^2\right)^2 = c^3 p^4$. Continuing this process, if we concatenate the same code a times, the probability of failure of any top level component is bounded by $c^{-1}(cp)^{2^a}$. Let s^a be the the circuit size at the ath level of concatenation. While the size of the circuit grows exponentially, the error is reduced double exponentially. Hence, fixing some desired error rate $\epsilon = c^{-1}(cp)^{2^a}$ leads to $s^a = \Theta\left(\text{polylog}\left(\frac{1}{\epsilon}\right)\right)$. In summary, concatenation allows us to simulate a quantum circuit with component failure rate bounded by an arbitrarily small ϵ using components with error rate bounded by some constant error rate p, as long as the initial error rate is below a threshold value set by the combinatorial factor c.

Theorem 3.9 ([20], Theorem 12)**.** *There exists a noise threshold $\eta_c > 0$ such that for any $\eta < \eta_c$, $\varepsilon > 0$ the following holds. For any n-qubit quantum circuit C with s gates, ℓ locations, and depth D, there exists a quantum circuit $\tilde{C}$ of size $s\,\text{polylog}(\ell/\varepsilon)$ (no measurements or classical operations are required) and depth $D\,\text{polylog}(\ell/\varepsilon)$ operating on $n\,\text{polylog}(\ell/\varepsilon)$ qubits such that in the presence of local depolarizing noise with error rate $\eta < \eta_c$, the encoded output of $\tilde{C}$ is ε-close to that of C.*

The theorem above does assume all-to-all connectivity, i.e. gates can be applied on arbitrary sets of qubits. We can also constrain the circuit to only operate locally on a d-dimensional grid of qubits, so that two qubit gates are only applied between neighbours on the grid. Note that an arbitrary circuit can be turned into a d-dimensional circuit by introducing SWAP gates and ancilla qubits, leading to the following result for any $d \geq 1$.

Corollary 3.10 ([20], Theorem 13)**.** *There exists a noise threshold $\eta_c > 0$ such that for any $\eta < \eta_c$, $\varepsilon > 0$, and $d \geq 1$ the following holds. For any d-dimensional n-qubit quantum circuit C with s gates, ℓ locations, and depth D, there exists a d-dimensional quantum circuit $\tilde{C}$ of size $s\,\text{polylog}(\ell/\varepsilon)$ (no measurements or classical operations are required) and depth $D\,\text{polylog}(\ell/\varepsilon)$ operating on $n\,\text{polylog}(\ell/\varepsilon)$ qubits such that in the presence of local depolarizing noise with error rate $\eta < \eta_c$, the encoded output of $\tilde{C}$ is ε-close to that of C.*

4 Soundness of the parent Hamiltonian

Our construction naturally gives rise to a mapping from a circuit to a Hamiltonian by considering the parent Hamiltonian of the injective tensor network, which we analyze in this section.

4.1 Adversarially noisy states

In the remainder of this section, we investigate the properties of combinatorial or low-energy states of $H_{\text{parent}} = H_{\text{in}} + H_{\text{prop}}$ (defined in Definition 2.2). Informally, local terms in H_{parent} that are violated by these states can be converted into errors in the quantum circuit. These errors, however, are *adversarial* in the sense that the faulty locations are chosen arbitrarily by the adversary. Informally, the violated terms in H_{in} correspond to errors at a set of locations S_0 in the initialization step, the violated terms in the first layer of H_{prop} correspond to gate errors at locations S_1 in the circuit's first layer, and so on. For this, let us define the notion of adversarially noisy states.

Definition 4.1. *Suppose $S = \{S_0, \ldots, S_D\}$, where $S_\ell \subseteq [n]$ for $0 \le \ell \le D$, is a set of locations in a depth-D n-qubit circuit. We define* $\mathrm{err}(S) = \{\vec{E} \in \mathcal{P}^{\otimes n(D+1)} : \mathrm{loc}(\tilde{E}_\ell) \subseteq S_\ell,\ 0 \le \ell \le D\}$ *to be the set of Pauli errors supported within the set of locations S.*

Definition 4.2 (Noisy states)**.** *For any sets of locations $S_\ell \subseteq [n]$, for $0 \le \ell \le D$, a pure state $|\psi\rangle$ is said to be an* adversarially noisy *state at locations $S = \{S_0, S_1, \ldots, S_D\}$ if*

$$|\psi\rangle \in \mathrm{adv}(W, S) \triangleq \mathrm{span}\{\tilde{E}_D W_D \ldots \tilde{E}_1 W_1 \tilde{E}_0 |0^a\rangle |\xi\rangle : \forall |\xi\rangle, \vec{E} \in \mathrm{err}(S)\}.$$

We consider noisy states such that at most εn adversarial errors are present in the circuit. In particular, we say a pure state $|\psi\rangle$ is an ε-noisy state *if*

$$|\psi\rangle \in \mathrm{adv}_\varepsilon(W) \triangleq \mathrm{span}\{\mathrm{adv}(W, S) : \sum_{\ell=0}^{D} |S_\ell| \le \varepsilon n\}.$$

A mixed state ρ is ε-noisy if it is a convex combination of ε-noisy pure states.

Our main theorems are the following soundness results.

Theorem 4.3 (Soundness)**.** *Suppose the depth $D = o(\log n)$ and consider any injectivity parameter $\delta = O(D^{-0.51})$. For any state $|\psi\rangle$ with energy density $\frac{\delta^{200D}}{D+1}$ with respect to H_{parent}, the reduced ψ_{out} in the output column is $\frac{1}{10}$-close in trace distance to a $400\delta^2 D$-noisy mixed state.*

We also prove a "combinatorial" version.

Theorem 4.4 (Combinatorial soundness)**.** *There exists a constant ε_0 such that the following holds. Consider any injectivity parameter $\delta = O(D^{-0.51})$ and any $10\delta\sqrt{D} < \varepsilon < \varepsilon_0$. Then for any state $|\psi\rangle$ that satisfies all but $\frac{\varepsilon}{D+1}$ fraction of terms in H_{parent}, the reduced state ψ_{out} in the output column is e^{-99n}-close in infidelity to an 8ε-noisy mixed state.*

Remark 4.5. *The theorem statements and proofs below are presented assuming all n qubits are intialized at the beginning of the computation for simplicity. However, they can be readily adapted to the setting where qubits are initialized at varying times such as in quantum fault tolerance. In this case, D is defined to be the longest elapse time between an output qubit and the initialization of any qubit causally connected to it.*

4.2 Proof of Theorem 4.4 (Combinatorial soundness)

Consider a $\frac{\varepsilon}{D+1}$-combinatorial state $|\psi\rangle$ and let $S = \{S_0, S_1, \ldots, S_D\}$ be the sets of faulty locations in each layer of the circuit corresponding to the violated Hamiltonian terms in $H_{\mathrm{parent}} = H_{\mathrm{in}} + H_{\mathrm{prop}}$. Since H_{prop} consists of at most nD terms and H_{in} consists of $a \le n$ terms, it holds that $\sum_\ell |S_\ell| \le 2\varepsilon n$ (assuming circuit consists of two-qubit gates). Below, we refer to the last column of qubits in the tensor network as the *output column*, the first two columns as the *first layer*, the next two columns as the *second layer*, and so on. Given a $n(2D+1)$-qubit PEPS state $|\psi\rangle$, we denote by $\psi_j^{(\ell)}$ the two-qubit reduced state on the j-th row of the ℓ-th layer and by ψ_j^{out} the one-qubit reduced state on the j-th row of the output column.

Proof idea: The combinatorial state $|\psi\rangle$ has the property that the (unnormalized) state $\Lambda^{\otimes nD} |\psi\rangle$ has a nice form - $\left(\bigotimes_{i \notin S_0} |0\rangle_i\right) \left(\bigotimes_{\mathrm{loc}(U) \notin S} |\Phi_U\rangle\right) \otimes |\psi''\rangle$. This means that we have the correct state $|0\rangle$ or $|\Phi_U\rangle$ corresponding to the satisfied Hamiltonian terms and an arbitrary state $|\psi''\rangle$ at the violated terms. If $|\psi''\rangle$ were of the form $\bigotimes_{j \in S} |\Phi_{U_j}\rangle$ for some 2-qubit unitaries U_j, then we could simply view the state $|\psi\rangle$ as encoding the circuit with iid noise on non-faulty locations, and adversarial noise at faulty locations. This would be a perfectly fine combination of stochastic error and small number of adversarial errors. But $|\psi''\rangle$ can be a superposition of the states of above form, which can arbitrarily correlate the noise at non-faulty locations! We appeal to the injectivity of the local maps Λ to argue that despite this possible correlation of

noise at non-faulty locations, the fraction of errors stays at $O(\delta^2)$ (with high probability). Thus a damaging situation, for example all the non-faulty locations experiencing a Pauli error, continues to occurs with very small probability.

Proof: We first prove the following lemma which asserts that the reduced state on the output column of the combinatorial state $|\psi\rangle$ contains the result of a quantum computation with both stochastic noise *and* adversarial noise. Later we will combine these two noise models into just adversarial noise.

Claim 4.6. *Let $\widetilde{W}_{\vec{P},\vec{E}} = \tilde{E}_D W_D \tilde{P}_D \dots \tilde{E}_1 W_1 \tilde{P}_1 \tilde{E}_0$ denote the erroneous circuit with Pauli errors $\vec{P}$ and $\vec{E}$. Suppose $|\psi\rangle$ is a normalized state which satisfies all but terms at locations S in H_{parent}. Then the reduced state on the output column is*

$$\psi_{\text{out}} \propto \sum_{\vec{P}\in\mathcal{P}^{\otimes nD}} \delta^{2|\vec{P}|} \left(\sum_{\vec{E}\in\text{err}(S)} c_{\vec{E}} \widetilde{W}_{\vec{P},\vec{E}} (|0^a\rangle \otimes |\xi_{\vec{E}}\rangle) \right) \left(\sum_{\vec{E}\in\text{err}(S)} c_{\vec{E}} (\langle 0^a| \otimes \langle \xi_{\vec{E}}|) \widetilde{W}^\dagger_{\vec{P},\vec{E}} \right), \tag{9}$$

where $|\xi_{\vec{E}}\rangle$ are normalized states and the real coefficients $c_{\vec{E}}$ satisfy $\sum_{\vec{E}\in\text{err}(S)} c_{\vec{E}}^2 = 1$. In other words, the state $|\psi\rangle$ encodes a noisy computation where the errors come from two sources: (1) stochastic noise coming from the tensor network injectivity and (2) adversarial errors coming from the energy violations.

Proof. Let us analyze the terms in each of H_{in} and H_{prop}. Consider the state $|\psi'\rangle \triangleq \frac{\Lambda^{\otimes nD}|\psi\rangle}{\|\Lambda^{\otimes nD}|\psi\rangle\|}$, where we recall that

$$\Lambda = \delta |\Phi_I\rangle\langle\Phi_I| + \sum_{p\in\{X,XZ,Z\}} |\Phi_p\rangle\langle\Phi_p|. \tag{10}$$

Consider a satisfied initialization term $h_j^{\text{in}} = \Lambda(|1\rangle\langle 1|_j)\Lambda$ (which acts on the j-th qubits of the first and second columns) in H_{in}. Since $h_j^{\text{in}}|\psi\rangle = 0$, the reduced state of $|\psi'\rangle$ on the first column's qubit i is $|0\rangle_i$. Similarly, for a satisfied propagation term $h_U = \Lambda^{\otimes 4}(\mathrm{I} - |\Phi_U\rangle\langle\Phi_U|)\Lambda^{\otimes 4}$ corresponding to a gate U acting on qubits i, j at time t, the reduced state of $|\psi'\rangle$ on the 4 qubits of $\text{loc}(U)$ must exactly be $|\Phi_U\rangle$. So it holds that

$$|\psi'\rangle = \left(\bigotimes_{i\notin S_0} |0\rangle_i \right) \left(\bigotimes_{\text{loc}(U)\notin S} |\Phi_U\rangle \right) \otimes |\psi''\rangle, \tag{11}$$

where $|\psi''\rangle$ is an arbitrary state supported on the remaining qubits (corresponding to the faulty locations in the circuit, including initialization, and the arbitrary witness state). The state $|\psi''\rangle$ can be further expressed in the orthonormal bases $\{|0\rangle, X|0\rangle\}$ on the input qubits and $\{|\Phi_U\rangle, \mathrm{I}\otimes X|\Phi_U\rangle, \mathrm{I}\otimes XZ|\Phi_U\rangle, \mathrm{I}\otimes Z|\Phi_U\rangle\}$ on the qubit pairs in the "bulk", such that

$$|\psi'\rangle = \sum_{\vec{E}\in\text{err}(S)} c_{\vec{E}} \left(\tilde{E}_0 |0^a\rangle |\xi_{\vec{E}}\rangle \right) \bigotimes_{\ell\in D} \left((\mathrm{I}\otimes\tilde{E}_\ell) \bigotimes_{U\in W_\ell} |\Phi_U\rangle \right), \tag{12}$$

where $|\xi_{\vec{E}}\rangle$ are normalized states and the coefficients $c_{\vec{E}}$ are real (w.l.o.g) and satisfy $\sum_{\vec{E}\in\text{err}(S)} c_{\vec{E}}^2 = 1$. Note that $c_{\vec{E}}$ is nonzero only when $\tilde{E}_0$ consists of only the Pauli operators I and X.

Next, we undo the maps Λ to obtain the original combinatorial state by applying the map $Q = |\Phi_I\rangle\langle\Phi_I| + \delta\sum_{p\in\{X,XZ,Z\}} |\Phi_p\rangle\langle\Phi_p| \propto \Lambda^{-1}$ (see Equation (5)) on $|\psi'\rangle$. We have that

$$|\psi\rangle \propto |\chi\rangle = Q^{\otimes nD}|\psi'\rangle \tag{13}$$

$$= \sum_{\vec{P}\in\mathcal{P}^{\otimes nD}} \delta^{|\vec{P}|} |\Phi_{\vec{P}}\rangle \bigotimes \left(\sum_{\vec{E}\in\text{err}(S)} \tilde{E}_D W_D \tilde{P}_D \dots \tilde{E}_1 W_1 \tilde{P}_1 \tilde{E}_0 |0^a\rangle |\xi_{\vec{E}}\rangle \right), \tag{14}$$

where the last equality follows from linearly extending Claim 2.1.

Tracing out the bulk EPR states $|\Phi_{\vec{P}}\rangle$ we obtain the statement of the claim. □

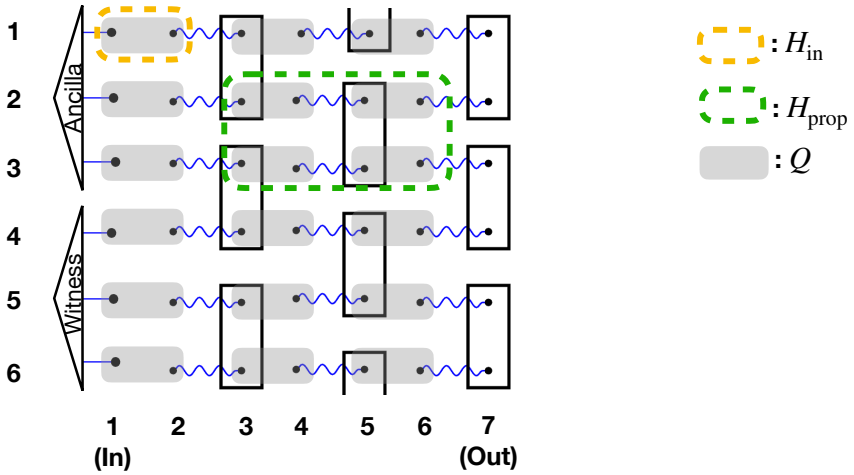


Figure 4: An injective PEPS encoding noisy quantum computation shown with $n = 6$ qubits (black dots), of which $a = 3$ are ancillas, and $D = 3$ layers of two-qubit gates in the brickwork architecture. The computation goes from left to right, with qubits on column 1 being the input. **Gates**: encoded in rotated 4-qubit EPR states (see Figure 2) placed on columns (2,3), (4,5), and so on. Applying the invertible map Q (gray box) as defined in Equation (5) generates a noisy computation on the last column (indexed 7). The qubit pairs where Q is applied are called *shifted* EPR locations. We refer to the last column of qubits in the PEPS as the *output column*. **Noisy computation**: After Q is applied, the output column can be interpreted as a noisy computation where for each layer of the circuit, the present noise pattern is specified by the EPR states at the shifted EPR locations. Due to this correspondence, we refer to the first two columns (indexed 1,2) as the *first layer*, the next two columns (indexed 3,4) as the *second layer*, and so on. **Parent Hamiltonian**: A propagation term (dashed green) acts on 8 qubits, while an initialization term (dashed yellow) acts on the first 2 qubits and only on each ancilla row (indexed 1,2,3).

We now show that the normalized state ψ_{out} is exponentially close to an $(\alpha + 2\varepsilon)$-noisy (mixed) state ρ which is obtained by removing from $|\chi\rangle$ the summands $\vec{P}$ whose weight is larger than αn and then normalizing properly, for some constant α to be specified shortly.

Claim 4.7. *Let Π be the projector onto the hight-weight EPR-basis states in the bulk which contains at least αn nontrivial EPR states*

$$\Pi = \sum_{\vec{P} \in \mathcal{P}^{\otimes nD} : |\vec{P}| \geq \alpha n} \left|\Phi_{\vec{P}}\right\rangle \left\langle \Phi_{\vec{P}}\right| . \tag{15}$$

If $\delta = O(D^{-0.51})$, then choosing $\alpha = 6\varepsilon$ we can get the following bound as long as $\varepsilon = \Omega(D^{-0.01})$

$$\langle\psi| \Pi |\psi\rangle \leq e^{-\Omega(n)}. \tag{16}$$

In other words, $|\psi\rangle$ is $e^{-\Omega(n)}$-close in fidelity to a 8ε-noisy mixed state.

Proof. Note that the distribution over $\vec{P}$ is not simply the i.i.d. distribution $\delta^{|\vec{P}|}$ since the linear combination over the adversarial error $\vec{E}$ in Equation (14) can change the norm of the state in the output column. So we need a more careful analysis.

With a slight abuse of notation, we use $\{S_1, S_2, \ldots, S_D\}$ to denote the locations of the *non-shifted* EPR states (which encode the gates) and S_0 to denote input column qubits (which are ancilla qubits) that correspond to the violated Hamiltonian terms (see Figure 4). Recall that $|S| \leq 2\varepsilon n$ by assumption. Let R^c be the *shifted* EPR locations that do *not* overlap with S and let R be the rest of shifted EPR locations. Note that $|R| \leq 2|S| \leq 4\varepsilon n$ and $|R| + |R^c| = nD$. Consider the "partially undone" state $|\chi'\rangle = Q_{R^c} |\psi'\rangle$, in which we only apply Q on R^c, such that $|\chi\rangle = Q_R |\chi'\rangle$. Let Π' be the projector onto the high-weight (shifted) EPR-basis states in R^c defined as

$$\Pi' = \sum_{\vec{P} \in \mathcal{P}^{R^c} : |\vec{P}| \geq (\alpha - 4\varepsilon)n} \left|\Phi_{\vec{P}}\right\rangle \left\langle \Phi_{\vec{P}}\right| . \tag{17}$$

Note that $\Pi \preceq \Pi'$ and $\delta \cdot \mathrm{I} \preceq Q \preceq \mathrm{I}$. Furthermore, Π' and Q (and Λ) commute for being both diagonal in the EPR basis. So we have that

$$\langle\psi| \Pi |\psi\rangle \leq \langle\psi| \Pi' |\psi\rangle = \frac{\langle\chi'| Q_R \Pi' Q_R |\chi'\rangle}{\langle\chi'| Q_R Q_R |\chi'\rangle} \leq \frac{1}{\delta^{2|R|}} \frac{\langle\chi'| \Pi' |\chi'\rangle}{\langle\chi'|\chi'\rangle}. \tag{18}$$

Substituting $|\chi'\rangle = Q_{R^c} |\psi'\rangle$ into the RHS we get

$$\langle\psi| \Pi |\psi\rangle \leq \frac{1}{\delta^{2|R|}} \frac{\sum_{\vec{P} \in \mathcal{P}^{R^c} : |\vec{P}| \geq (\alpha - 4\varepsilon)n} \delta^{2|\vec{P}|} \| (\mathrm{I}_R \otimes \left\langle \Phi_{\vec{P}}\right|) |\psi'\rangle \|^2}{\sum_{\vec{P} \in \mathcal{P}^{R^c}} \delta^{2|\vec{P}|} \| (\mathrm{I}_R \otimes \left\langle \Phi_{\vec{P}}\right|) |\psi'\rangle \|^2}, \tag{19}$$

where $|\psi'\rangle$ is defined in Equation (11) and repeated here for convenience

$$|\psi'\rangle = \left(\bigotimes_{i \notin S_0} |0\rangle_i \right) \left(\bigotimes_{\mathrm{loc}(U) \notin S} |\Phi_U\rangle \right) \otimes |\psi''\rangle . \tag{20}$$

However, observe that $(\mathrm{I}_R \otimes \left\langle \Phi_{\vec{P}}\right|) |\psi'\rangle$ is a unit vector since $\left\langle \Phi_{\vec{P}}\right|$ does not act on $|\psi''\rangle$.

Therefore,

$$\langle\psi| \Pi |\psi\rangle \leq \frac{1}{\delta^{2|R|}} \frac{\sum_{\vec{P} \in \mathcal{P}^{R^c} : |\vec{P}| \geq (\alpha - 4\varepsilon)n} \delta^{2|\vec{P}|}}{\sum_{\vec{P} \in \mathcal{P}^{R^c}} \delta^{2|\vec{P}|}}. \tag{21}$$

The RHS can now be straightforwardly bounded by the Chernoff bound.

Fact 4.8. *Let $X = X_1 + \ldots + X_N$ where $X_i \in \{0, 1\}$ are i.i.d. binary random variables with $\mathbb{E}[X_i] = \mu$. Then $\Pr[X \geq (1 + \eta)\mu N] \leq 2e^{-\eta^2 \mu N/3}$.*

We apply the Chernoff's bound with $N = |R^c|$, $\mu = 3\delta^2/(1+3\delta^2)$, and $\eta = (\alpha - 4\varepsilon)/\mu D - 1$. Note that $|R| \leq 4\varepsilon n$. Choosing $\alpha = 6\varepsilon$, assuming $\varepsilon \ll 1$ so that $nD/2 \leq |R^c| \leq nD$ and δ is sufficiently small such that $\eta \approx 2\varepsilon/\mu D$, we obtain the following bound on the RHS of Equation (21)

$$\text{RHS of Equation (21)} \leq 2e^{8\log(1/\delta)\varepsilon n} e^{-\varepsilon^2 n/\delta^2 D}. \tag{22}$$

The above bound to decays exponentially when $\varepsilon > 8\log(1/\delta)\delta^2 D$. We can choose, say, $\delta = O(D^{-0.51})$. Then for any $\varepsilon > 10\delta\sqrt{D}$, we obtain a bound of e^{-99n} on $\langle\psi|\,\Pi\,|\psi\rangle$. □

4.3 Proof of Theorem 4.3 (Soundness)

We assume depth $D = o(\log n)$ here.

Proof idea: We take inspiration from Kitaev's analysis where the clock Hamiltonian is analyzed in a suitable rotated basis. Here as well, we will carry out the proof in a "rotated" basis, which is defined by the $n(2D+1)$-qubit unitary V in Equation (8). In particular, we analyze the properties of a low-energy state $|\psi\rangle$ of $H_{\text{parent}} = H_{\text{in}} + H_{\text{prop}}$ and its rotated version $|\psi'\rangle = V^\dagger|\psi\rangle$. Note that in the rotated basis, the ground states (Claim 2.1) are of the form

$$V^\dagger|\Psi_{W,\xi}\rangle \propto \left(|\Phi_I\rangle + \delta \sum_{p\in\{X,XZ,Z\}} |\Phi_p\rangle\right)^{\otimes nD} |0^a\rangle\,|\xi\rangle\,. \tag{23}$$

Our starting observation is based on a surprising effect - despite the fact that H_{in} enforces $|0\rangle^{\otimes n}$ on the first column, the state $|0\rangle^{\otimes n}$ appears on the last column in $V^\dagger|\Psi_{W,\xi}\rangle$. We view this as a *teleportation of* H_{in}, highlighting that its a noiseless teleportation under 'zero energy' constraint, despite the tensor network performing noisy gate-by-gate teleportation. Given this, we focus on establishing two properties for low energy states:

- *Robust teleportation of* H_{in}*:* Upon rotating with V, the low energy states should look like $|0\rangle$ in most of the qubits (that do not include witness qubits) in the last column. This amounts to H_{in} effectively acting on the last column under the constraint of low energy.
- The number of Pauli errors is small enough in a low energy state.

The proof below carries both these properties.

Proof: We refer to the last column of qubits in the PEPS as the *output column* and note that the layers of shifted EPR locations have a correspondence with circuit layers. In particular, the first two columns as the *first layer*, the next two columns as the *second layer*, and so on (see Figure 4). The unitary V can be interpreted as applying a noisy circuit on the output column conditioned on the noise pattern in the bulk.

Definition 4.9. *Given a $n(2D+1)$-qubit PEPS state $|\psi\rangle$, we denote by $\psi_j^{(\ell)}$ the two-qubit reduced state on the j-th row of the ℓ-th layer and by ψ_j^{out} the one-qubit reduced state on the j-th row of the output column.*

The advantage of working in the rotated basis is that we can employ the following lemmas, whose proofs are provided in Section 4.4.

Remark 4.10. *W.l.o.g., we assume the last layer of the circuit consists of single-qubit identity gates.*

Lemma 4.11 (Last layer)**.** *For each $j \in [n]$, let $h_j^{(D)}$ be the 3-qubit term in H_{prop} corresponding to the identity gate on qubit j. Furthermore, let $|\phi_0\rangle = \frac{1}{\sqrt{1+3\delta^2}}(|\Phi_I\rangle + \delta\sum_{p\in\{X,XZ,Z\}}|\Phi_p\rangle)$. If $\langle\psi|\,h_j^{(D)}\,|\psi\rangle \leq \alpha$, then it holds that*

$$\text{Tr}\left(\psi_j^{\prime(D)}\,|\phi_0\rangle\langle\phi_0|\right) \geq 1 - 4\alpha. \tag{24}$$

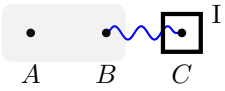


Figure 5: Propagation term $h_j^{(D)} = \Lambda_{AB}(\mathrm{I} - |\Phi_I\rangle\langle\Phi_I|_{BC})\Lambda_{AB}$ corresponding to a single-qubit identity gate in the last layer. If a global state $|\psi\rangle$ has low energy with restpect to $h_j^{(D)}$, then Lemma 4.11 asserts that $V^\dagger|\psi\rangle$ is locally close to $|\phi_0\rangle$ (Equation (26)) on qubits A, B.

In the proof of Lemma 4.11, we will show that $V^\dagger h_j^{(D)} V$ is in fact a local Hamiltonian term, despite V being global. In particular, denoting $h_j^{(D)} = \Lambda_{AB}(\mathrm{I} - |\Phi_I\rangle\langle\Phi_I|_{BC})\Lambda_{AB}$ where A, B, and C denote the qubits acted upon by $h_j^{(D)}$ (see Figure 5), then $V^\dagger h_j^{(D)} V$ is a 2-local term acting on qubits A, B in the rotated basis

$$V^\dagger (h_j^{(D)})_{ABC} V = \Lambda_{AB}\left(\sum_{p\in\mathcal{P}} |\Phi_p\rangle\langle\Phi_p|_{AB} - \frac{1}{4}\sum_{p,p'\in\mathcal{P}} |\Phi_p\rangle\langle\Phi_{p'}|_{AB}\right)\Lambda_{AB}. \tag{25}$$

The rotated propagation terms $V^\dagger h_{i,j}^{(\ell)} V$ for $\ell < D$ (corresponding to two-qubit gates acting on qubits i, j in layer ℓ) are, however, generally non-local[3]. Here, we will instead utilize the following lemma about a property of them in a certain subspace related to the state

$$|\phi_0\rangle \triangleq \frac{1}{\sqrt{1+3\delta^2}}(|\Phi_I\rangle + \delta \sum_{p\in\{X,XZ,Z\}} |\Phi_p\rangle). \tag{26}$$

Lemma 4.12 (Bulk propagation)**.** *Consider a propagation term $h_{i,j}^{(\ell)} = \Lambda^{\otimes 4}(\mathrm{I} - |\Phi_U\rangle\langle\Phi_U|)\Lambda^{\otimes 4}$, where $|\Phi_U\rangle$ is the EPR state encoding the two-qubit gate U acting on qubits i, j in layer $\ell < D$. It holds that*

$$\langle\phi_0|^{\otimes 2} V^\dagger h_{i,j}^{(\ell)} V |\phi_0\rangle^{\otimes 2} = \frac{16\delta^4}{(1+3\delta^2)^2}\Lambda^{\otimes 2}\left(\sum_{\vec{p}\in\mathcal{P}^{\otimes 2}} |\Phi_{\vec{p}}\rangle\langle\Phi_{\vec{p}}| - \frac{1}{16}\sum_{\vec{p},\vec{q}\in\mathcal{P}^{\otimes 2}} |\Phi_{\vec{p}}\rangle\langle\Phi_{\vec{q}}|\right), \tag{27}$$

where $|\phi_0\rangle^{\otimes 2}$ acts on the shifted EPR locations (i, j) in EPR layer $\ell + 1$. Furthermore, a robust version of the previous statement also holds. Let $|\psi\rangle$ be a normalized state such that $\mathrm{Tr}\left(\psi_{i,j}'^{(\ell+1)}\phi_0^{\otimes 2}\right) \geq 1 - \eta$ in for some $\ell \leq D - 1$. If additionally $\langle\psi| h_{i,j}^{(\ell)} |\psi\rangle \leq \alpha$, then $\mathrm{Tr}\left(\psi_{i,j}'^{(\ell,\ell+1)}\phi_0^{\otimes 4}\right) \geq 1 - \frac{\eta}{\delta^8} - \frac{\alpha}{\delta^{16}}$.

Intuitively, the above lemma says that, if the qubits i, j in layer $\ell + 1$ of a slighly violated propagation Hamiltonian term are in the "good" state ϕ_0, then the qubits i, j in layer ℓ are also in the good state ϕ_0.

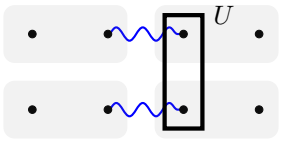


Figure 6: Propagation term h_U corresponding to a two-qubit gate U in the bulk of the circuit. According to Lemma 4.12, if a global state $|\psi\rangle$ has low energy with respect to h_U and its rotated version $V^\dagger|\psi\rangle$ is locally close to $|\phi_0\rangle^{\otimes 2}$ on the EPR locations to the right, then $V^\dagger|\psi\rangle$ is close to $|\phi_0\rangle^{\otimes 4}$ on all 4 EPR locations.

[3]We show in Appendix B that they are local if the associated gate is Clifford.

Lemma 4.13 (Robust teleportation of H_{in})**.** *Consider an initialization term $h_j^{\text{in}} = \Lambda\Pi_j\Lambda$ in H_{in}, where Π_j is the input check on qubit j. It holds that $\langle\phi_0| V^\dagger h_j^{\text{in}} V |\phi_0\rangle = \frac{4\delta^2}{1+3\delta^2}\Pi_j^{\text{out}}$, where $|\phi_0\rangle$ acts on the first layer at EPR location j, and Π_j^{out} means Π_j acts on the output column. Furthermore, a robust version of the previous statement also holds. Let $|\psi'\rangle$ be a normalized state such that $\text{Tr}\left(\psi_j'^{(1)} |\phi_0\rangle\langle\phi_0|\right) \geq 1-\eta$. If additionally $\langle\psi'| V^\dagger h_j^{\text{in}} V |'\psi\rangle \leq \alpha$, then $\text{Tr}\left(\psi'\Pi_j^{\text{out}}\right) \geq 1 - \frac{\eta}{\delta^2} - \frac{\alpha}{\delta^2}$.*

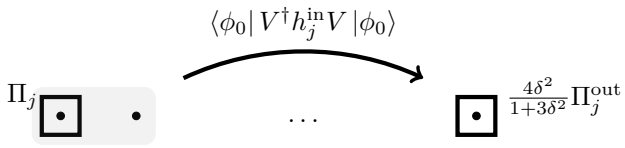


Figure 7: Initialization term $h_j^{\text{in}} = \Lambda\Pi_j\Lambda$ in H_{in} is teleported to the output column according to Lemma 4.13.

Therefore, if the first layer is in the good state $|\phi_0\rangle$, then H_{in} is teleported to the output column.

Proof of Theorem 4.3. Consider a state $|\psi\rangle$ with energy density $\frac{\varepsilon}{D+1}$ such that $\langle\psi| H_{\text{parent}} |\psi\rangle \leq \varepsilon n$, where $\varepsilon = \delta^{200D}$. It follows that at most $\varepsilon n/\alpha$ terms in H_{parent} have energy greater than α, for some value α to be specified later (we will choose $\alpha = \delta^{100D}$). We refer to these terms as "slightly violated" (as opposed to strongly violated). Let $|\psi'\rangle = V^\dagger |\psi\rangle$ be the rotated state.

The proof will proceed as follows. First, according to Lemma 4.11 $\psi_j'^{(D)}$ is close to $|\phi_0\rangle$ for many indices $j \in [n]$. Next, we repeatedly apply Lemma 4.12 to propagate the "good" states $|\phi_0\rangle$ to the first layer. Then we use Lemma 4.13 to conclude that most of intialization terms in H_{in} get teleported (approximately) to the output column. This makes sure that most of the ancilla qubits are initialized (approximately) correctly to $|0\rangle$. Finally, we use a similar argument to Theorem 4.4's proof.

We start by looking at the local reduced states on the last layer. For at least $(1-\varepsilon/\alpha)n$ many indices j, the propagation term $h_j^{(D)}$ in the last layer is slightly violated. So we invoke Lemma 4.11 to obtain $\text{Tr}\left(\psi_j'^{(D)} |\phi_0\rangle\langle\phi_0|\right) \geq 1 - 4\alpha$.

We now "propagate" these good states to the first layer. According to Lemma 4.12, a sufficient condition for $\psi_j'^{(1)}$ to be good is that all the Hamiltonian terms associated to gates in the forward lightcone of qubit j, have energy bounded by α. We denote the forward lightcone of qubit j by $\text{LC}(j)$. Note that we only consider $D = o(\log n)$. Assume all propagation terms in $\text{LC}(j)$ are slightly violated, then for any locations $r, s \in \text{LC}(j)$ in the last layer we have $\text{Tr}\left(\psi_{r,s}'^{(D)}(|\phi_0\rangle\langle\phi_0|^{\otimes 2})\right) \geq 1 - 8\alpha$ due to the previous paragraph. Next, we repeatedly apply Lemma 4.12 on the propagation terms in $\text{LC}(j)$ to obtain that $\text{Tr}\left(\psi_j'^{(1)} |\phi_0\rangle\langle\phi_0|\right) \geq 1 - \frac{\alpha}{\delta^{16D}}$ for sufficiently small δ. Thus, we can invoke the robust version of Lemma 4.13 to obtain $\text{Tr}\left(\psi'\Pi_j^{\text{out}}\right) \geq 1 - \frac{\alpha}{\delta^{16D+2}}$.

The number of locations whose forward lightcone is "bad" (i.e., it contains a strongly violated Hamiltonian term) is bounded above by $2^{D-1}\varepsilon n/\alpha$. Hence, $\text{Tr}\left(\psi'\Pi_j^{\text{out}}\right) \geq 1 - \frac{\alpha}{\delta^{16D+2}}$ for at least a fraction of $1 - 2^{D-1}\varepsilon/\alpha$ of the qubits j. We refer to the initialization locations without this guarantee as "strongly faulty" initialization locations. Denote these locations as S_0, we have $|S_0| \leq 2^{D-1}\varepsilon n/\alpha$. Similarly, there are at most $2^{D-1}\varepsilon n/\alpha$ EPR locations where we do not have the guarantee $\text{Tr}\left(\psi_j'^{(\ell)} |\phi_0\rangle\langle\phi_0|\right) \geq 1 - \frac{\alpha}{\delta^{16D}}$. We refer to them as "strongly faulty" gate locations in the circuit and denote $S = \{S_1, \ldots, S_D\}$.

For each slightly faulty location j at layer $1 \leq \ell \leq D$, we have the following distance guarantee due to Fuchs–van de Graf inequality and by choosing, say, $\alpha = \delta^{50D}$

$$\|\psi_j'^{(\ell)} - |\phi_0\rangle\langle\phi_0|\|_1 \leq 2\sqrt{\frac{\alpha}{\delta^{16D}}} \leq \delta^{10D}. \tag{28}$$

It follows that

$$\left|\mathrm{Tr}\left(\psi_j^{'(\ell)} \left|\Phi_p\right\rangle \left\langle\Phi_p\right|\right) - \left|\left\langle\phi_0|\Phi_p\right\rangle\right|^2\right| \leq \delta^{10D}, \qquad p \in \mathcal{P} \tag{29}$$

In other words, the Pauli errors at the slightly violated locations approximately follows the depolarizing channel with probability $p = \delta^2/(1+3\delta^2)$ for each of X, Y, Z errors.

Similarly, we obtain the following bound at the (approximately) correctly initialized locations in the output column

$$\mathrm{Tr}\left(\psi_j^{'\mathrm{out}} \left|0\right\rangle \left\langle 0\right|\right) \geq 1 - \delta^{10D}, \tag{30}$$

With $\alpha = \delta^{50D}$ and $\varepsilon = \delta^{200D}$ we also have the following bound on the total number of strongly faulty locations

$$|S| = \sum_{\ell=0}^{D} |S_\ell| \leq 2^D \varepsilon n/\alpha \leq \delta^{50D} n. \tag{31}$$

Similar to the proof of Theorem 4.4, we denote by $\tilde{E}_\ell$ the adversarial errors at locations S_ℓ coming from the strongly faulty locations and by $\tilde{P}_\ell$ be the *almost* local depolarizing noise coming from the slightly faulty locations. We can expand $|\psi'\rangle$ as

$$|\psi'\rangle = \sum_{\substack{\vec{E}\in\mathcal{P}^S \\ \vec{P}\in\mathcal{P}^{S^c}}} c_{\vec{E},\vec{P}} \bigotimes_{1\leq\ell\leq D} \left(\left|\Phi_{\tilde{P}_\ell}\right\rangle \left|\Phi_{\tilde{E}_\ell}\right\rangle\right) \bigotimes (\tilde{E}_0 \otimes \tilde{P}_0 \left|0^a\right\rangle) \otimes \left|\xi_{\vec{P},\vec{E}}\right\rangle, \tag{32}$$

where $\left|\xi_{\vec{P},\vec{E}}\right\rangle$ are normalized states and $c_{\vec{E},\vec{P}}$ are (w.l.o.g.) real coefficients such that $|c_{\vec{E},\vec{P}}|^2$ define a probability distribution whose local marginals on S^c are constrained by Equations (29), (30).

Finally, we can combine $\vec{P}$ and $\vec{E}$ together and treat them as adversarial errors by truncating the summands with high-weight $\vec{P}$. Let Π be the projector onto high-weight EPR states in S^c

$$\Pi = \sum_{\vec{P}\in\mathcal{P}^{S^c}:|\vec{P}|\geq(\beta-\delta^{50D})n} \left|\Phi_{\vec{P}}\right\rangle \left\langle\Phi_{\vec{P}}\right|, \tag{33}$$

where β is a parameter to be specified. Below, we will truncate the high-weight $\vec{P}$ in $|\psi'\rangle$ to obtain the state $|\chi'\rangle = \frac{(\mathrm{I}-\Pi)|\psi\rangle}{\|(\mathrm{I}-\Pi)|\psi\rangle\|}$, and our goal is to show $|\chi'\rangle$ is close to $|\psi'\rangle$. Note that $|\chi'\rangle$ only contains terms with at most βn adversarial errors as desired.

Observe that $\mathbb{E}_{\vec{P}\sim|\psi'\rangle}[|\vec{P}| : \vec{P} \in \mathcal{P}^{S^c}] \leq 3(\frac{\delta^2}{1+3\delta^2} + \delta^{10D})nD$ according to Equations (29), (30). So using Markov's inequality we can bound

$$\mathrm{Tr}(\Pi \left|\psi'\right\rangle \left\langle\psi'\right|) \leq \frac{4\delta^2 D}{\beta}. \tag{34}$$

Assuming $\delta^2 D \ll 1$ and chosing $\beta = 400\delta^2 D$ and using gentle measurement lemma we have

$$\frac{1}{2}\| \left|\chi'\right\rangle \left\langle\chi'\right| - \left|\psi'\right\rangle \left\langle\psi'\right| \|_1 \leq \frac{1}{10}. \tag{35}$$

The same trace distance bound holds on the unrotated states $V|\psi'\rangle$ and $V|\chi'\rangle \triangleq |\chi\rangle$, as well as their reduced states on the output column ψ_{out} and χ_{out} due to unitary-invariance and monoticity of the trace distance:

$$\frac{1}{2}\|\psi_{\mathrm{out}} - \chi_{\mathrm{out}}\| \leq \frac{1}{10}. \tag{36}$$

The reduced state χ_{out} is a β-noisy state

$$\chi_{\mathrm{out}} = \sum_{\tilde{E}_1 \ldots \tilde{E}_D : |\vec{E}|\leq\beta n} \left(\sum_{\tilde{E}_0 : |\vec{E}|\leq\beta n} c'_{\vec{E}} \widetilde{W}_{\vec{E}} (\tilde{E}_0 \left|0^a\right\rangle) \otimes \left|\xi_{\vec{E}}\right\rangle \right) (\ldots)^\dagger, \tag{37}$$

where $\widetilde{W}_{\vec{E}} \triangleq W_D \tilde{E}_D \ldots W_1 \tilde{E}_1$. This concludes the proof of Theorem 4.3. □

4.4 Analysis of H_{prop} and H_{in}: deferred proofs

For convenience, recall the change of basis

$$V = \sum_{\vec{P}\in\mathcal{P}^{\otimes nD}} |\Phi_{\vec{P}}\rangle\langle\Phi_{\vec{P}}| \otimes (W_D\tilde{P}_D \dots W_1\tilde{P}_1), \tag{38}$$

and the inverse local injective map

$$\Lambda = \delta\,|\Phi_I\rangle\langle\Phi_I| + \sum_{p\in\{X,XZ,Z\}} |\Phi_p\rangle\langle\Phi_p|\,. \tag{39}$$

4.4.1 Proof of Lemma 4.11 (Good states in last layer of H_{prop})

As stated in the lemma, we assume the last layer of gates in the circuit are identity gates for simplicity in calculating the rotated H_g terms. We have that

$$\begin{aligned}
V^\dagger(h_j^{(D)})_{ABC}V &= V^\dagger\Lambda_{AB}(\mathrm{I} - |\Phi_I\rangle\langle\Phi_I|_{BC})\Lambda_{AB}V \\
&= V^\dagger\left(\sum_{p,p'\in\mathcal{P}} \delta^{2-|(p,p')|}\,|\Phi_p\rangle\langle\Phi_{p'}|_{AB} \otimes \langle\Phi_p|_{AB}\,(\mathrm{I} - |\Phi_I\rangle\langle\Phi_I|_{BC})\,|\Phi_{p'}\rangle_{AB}\right)V \\
&= V^\dagger\left(\sum_{p,p'\in\mathcal{P}} \delta^{2-|(p,p')|}\,|\Phi_p\rangle\left(\mathbb{1}_{p,p'}\mathrm{I} - \frac{1}{4}(p^*p'^\top)_C\right)\right)V \\
&= \sum_{p\in\mathcal{P}} \delta^{2-2|p|}\,|\Phi_p\rangle\langle\Phi_p|_{AB} \\
&\quad - \frac{1}{4}\sum_{\vec{P}\in\mathcal{P}^{\otimes n(D-1)}}\sum_{p,p'\in\mathcal{P}} \delta^{2-|(p,p')|}\,|\Phi_p\rangle\langle\Phi_{p'}|_{AB} \otimes |\Phi_{\vec{P}}\rangle\langle\Phi_{\vec{P}}| \otimes \left((\widetilde{W}_{\vec{P}}^{\le D})^\dagger p^\dagger p^* p'^\top p'(\widetilde{W}_{\vec{P}}^{\le D})\right) \\
&= \sum_{p\in\mathcal{P}} \delta^{2-2|p|}\,|\Phi_p\rangle\langle\Phi_p|_{AB} - \frac{1}{4}\sum_{p,p'} \delta^{|(p,p')|}\,|\Phi_p\rangle\langle\Phi_{p'}|_{AB} \\
&= \Lambda_{AB}\left(\sum_{p} |\Phi_p\rangle\langle\Phi_p|_{AB} - \frac{1}{4}\sum_{p,p'} |\Phi_p\rangle\langle\Phi_{p'}|_{AB}\right)\Lambda_{AB},
\end{aligned}$$

where $\sum_{\vec{P}\in\mathcal{P}^{\otimes n(D-1)}}$ denotes the sum over the Pauli noise $\tilde{P}_\ell$ for $\ell < D$ and $\widetilde{W}_{\vec{P}}^{\le D} \triangleq W_{D-1}\tilde{P}_{D-1}\dots W_1\tilde{P}_1$.

Note that $V^\dagger h_j^{(D)} V$ has ground state $|\phi_0\rangle$ and spectral gap $\gamma \ge 1/4$, so

$$\begin{aligned}
\frac{1}{4}\operatorname{Tr}\left(\psi_j^{'(D)}(\mathrm{I} - |\phi_0\rangle\langle\phi_0|)\right) &\le \operatorname{Tr}\left(|\psi\rangle\langle\psi|\,h_j^{(D)}\right) \le \alpha, \\
\operatorname{Tr}\left(\psi_j^{'(D)}\,|\phi_0\rangle\langle\phi_0|\right) &\ge 1 - 4\alpha.
\end{aligned}$$

4.4.2 Proof of Lemma 4.12 (Bulk propagation of good states)

We first prove the following claim, which is the "noiseless" version of Lemma 4.12.

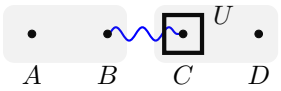
Claim 4.14. *Let $|\phi_0\rangle$ be defined as in Equation (26). Consider a propagation term h_U corresponding to a two-qubit gate as shown in Figure 6. Furthermore, define $|\Phi_{\vec{p}}\rangle \triangleq |\Phi_{p_1}\rangle|\Phi_{p_2}\rangle$, for $\vec{p}\in\mathcal{P}^{\otimes 2}$. It holds that*

$$\langle\phi_0|^{\otimes 2}\,V^\dagger h_U V\,|\phi_0\rangle^{\otimes 2} = \frac{16\delta^4}{(1+3\delta^2)^2}\Lambda^{\otimes 2}\left(\sum_{\vec{p}\in\mathcal{P}^{\otimes 2}} |\Phi_{\vec{p}}\rangle\langle\Phi_{\vec{p}}| - \frac{1}{16}\sum_{\vec{p},\vec{q}\in\mathcal{P}^{\otimes 2}} |\Phi_{\vec{p}}\rangle\langle\Phi_{\vec{q}}|\right)\Lambda^{\otimes 2}, \tag{40}$$

where $|\phi_0\rangle^{\otimes 2}$ ($|\Phi_{\vec{p}}\rangle$) acts on the EPR locations to the right (left) of h_U (see Figure 6).

The above claim implies that if on one side of h_U the state $V^\dagger |\psi\rangle$ is equal to $|\phi_0\rangle^{\otimes 2}$, then so is it on the other side since this is the unique ground state of the matrix to the RHS of Equation (40).

Proof of Claim. For simplicity, we prove the claim for terms corresponding to single-qubit gates. The generalization to the multi-qubit case is straightforward as explained later. Consider a gate U in layer $\ell < D$ acting on qubit j. We refer to the Hamiltonian term corresponding to this gate as $h_U = \Lambda^{\otimes 2}_{AB,CD}(\mathrm{I} - |\Phi_U\rangle\langle\Phi_U|_{BC})\Lambda^{\otimes 2}_{AB,CD}$, which acts on qubits A, B (EPR layer ℓ) and C, D (EPR layer $\ell+1$).



Let $|\Phi_{\vec{p}}\rangle_{AB,CD} \triangleq |\Phi_{p_1}\rangle_{AB} |\Phi_{p_2}\rangle_{CD}$, for $p_1, p_2 \in \mathcal{P}$. We will omit the system labels when they are clear from the context. The rotated term is of the form

$$\begin{aligned}
V^\dagger h_U V &= V^\dagger \Lambda^{\otimes 2}_{AB,CD}(\mathrm{I} - |\Phi_U\rangle\langle\Phi_U|_{BC})\Lambda^{\otimes 2}_{AB,CD} V \\
&= V^\dagger \left(\sum_{\vec{p},\vec{q}\in\mathcal{P}^{\otimes 2}} \delta^{4-|(\vec{p},\vec{q})|} |\Phi_{\vec{p}}\rangle\langle\Phi_{\vec{q}}|_{AB,CD} \cdot \langle\Phi_{\vec{p}}| (\mathrm{I} - |\Phi_U\rangle\langle\Phi_U|_{BC}) |\Phi_{\vec{q}}\rangle \right) V \\
&= V^\dagger \left(\sum_{\vec{p},\vec{q}\in\mathcal{P}^{\otimes 2}} \delta^{4-|(\vec{p},\vec{q})|} |\Phi_{\vec{p}}\rangle\langle\Phi_{\vec{q}}|_{AB,CD} \cdot \left(\mathbb{1}_{\vec{p},\vec{q}} - \frac{1}{8}\mathrm{Tr}(p_1^* q_1^\top U^\dagger q_2^\top p_2^* U) \right) \right) V \\
&= \sum_{\vec{p}\in\mathcal{P}^{\otimes 2}} \delta^{4-2|\vec{p}|} |\Phi_{\vec{p}}\rangle\langle\Phi_{\vec{p}}| \\
&\quad - \frac{1}{8} \sum_{\vec{P}\in\mathcal{P}^{\otimes n(\ell-1)}} \sum_{\vec{p},\vec{q}} \delta^{4-|(\vec{p},\vec{q})|} |\Phi_{\vec{p}}\rangle\langle\Phi_{\vec{q}}| \mathrm{Tr}(p_1^* q_1^\top U^\dagger q_2^\top p_2^* U) \otimes |\Phi_{\vec{P}}\rangle\langle\Phi_{\vec{P}}| \otimes (\widetilde{W}^{\leq\ell}_{\vec{P}})^\dagger p_1^\dagger U^\dagger p_2^\dagger q_2 U q_1 (\widetilde{W}^{\leq\ell}_{\vec{P}}),
\end{aligned}$$

where $\widetilde{W}^{\leq\ell}_{\vec{P}} \triangleq W_{\ell-1}\tilde{P}_{\ell-1}\ldots W_1\tilde{P}_1$. Above, $\mathbb{1}_{\vec{p},\vec{q}}$ denotes the Kronecker delta symbol. The sum $\sum_{\vec{P}\in\mathcal{P}^{\otimes n(\ell-1)}}$ is over the Pauli noise $\vec{P}$ in layers preceding the gate U. We can also drop the complex conjugate "$*$" because $\mathcal{P} = \{I, X, XZ, Z\}$ are real matrices.

Next, we project qubits C, D onto $|\phi_0\rangle$. Doing so on the term $\sum_{\vec{p}} \delta^{4-2|\vec{p}|} |\Phi_{\vec{p}}\rangle\langle\Phi_{\vec{p}}|_{AB,CD}$ yields the following two-qubit term acting on qubits A, B

$$\frac{4\delta^2}{1+3\delta^2} \sum_{p_1\in\mathcal{P}} \delta^{2-2|p_1|} |\Phi_{p_1}\rangle\langle\Phi_{p_1}|_{AB}\,. \tag{41}$$

We analyze the second term in $V^\dagger h_U V$. For each summand $\vec{P}$, projecting project qubits C, D onto $|\phi_0\rangle$ gives

$$\frac{\delta^2}{1+3\delta^2} \sum_{\vec{p},\vec{q}\in\mathcal{P}^{\otimes 2}} \delta^{2-|(p_1,q_1)|} |\Phi_{p_1}\rangle\langle\Phi_{q_1}| \mathrm{Tr}(p_1 q_1^\top U^\dagger q_2^\top p_2 U) \otimes |\Phi_{\vec{P}}\rangle\langle\Phi_{\vec{P}}| \otimes (\widetilde{W}^{\leq\ell}_{\vec{P}})^\dagger p_1^\top U^\dagger p_2^\top q_2 U q_1 (\widetilde{W}^{\leq\ell}_{\vec{P}}). \tag{42}$$

Next, we apply the following identity

$$\sum_{p_2,q_2\in\mathcal{P}} \mathrm{Tr}(p_1 q_1^\top U^\dagger q_2^\top p_2 U) p_2^\top q_2 = 8 U p_1 q_1^\top U^\dagger \tag{43}$$

to simplify Equation (42) to

$$\frac{8\delta^2}{1+3\delta^2} \sum_{p_1,q_1} \delta^{2-|(p_1,q_1)|} |\Phi_{p_1}\rangle\langle\Phi_{q_1}| \otimes |\Phi_{\vec{P}}\rangle\langle\Phi_{\vec{P}}| \otimes \mathrm{I}. \tag{44}$$

Overall, summing over $\vec{P}$, the second term in $V^\dagger h_U V$ is equal to

$$-\frac{\delta^2}{1+3\delta^2}\sum_{p_1,q_1}\delta^{2-|(p_1,q_1)|}\left|\Phi_{p_1}\right\rangle\left\langle\Phi_{q_1}\right|. \tag{45}$$

Combining this with Equation (41) we get

$$\left\langle\phi_0\right|_{CD} V^\dagger h_U V \left|\phi_0\right\rangle_{CD} = \frac{4\delta^2}{1+3\delta^2}\Lambda_{AB}\left(\sum_{p_1}\left|\Phi_{p_1}\right\rangle\left\langle\Phi_{p_1}\right|_{AB} - \frac{1}{4}\sum_{p_1,q_1}\left|\Phi_{p_1}\right\rangle\left\langle\Phi_{q_1}\right|_{AB}\right)\Lambda_{AB}. \tag{46}$$

A completely similar analysis for two-qubit gates gives the lemma statement. □

We now prove Lemma 4.12.

Proof of Lemma 4.12. Let $\Pi_1 = \mathrm{I}\otimes\phi_0^{\otimes 2}$ and Π_2 be the projector onto the ground space of $V^\dagger h_U V$. As a reminder, it is assumed that $\mathrm{Tr}(\Pi_1\psi') \geq 1-\eta$ and $\mathrm{Tr}\left(V^\dagger h_U V\psi'\right) \leq \alpha$, and the goal is to show $\mathrm{Tr}(\Pi_2\psi') \geq 1 - \frac{\eta}{\Theta(\delta^8)} - \frac{\alpha}{\Theta(\delta^{16})}$.

According to Claim 4.14, the operator $\left\langle\phi_0\right|^{\otimes 2} V^\dagger h_U V \left|\phi_0\right\rangle^{\otimes 2}$ has a spectral gap $\geq 15\delta^8$ for sufficiently small δ. Therefore,

$$\delta^8(\Pi_1 - \phi_0^{\otimes 4}) \leq \Pi_1 V^\dagger h_U V \Pi_1. \tag{47}$$

However, observe the following inequality which follows from $\|h_U\| \leq 1$

$$\Pi_1 V^\dagger h_U V \Pi_1 \leq (\Pi_1 - \Pi_1\Pi_2\Pi_1) \tag{48}$$

Combining the previous inequalities we obtain

$$\Pi_1\Pi_2\Pi_1 \leq \Pi_1 - 15\delta^8(\Pi_1 - \phi_0^{\otimes 4}). \tag{49}$$

Next, we apply Jordan's lemma to decompose Π_1 and Π_2 into 1×1 and 2×2 blocks. Observe that Claim 4.14 implies $\phi_0^{\otimes 4}$ is the *unique* intersection of Π_1 and Π_2, as also evident from Equation (49). Consider two corresponding 2×2 blocks $\left|u\right\rangle\left\langle u\right|$ in Π_1 and $\left|v\right\rangle\left\langle u\right|$ in Π_2, Equation (49) then implies that $|\left\langle u|v\right\rangle|^2 \leq 1 - 15\delta^8$.

On the other hand, letting $\gamma \geq \delta^8$ be the spectral gap of h_U[4], we have

$$\gamma(\mathrm{I}-\Pi_2) \leq V^\dagger h_U V \implies \mathrm{Tr}(\Pi_2\psi') \geq 1 - \frac{\alpha}{\delta^8}. \tag{50}$$

The following expressions follows by writing the projectors Π_1, Π_2 according to Jordan's lemma $\Pi_1 = \phi_0^{\otimes 4} + \sum_i \left|u_i\right\rangle\left\langle u_i\right|$ and $\Pi_2 = \phi_0^{\otimes 4} + \sum_i \left|v_i\right\rangle\left\langle v_i\right|$

$$\mathrm{Tr}\left(\psi'\phi_0^{\otimes 4}\right) + \sum_i \mathrm{Tr}(\left|u_i\right\rangle\left\langle u_i\right|\psi') \geq 1-\eta, \tag{51}$$

$$\mathrm{Tr}\left(\psi'\phi_0^{\otimes 4}\right) + \sum_i \mathrm{Tr}(\left|v_i\right\rangle\left\langle v_i\right|\psi') \geq 1-\frac{\alpha}{\delta^8}. \tag{52}$$

Using $\left|u_i\right\rangle\left\langle u_i\right| + \left|v_i\right\rangle\left\langle v_i\right| \leq (1 + |\left\langle u_i|v_i\right\rangle|)P_i \leq (2-\delta^8)P_i$, where P_i is the projector onto the 2×2 Jordan block i, we get

$$2\,\mathrm{Tr}\left(\psi'\phi_0^{\otimes 4}\right) + (2-\delta^8)\,\mathrm{Tr}\left(\sum_i P_i\psi'\right) \geq 2 - \eta - \frac{\alpha}{\delta^8} \tag{53}$$

Using $\sum_i P_i + \phi_0^{\otimes 4} \leq \mathrm{I}$ and rearranging we get

$$\mathrm{Tr}\left(\psi'\phi_0^{\otimes 4}\right) \geq 1 - \frac{\eta}{\delta^8} - \frac{\alpha}{\delta^{16}}, \tag{54}$$

This concludes the proof of Lemma 4.12. □

[4] We have $h_U^2 = \Lambda^{\otimes 4}(\mathrm{I}-\left|\Phi_U\right\rangle\left\langle\Phi_U\right|)(\Lambda^2)^{\otimes 4}(\mathrm{I}-\left|\Phi_U\right\rangle\left\langle\Phi_U\right|)\Lambda^{\otimes 4} \geq \delta^8\Lambda^{\otimes 4}(\mathrm{I}-\left|\Phi_U\right\rangle\left\langle\Phi_U\right|)\Lambda^{\otimes 4} = \delta^8 h_U$.

4.4.3 Proof of Lemma 4.13 (Teleportation of H_{in} to output column)

We have the "noiseless" version

$$\langle\phi_0| V^\dagger h_j^{\text{in}} V |\phi_0\rangle = \langle\phi_0| \left(\sum_{p,p' \in \mathcal{P}} \delta^{2-|(p,p')|} |\Phi_p\rangle \langle\Phi_{p'}| \otimes (p^\dagger p')_j^{\text{out}} \langle\Phi_p| (\Pi_j \otimes \mathrm{I}) |\Phi_{p'}\rangle \right) |\phi_0\rangle \tag{55}$$

$$= \langle\phi_0| \left(\sum_{p,p' \in \mathcal{P}} \delta^{2-|(p,p')|} |\Phi_p\rangle \langle\Phi_{p'}| \otimes (p^\dagger p')_j^{\text{out}} \frac{1}{2} \operatorname{Tr}\left(p'^\top p^* \Pi_j\right) \right) |\phi_0\rangle \tag{56}$$

$$= \frac{1}{1+3\delta^2} \sum_{p,p' \in \mathcal{P}} \delta^2 (p^\dagger p')_j^{\text{out}} \frac{1}{2} \operatorname{Tr}\left(p'^\top p^* \Pi_j\right) \tag{57}$$

$$= \frac{4\delta^2}{1+3\delta^2} \Pi_j^{\text{out}}. \tag{58}$$

The proof of the robust version is completely similar to that of Lemma 4.12.

5 Verifying QMA via shallow circuits

As shown in Section 4, the parent Hamiltonian robustness properties only depend on circuit depth, so it is desirable to restrict our attention to shallow circuits. Here we show that any QMA protocol can be replaced by one involving a constant depth quantum circuit followed a logarithmic depth classical circuit. The high-level idea is to first use the Feynman-Kitaev mapping to turn an arbitrary QMA protocol into a local Hamiltonian, and then construct a short-depth QMA circuit to measure the energy of the resulting Hamiltonian. For this, we need a low-degree version of the FK mapping.

Claim 5.1 (Degree reduction for FK Hamiltonian)**.** *Any* QMA *protocol involving an n-qubit verifier circuit V with $T = \text{poly}(n)$ two-qubit gates can be mapped into a* 5-LH$[a, b]$ *on* $\text{poly}(n)$ *qubits with $a = 2^{-\text{poly}(n)}$ and $b = a + 1/\text{poly}(n)$. Furthermore, each qubit is involved in at most 7 terms in the Hamiltonian.*

Proof. W.l.o.g., we assume the circuit has been amplified by Lemma 3.2 or Lemma 3.3, such that its completeness is $c = 1 - 2^{-r}$ and $s = 2^{-r}$ with $r = \text{poly}(n)$.

We first recall the FK Hamiltonian [1] here to observe that it is not sparse. For all $T \in \mathbb{N}$ and $t \leq T$, we define the unary clock states as $|u(t,T)\rangle = |1^t\rangle \otimes |0^{T-t}\rangle$. The clock qubits are index by $t \in [T]$ and the data qubits are indexed by $i \in [n]$. Let m be the number of ancilla qubits, so that the witness has $n - m$ qubits. The FK Hamiltonian consists of four parts acting on a unary clock register and a data register: (1) initialization terms

$$H_{\text{in}} = |0\rangle\langle 0|_{t=0} \otimes \left(\sum_{i=1}^{m} |1\rangle\langle 1|_i \right),$$

(2) propagation terms (note there are no clock qubits -1 and $T+1$)

$$H_{\text{prop}} = \frac{1}{2} \left(\sum_{t=1}^{T} (|100\rangle\langle 100| + |110\rangle\langle 110|)_{t-1,t,t+1} - |110\rangle\langle 100_{t-1,t,t+1}| \otimes U_t - |100\rangle\langle 110|_{t-1,t,t+1} \otimes U_t^\dagger \right),$$

(3) clock validity terms

$$H_{\text{clock}} = \sum_{t=1}^{T} |01\rangle\langle 01|_{t-1,t},$$

(4) and output check term

$$H_{\text{out}} = |1\rangle\langle 1|_T \otimes |0\rangle\langle 0|_1 .$$

As it is, the FK Hamiltonian has high degree due to the $t = 0$ clock qubit, which participates in m terms in H_{in} and the data qubits, which participate in as many terms in H_{prop} as the number of nontrivial gates acting on the qubit.

Note: *Here, we justify the claim in Table 1 that there is a combinatorial state that violates a $O(\frac{1}{T})$ fraction of terms. For example, the state $|0100\ldots\rangle_{\text{clock}} \otimes |0\rangle_{\text{data}}^{\otimes n}$ contains a fixed invalid clock configuration and hence satisfies all the terms in the Feynman-Kitaev Hamiltonian, except 2 terms from H_{clock}.*

We reduce the degree of data qubits by transforming V into a new circuit V' that acts on $n' = nT$ qubits divided into T n-qubit blocks. After applying the first gate in V on the first qubit block, we apply n SWAP gates to swap the first and second blocks. Then, the second gate in V is applied on the second block of V', and so on. This way, the qubits in V' are acted on by at most 3 nontrivial gates. The number of nontrivial gates in V' is $T' = O(nT)$.

We reduce the degree of the $t = 0$ clock qubit by observing that the initialization of ancilla qubit i only need to be verified right before the first gate acting on it. Let $t_i \in [T]$ be this gate, then we use the following initialization term (note there are no clock qubits -1 and $T + 1$)

$$H_{\text{in},i} = |10\rangle\langle 10|_{t_i-1,t_i} \otimes |1\rangle\langle 1|_i . \tag{59}$$

Applying this modified FK mapping (with modified H_{in}) to the circuit V' we obtain a 5-local Hamiltonian H_{FK} in which each qubit involves in at most 7 terms. The energy analysis in [1] still applies for this modified construction. Indeed, according to [1], $H_{\text{prop}} + H_{\text{clock}}$ has ground states of the form

$$|\Psi\rangle := \frac{1}{\sqrt{T'+1}} \sum_{t=0}^{T'} |u(t, T')\rangle \otimes U_t \cdots U_1 |\psi\rangle , \text{ for any } |\psi\rangle \in (\mathbb{C}^2)^{n'} \tag{60}$$

In the completeness case, setting $|\psi\rangle = |0^{m'}\rangle |\xi\rangle$, where $|\xi\rangle$ is the witness that V' accepts with probability c, gives an energy of $a = O((1 - c)/T') = 2^{-\text{poly}(n')}$.

In the soundness case, the main step of the proof is Equation 14.17 in [1] in which the author bounds $\max_{|\psi\rangle} \langle\Psi| \Pi_1 |\Psi\rangle$ where Π_1 is the projector onto the nullspace of $H_{\text{in}} + H_{\text{out}}$. However, it can be seen that modifying H_{in} as in Equation (59) does not change this quantity which remains to be

$$\langle\Psi| \Pi_1 |\Psi\rangle = 1 - \frac{1}{T'+1} \left(\langle\psi| (\sum_{i=1}^{m'} |1\rangle\langle 1|_i) |\psi\rangle + \langle\psi| V'^{\dagger} |0\rangle\langle 0|_1 V' |\psi\rangle \right) \tag{61}$$

by noting that $U_1^{\dagger} \ldots U_{t_i-1}^{\dagger} (|1\rangle\langle 1|_i) U_{t_i-1} \ldots U_1 = |1\rangle\langle 1|_i$ for any i. Therefore, according to [1], any state has energy no smaller than $b = \Omega((1 - \sqrt{s})/T'^3) = 1/\text{poly}(n')$. □

Claim 5.2 (Log-depth QMA). *Any* QMA *protocol involving an n-qubit verifier circuit V with $T = \text{poly}(n)$ two-qubit gates can be converted into a $O(\log n)$-depth* QMA *protocol on $\text{poly}(n)$ qubits, whose completeness is $1 - 2^{-r}$ and soundness is 2^{-r} with $r = \text{poly}(n)$. More specifically, the $O(\log n)$-depth circuit involves a constant-depth quantum circuit that ends with computational basis measurements, followed by a $O(\log n)$-depth classical circuit.*

Proof. Given any QMA protocol V_0 (with n_0 qubits including the size of the witness and number of gates $T_0 = \text{poly}(n_0)$), we first convert it into the low-degree FK Hamiltonian using Claim 5.1. The Hamiltonian $H_{\text{FK}} = \sum_{i=1}^{m} h_i$ acts on $n = \Theta(n_0 T_0)$ qubits, contains $m = \Theta(n_0 T_0) = \Theta(n)$ projectors that are at most 5-local, and has a promise gap of $b - a = \Omega(m^{-3})$. Each qubit participates in at most 7 terms h_i.

Next, we construct a constant-depth circuit V extracting the satisfiability of the Hamiltonian terms in H_{FK}. The circuit consists of m ancillas initialized to $|0\rangle$. Upon receiving an n-qubit witness state $|\xi\rangle$, V applies unitaries of the form $\mathrm{C_{h_i}NOT} = (\mathrm{I} - h_i) \otimes I_i + h_i \otimes X_i$, which, conditioned on the reduced state of $|\xi\rangle$ being in $\mathrm{supp}(h_i)$, flip ancilla qubit i. In particular, consider the decomposition of the terms h_i into $L = O(1)$ groups, $H_1, \ldots, H_L$ such that the terms in each group are pairwise non-overlapping. Let $\Pi_\ell = \bigotimes_{i:h_i \in H_\ell} (\mathrm{I} - h_i)$. The layer $\ell \in [L]$ of V is $V_\ell = \prod_{i:h_i \in H_\ell} \mathrm{C_{h_i}NOT}$. After applying $V_Q \triangleq V_L \ldots V_2 V_1$, we measure the ancillas in the Z basis to get a bitstring $x \in \{0,1\}^m$, and compute the OR function on x and output $\overline{\mathrm{OR}(x)}$. The OR function on m bits can be computed by a $O(\log m)$-depth Boolean circuit[5], which can in turn be made reversible with constant space overhead [31, Section 3.2.5].

The circuit V output 1 if and only if the ancillas are measured in the all-zeros string, which happens with probability $\Pr[x = 0^m] = |\langle 0^m | V_Q |\xi\rangle \otimes |0^m\rangle|^2 = \mathrm{Tr}\Big(\mathrm{DL}^\dagger \mathrm{DL} |\xi\rangle\langle\xi|\Big)$ where $\mathrm{DL} \triangleq \Pi_L \ldots \Pi_2 \Pi_1$ is the detectability lemma operator [26].

Below we show that, if V_0 accepts, then V accepts $1 - 2^{-\mathrm{poly}(n)}$ and if V_0 rejects then V accepts with $1 - \Omega(n^{-2})$. In addition, the soundness can be depth-efficiently improved to $2^{-\mathrm{poly}(n)}$.

Completeness The prover sends the witness state $|\xi\rangle$ such that $\langle\xi| H_{\mathrm{FK}} |\xi\rangle \leq 2^{-\mathrm{poly}(n)}$ (the case of mixed state witness follows by linear extension). Also $\langle\xi| h_i |\xi\rangle \leq 2^{-\mathrm{poly}(n)}$ for any $i \in [m]$. Using the quantum union bound (Lemma 3.7) on H_{FK} and $|\xi\rangle$ we can bound

$$1 - \mathrm{Tr}\Big(\mathrm{DL}^\dagger \mathrm{DL} |\xi\rangle\langle\xi|\Big) \leq 4 \sum_i \langle\xi| h_i |\xi\rangle \leq 2^{-\mathrm{poly}(n)}. \tag{62}$$

So V outputs 1 with probability at least $c = 1 - 2^{-\mathrm{poly}(n)}$.

Soundness According to Claim 5.1, for any state $|\xi\rangle$ we have $\langle\xi| H_{\mathrm{FK}} |\xi\rangle \geq \Omega(n^{-3})$. Observe that the terms in H_{FK} are projectors and each of them overlaps with at most $g = 34$ others, so we can apply the detectability lemma (Lemma 3.6) on H_{FK} and $|\xi\rangle$

$$\mathrm{Tr}\Big(\mathrm{DL}^\dagger \mathrm{DL} |\xi\rangle\langle\xi|\Big) \leq \frac{1}{\Omega(n^{-3}) + 1} \leq 1 - \Omega(n^{-3}) = s. \tag{63}$$

Finally, soundness can be amplified to $2^{-\mathrm{poly}(n)}$ while keeping the depth logarithmic via the weak amplification procedure in Lemma 3.2. In particular, this procedure [1] works by using $q = \mathrm{poly}(n)$ copies of V in parallel. The prover is expected to send q copies of an accepting state. We perform OR on the q decision bits of the copies in depth $O(\log q) = O(\log n)$. It is a standard fact that we can assume w.l.o.g. the prover sends an unentangled state between these q copies (e.g., see [1, Lemma 14.1]). Thus, a simple application of Chernoff's bound achieves the amplified soundness whenever q is a sufficiently large polynomial in n.

□

We note that the technique in [32], where the author studies the hardness of distinguishing log-depth circuits, seems to also give a log-depth verification procedure for QMA. However, their quantum circuit is necessarily logarithmic-depth due to the use of n-qubit controlled SWAP gates. This is to be compared with our construction, where the quantum circuit is constant-depth and followed by log-depth classical circuit. This could be a useful feature for fault tolerance protocols and possible implications for the quantum PCP conjecture that we discuss in this work.

6 Computational complexity of injective PEPS

We now discuss a hardness result on the creation and contraction of certain tensor network states that follows from our construction. A summary of our results can be found in Table 2.

[5] A log-depth OR circuit is as follows: the first layer computes pairwise OR's $(x_1 \vee x_2)$, $(x_3 \vee x_4)$, ..., the second layer similarly computes OR pairwise on the output of the first layer, and so on.

Definition 6.1 (PEPS)**.** *A* projected entangled pair state (PEPS) *is any (unnormalized) state that can be obtained by the following procedure: consider a graph and associate to each vertex v as many D-dimensional spins as there are edges incident to v. Assume that the spins associated to the end points of an edge form maximally entangled states $|\mathrm{EPR}_D\rangle = \sum_{i=1}^{D} |i\rangle |i\rangle$. The PEPS is obtained by applying a linear map $P_v : \mathbb{C}^D \otimes \cdots \otimes \mathbb{C}^D \to \mathbb{C}^d$ at each vertex v. Without affecting the computational complexity, we further allow the virtual states to be any maximally entangled states of the form $(\mathrm{I} \otimes U) |\mathrm{EPR}_D\rangle$. We can also assume $\|P_v\| \leq 1$.*

In [15] it was shown that preparing PEPS as a quantum state is PostBQP-hard, where PostBQP is a large complexity class that contains QMA. The idea of the proof is that measurement-based quantum computation with the power to post-select on the measurement outcomes reduces to preparing a PEPS. The power to post-select on the outcomes of a quantum computation is due to the fact that the local maps P_v are allowed to be non-invertible. Hence, it is natural to ask what happens when we reduce the power of preparing arbitrary PEPS by removing the ability to post-select. We do this by considering a subclass of tensor networks called *injective PEPS* [33].

Definition 6.2 (Injective PEPS)**.** *A PEPS on n spins is* $\delta(n)$-injective *if the local maps P_v are non-singular matrices with singular values bounded from below by $\Omega(\delta(n))$.*

Our construction gives the following hardness result on the preparation of injective PEPS.

Theorem 6.3. *Preparing constant-injective PEPS states in two or higher dimensions with bond dimension $D \geq 4$ and physical dimension $d \geq 4$ allows solving* BQP*-hard problems.*

Proof. Claim 2.1 tells us that we can encode a noisy quantum computation suffering from i.i.d. depolarizing noise with constant error probability $\delta^2/(1+3\delta^2)$ into a δ-injective PEPS state. Choosing δ to be a sufficiently small constant, the quantum fault-tolerance threshold theorem (Theorem 3.9) states that we can ε-approximate any noise-free circuit C with a noisy circuit $\tilde{C}$ with polylogarithmic overhead in $1/\varepsilon$. Note that the threshold theorem holds even when the circuit connectivity is restricted to one dimension (Corollary 3.10), so the computational hardness persists on two-dimensional injective PEPS. □

The above (state) BQP-hardness result can be understood as a complement to previous works in efficient quantum algorithms for preparing injective PEPS under assumptions on the parent Hamiltonian spectral gap [17, 18]. We leave it as an open question to obtain tight upper bound on the complexity of preparing injective PEPS states (an upper bound is (state) PostBQP due to [15]).

We next discuss the classical complexity of injective PEPS. PEPS is conceived as an efficient classical description of quantum states and an important application is contracting a PEPS in order to evaluate the value of a given observable.

Task 6.4 (PEPS observable contraction)**.** *Given a PEPS describing an unnormalized state $|\psi\rangle$ and a local observable O, calculate the normalized expectation value $\frac{\langle\psi|O|\psi\rangle}{\langle\psi|\psi\rangle}$.*

Ref. [19] gave a quasi-polynomial time classical algorithn to contract injective PEPS under assumptions on the parent Hamiltonian spectral gap. The complexity of injective PEPS has also been implicitly studied in [34], where the authors showed that random PEPS, whose local maps are i.i.d. Gaussian are #P-hard to contract to exponential additive precision. However, the random PEPS ensemble of [34] has injectivity $1/\Omega(\mathrm{poly}(n))$ with high probability, and their result does not necessarily indicate that a #P-hard PEPS instance would have constant injectivity. Similarly, the #P-hard PEPS instances in [15] can be seen to be non-injective, even after blocking[6] [33]. Using the construction in this work, we obtain the following hardness results for contracting PEPS with *constant* injectivity.

Theorem 6.5. *For constant-injective PEPS states in two or higher dimensions with bond dimension $D \geq 4$ and physical dimension $d \geq 4$, evaluating local observable expectation values to $O(1)$ additive error is* BQP*-hard.*

[6] This is because their local maps have the form $|0\rangle\langle 0|$, which will remain being rank-1 after blocking.

Proof. The BQP-hardness, similar to Theorem 6.3, follows from encoding a noisy BQP computation C into our injective PEPS $|\Psi\rangle$ and invoking the threshold theorem. The difference is that at the end of the fault-tolerant circuit $\tilde{C}$ that simulates the noiseless circuit C in Theorem 3.9, we further perform a fault-tolerant decoding circuit that transforms the encoded output into a physical output state. For concatenated-code fault tolerance, this procedure is described in Section 4 of [35] (also see Section 6 of [36]). This decoding circuit results in a physical error rate per physical qubit of the output state which is bounded by some constant value. We encode the entire fault-tolerant circuit, including the fault-tolerant decoding part, into our injective PEPS with noise rate below the threshold. Then evaluating to $O(1)$-additive error the Pauli-Z expectation on the first qubit of the PEPS output column decides the BQP computation C. □

Task	**PEPS**	**Injective PEPS**
State preparation	PostBQP-complete	BQP-hard
Multiplicative-error contraction	#P-complete	#P-complete*
Additive-error contraction	BQP-hard	BQP-hard

Table 2: Computational complexity of general PEPS [15] and constant-injective PEPS. *The #P-hardness of injective PEPS requires a specific non-local observable in Theorem 6.6.

If we instead consider $O(1)$-multiplicative error expectation value evaluation of PEPO non-local observables in Task 6.4, then we obtain a classical hardness matching that of general non-injective PEPS [15]. Is is a simple observation that *exact* observable evaluation of Task 6.4 for general PEPS and PEPO observables[7] is in #P. For this, we reduce this task to norm evaluation of PEPS, which was shown to be in #P in [15]. Observe that $\langle\psi| O |\psi\rangle = (\langle\psi| (O+\mathrm{I})(O+\mathrm{I}) |\psi\rangle - \langle\psi| OO |\psi\rangle - \langle\psi|\psi\rangle)/2$. Each of $(O+\mathrm{I})|\psi\rangle$, $O|\psi\rangle$, and $|\psi\rangle$ are PEPS states since O is a PEPO. Thus, evaluating $\langle\psi| O |\psi\rangle / \langle\psi|\psi\rangle$ is in #P.

In order for our construction to go through for multiplicative errors, we have to constrain the type of observable to be describable as a *tree tensor network (TTN)* which is a PEPO defined on a tree graph [37].

Theorem 6.6. *For constant-injective PEPS states in two or higher dimensions with bond dimension $D \geq 4$ and physical dimension $d \geq 4$, evaluating the expectation value of a tree tensor network observable to $O(1)$-multiplicative error is #P-hard.*

Proof. We encode the "quantum sum" problem[8], well-known in the random circuit sampling literature [38], into a fault-tolerant quantum circuit that exponentially suppresses the local depolarizing noise:

$$S = \sum_{x\in\{0,1\}^n} (-1)^{f(x)}, \qquad f : \{0,1\}^n \mapsto \{0,1\}. \tag{64}$$

It is well known that $O(1)$-multiplicative error approximation of S^2 remains #P-hard (under Turing reduction) and can be converted into $O(1)$-multiplicative error estimation of the amplitude $|\langle 0^n| C |0^n\rangle|^2 = S^2$ of a quantum circuit C [38].

This can be further simplified to a measurement on one qubit as follows. Let W be a reversible circuit that computes the inverse OR function, using additional ancillas initialized to a computational basis state $|0^m\rangle_{\mathrm{anc}}$, and then writing the result on the $(n+1)$-st qubit. Note the identity $(\mathrm{I}\otimes\langle 0^m|_{\mathrm{anc}})W^\dagger |0\rangle\langle 0|_{n+1} W(\mathrm{I}\otimes |0^m\rangle_{\mathrm{anc}}) = |0\rangle\langle 0|^n$. This implies - defining $C' = WC$ and abbreviating $|0^n\rangle \otimes |0^m\rangle_{\mathrm{anc}}$ as $|0^{n+m}\rangle$ - that $|\langle 0| C' |0^{n+m}\rangle|^2 = |\langle 0^n| C |0^n\rangle|^2$.

[7] *Projected entangled pair operators (PEPO)* are an efficiently describable family of operators, which can be thought of as the operator version of PEPS, i.e., local maps are $P_v : \mathbb{C}^D \otimes \cdots \otimes \mathbb{C}^D \to \mathbb{C}^{d\times d}$.

[8] Ref. [15] instead used the "classical sum" $\sum_x f(x)$, whose multiplicative error estimation is significantly easier than #P. So strictly speaking, #P-hardness of multiplicative error estimation of general PEPS contraction does not follow from Ref. [15].

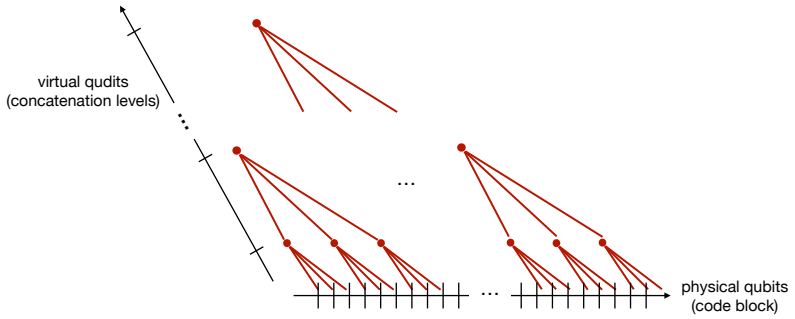


Figure 8: The recursive-majority readout of a codeblock is encoded into a tree tensor network observable.

MAJ

Next, we use a concatenated-code fault-tolerant circuit $\tilde{C}$ from Theorem 3.9 that approximates C' to additive error $\varepsilon = e^{-\Omega(n)}$, which implies that $|\langle\bar{0}|\tilde{C}|\bar{0}^{n+m}\rangle|^2$ approximates $|\langle 0|C'|0^{n+m}\rangle|^2 = S^2$ to $O(1)$-multiplicative error. Unlike in the proof of Theorem 6.5, we cannot afford to perform the noisy fault-tolerant decoding circuit, as doing so would no longer guarantee a multiplicative-closeness between $|\langle\bar{0}|\tilde{C}|\bar{0}^{n+m}\rangle|^2$ and S^2. Instead, we directly read out the logical information in the output of the noisy execution of $\tilde{C}$ by performing a recursive majority vote on the first logical qubit register (cf. Lemma 9 in [20]). The observable that represents this majority vote is a tree tensor network. For concreteness, suppose $\tilde{C}$ is constructed from concatenating L levels of the $[[7,1,3]]$-Steane code, then this recursive majority vote means that we first take the majority in each block of size 7, then we take the majority, of 7 such majority bits, and so on for L levels, to give one output bit. Here, $L = \Theta(\log n)$ for the desired error suppression, which means each logical code block consists of $\mathrm{poly}(n)$ physical qubits. This recursive majority vote can readily be encoded in an L-level tree tensor network operator O as illustrated in Figure 8. Thus, evaluating the expectation value of O to multiplicative error on our injective PEPS gives a multiplicative error estimation of S^2. □

Since tree tensor networks are themselves easy to contract (similar to an MPO), we conclude from Theorem 6.6 that the #P hardness must be arising from the injective PEPS itself. Proving the same theorem with local - or product - observables is an interesting open question.

7 Open questions

This work brings up a series of relevant open questions.

- Our main question is if we can achieve a soundness of $1/\mathrm{poly}(D)$. In our proof of Theorem 4.3, two steps are needed - robust teleportation of H_{in} and a small number of Pauli errors in the low energy states. The challenging part is the robust teleportation of H_{in}, which we do not know how to achieve when the energy is $1/\mathrm{poly}(D)$. It turns out that we can enforce low Pauli errors by adding new Hamiltonian terms with $\mathrm{poly}(D)$ locality. Specifically, for each local region A of $\frac{D}{\delta^4}$ EPR locations we can define a Hamiltonian term that penalizes $\geq 10\delta^2 \cdot \frac{D}{\delta^4}$ Pauli errors, i.e.,

$$h_A^{\mathrm{low}} = \sum_{\vec{P}\in\mathcal{P}^A : |\vec{P}|\geq 10\delta^2\cdot\frac{D}{\delta^4}} |\Phi_{\vec{P}}\rangle\langle\Phi_{\vec{P}}| .$$

  Since δ is chosen $\frac{1}{\mathrm{poly}(D)}$, there new Hamiltonian terms are $\mathrm{poly}(D)$ local, which is $\mathrm{polylog}(n)$ when $D = \mathrm{polylog}(n)$. In addition, these Hamiltonian terms are unchanged under the unitary V defined in Section 2.3. Now, lets choose a collection of regions $A_1, \ldots A_m$ (with $m = O(n\delta^4)$ such that each of nD EPRs is involved in $O(1)$ regions) and add the following Hamiltonian to the existing Hamiltonian

$H^{\text{low}} = \frac{1}{m}\sum_i h_{A_i}^{\text{low}}$. Note that H^{low} does not change the ground state energy density too much: The ground state achieves energy density of at most e^{-27D/δ^2} by Chernoff bound. On the other hand, any state with energy density at most $\frac{1}{100D^2}$ has the property that it has at most $O(D\delta^2 + \frac{1}{D})$ fraction of Pauli errors with probability $1 - 1/D$. For this, consider the 'fraction of Pauli errors' operator in local regions $\frac{\delta^4}{D}\sum_{j\in A_i}(\mathrm{I} - \Phi_{I_j})$. Note that $\frac{\delta^4}{D}\sum_{j\in A_i}(\mathrm{I} - \Phi_{I_j}) \preceq 10\delta^2(\mathrm{I} - h_{A_i}^{\text{low}}) + h_{A_i}^{\text{low}} \preceq 10\delta^2\mathrm{I} + h_{A_i}^{\text{low}}$. Thus, the total fraction of Pauli errors satisfies

$$\frac{1}{nD}\sum_j(\mathrm{I} - \Phi_{I_j}) \preceq O(1)\frac{1}{m}\sum_{i=1}^{m}(\frac{\delta^4}{D}\sum_{j\in A_i}(\mathrm{I} - \Phi_{I_j})) \preceq O(1)\cdot 10\delta^2\mathrm{I} + \frac{O(1)}{m}\sum_{i=1}^{m} h_{A_i}^{\text{low}}.$$

Thus, the expected fraction of Pauli errors in such low energy states is $O(\delta^2 + 1/D^2)$. The claim follows from Markov's inequality.

- In the introduction and Appendix A, we outline a connection between the 'polylog weaker' classical PCP result and adversarial fault tolerance. It is expected that adversarial fault tolerance may use good classical codes, but we do not see a clear use of local testability (beyond locally testable repetition codes for reading the answer from fault-tolerant computation). Could 'polylog weaker' classical PCP be achieved without strong reliance on local testability (or local decodability)?

- Can the depth of BQP protocol be reduced to polylogarithmic in the input size? This does not follow from the depth reduction of QMA due to the presence of witness. Thus, the heart of the question is if the ground state of the tensor network Hamiltonian can be prepared in low depth when witness is absent. One possibility is to run an adiabatic algorithm tuning δ from 1 to a smaller value. We do not know about the spectral gap in this process - and it is likely small. But suppose that we go ahead and tune δ adiabatically for small duration. Can we argue that we end up in a low energy state of the parent Hamiltonian? If that is the case, we would still encode the answer to the computation if we started from a fault-tolerant circuit.

- Can we recover QMA-hardness of the local Hamiltonian problem as an alternative to Feynman-Kitaev clock? This is not the goal of this paper, but is interesting in its own right as it would directly map a QMA protocol to a local Hamiltonian in two dimensions. We expect that a resolution of our main open question would lead to this result.

7.1 Acknowledgements

We thank Fernando Brandão, Steve Flammia, Yeongwoo Hwang, Zeph Landau, Spiros Michalakis, Chinmay Nirkhe, Mehdi Soleimanifar and Umesh Vazirani for helpful discussions. We especially thank Steve Flammia for pointing us to the reference [23] and Chinmay Nirkhe for pointing us to [16]. AA and QTN acknowledge support through the NSF Award No. 2238836. AA acknowledges support through the NSF award QCIS-FF: Quantum Computing & Information Science Faculty Fellow at Harvard University (NSF 2013303). QTN acknowledges support through the Harvard Quantum Initiative PhD fellowship.

References

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

[2] Julia Kempe and Oded Regev. 3-Local Hamiltonian is QMA-complete. Quantum Computation and Information, Vol. 3(3), p. 258-64, 2003, 2003.

[3] Dorit Aharonov, Daniel Gottesman, Sandy Irani, and Julia Kempe. The power of quantum systems on a line. *Communications in mathematical physics*, 287(1):41–65, 2009.

[4] Julia Kempe, Alexei Kitaev, and Oded Regev. The complexity of the local hamiltonian problem. *Siam journal on computing*, 35(5):1070–1097, 2006.

[5] Roberto Oliveira and Barbara M. Terhal. The complexity of quantum spin systems on a two-dimensional square lattice. *Quant. Inf, Comp.*, 8(10):0900–0924, 2008.

[6] Daniel Nagaj. Fast universal quantum computation with railroad-switch local Hamiltonians. *Journal of Mathematical Physics*, 51:062201, 2010.

[7] Nikolas P. Breuckmann and Barbara M. Terhal. Space-time circuit-to-Hamiltonian construction and its applications. *Journal of Physics A: Mathematical and Theoretical*, 47(19):195304, 2014.

[8] Sandy Irani. Ground state entanglement in one-dimensional translationally invariant quantum systems. *Journal of Mathematical Physics*, 51(2):022101, 02 2010.

[9] Dorit Aharonov, Aram W. Harrow, Zeph Landau, Daniel Nagaj, Mario Szegedy, and Umesh Vazirani. Local tests of global entanglement and a counterexample to the generalized area law. In *2014 IEEE 55th Annual Symposium on Foundations of Computer Science*, pages 246–255, 2014.

[10] Chinmay Nirkhe, Umesh Vazirani, and Henry Yuen. Approximate Low-Weight Check Codes and Circuit Lower Bounds for Noisy Ground States. In Ioannis Chatzigiannakis, Christos Kaklamanis, Dániel Marx, and Donald Sannella, editors, *45th International Colloquium on Automata, Languages, and Programming (ICALP 2018)*, volume 107 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 91:1–91:11, Dagstuhl, Germany, 2018. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik.

[11] Dorit Aharonov, Wim Van Dam, Julia Kempe, Zeph Landau, Seth Lloyd, and Oded Regev. Adiabatic quantum computation is equivalent to standard quantum computation. *SIAM review*, 50(4):755–787, 2008.

[12] Alexandru Gheorghiu, Theodoros Kapourniotis, and Elham Kashefi. Verification of quantum computation: An overview of existing approaches. *Theory of computing systems*, 63:715–808, 2019.

[13] Dorit Aharonov and Tomer Naveh. Quantum NP-a survey. *arXiv preprint quant-ph/0210077*, 2002.

[14] Dorit Aharonov, Itai Arad, and Thomas Vidick. Guest column: the quantum pcp conjecture. *Acm sigact news*, 44(2):47–79, 2013.

[15] Norbert Schuch, Michael M Wolf, Frank Verstraete, and J Ignacio Cirac. Computational complexity of projected entangled pair states. *Physical review letters*, 98(14):140506, 2007.

[16] Lijie Chen. Unpublished work. 2020.

[17] Martin Schwarz, Kristan Temme, and Frank Verstraete. Preparing projected entangled pair states on a quantum computer. *Physical Review Letters*, 108(11):110502, 2012.

[18] Yimin Ge, András Molnár, and J Ignacio Cirac. Rapid adiabatic preparation of injective projected entangled pair states and gibbs states. *Physical Review Letters*, 116(8):080503, 2016.

[19] Martin Schwarz, Olivier Buerschaper, and Jens Eisert. Approximating local observables on projected entangled pair states. *Physical Review A*, 95(6):060102, 2017.

[20] D Aharonov and M Ben-Or. Fault-tolerant quantum computation with constant error rate. *arXiv preprint quant-ph/9906129*, 1999.

[21] Anurag Anshu and Nikolas P. Breuckmann. A construction of combinatorial NLTS. *Journal of Mathematical Physics*, 63(12):122201, 2022.

[22] Dave Bacon, Steven T Flammia, Aram W Harrow, and Jonathan Shi. Sparse quantum codes from quantum circuits. *IEEE Transactions on Information Theory*, 63(4):2464–2479, 2017.

[23] Stephen D Bartlett and Terry Rudolph. Simple nearest-neighbor two-body hamiltonian system for which the ground state is a universal resource for quantum computation. *Physical Review A*, 74(4):040302, 2006.

[24] Dorit Aharonov and Sandy Irani. Translationally invariant constraint optimization problems. *arXiv preprint arXiv:2209.08731*, 2022.

[25] Chris Marriott and John Watrous. Quantum Arthur–Merlin games. *computational complexity*, 14:122–152, 2005.

[26] Anurag Anshu, Itai Arad, and Thomas Vidick. Simple proof of the detectability lemma and spectral gap amplification. *Physical Review B*, 93(20):205142, 2016.

[27] Jingliang Gao. Quantum union bounds for sequential projective measurements. *Physical Review A*, 92(5):052331, 2015.

[28] Camille Jordan. Essai sur la géométrie à n dimensions. *Bulletin de la Société mathématique de France*, 3:103–174, 1875.

[29] Barbara M Terhal. Quantum error correction for quantum memories. *Reviews of Modern Physics*, 87(2):307, 2015.

[30] Emanuel Knill, Raymond Laflamme, and Wojciech H Zurek. Resilient quantum computation. *Science*, 279(5349):342–345, 1998.

[31] Michael A. Nielsen and Isaac L. Chuang. *Quantum Computation and Quantum Information*. Cambridge University Press, Cambridge, 2000.

[32] Bill Rosgen. Distinguishing short quantum computations. *arXiv preprint arXiv:0712.2595*, 2007.

[33] Norbert Schuch, Ignacio Cirac, and David Pérez-García. Peps as ground states: Degeneracy and topology. *Annals of Physics*, 325(10):2153–2192, 2010.

[34] Jonas Haferkamp, Dominik Hangleiter, Jens Eisert, and Marek Gluza. Contracting projected entangled pair states is average-case hard. *Physical Review Research*, 2(1):013010, 2020.

[35] Emanuel Knill and Raymond Laflamme. Concatenated quantum codes. *arXiv preprint quant-ph/9608012*, 1996.

[36] Daniel Gottesman. Fault-tolerant quantum computation with constant overhead. *Quantum Information and Computation*, (15-16):1338–1372, 2014.

[37] Shi-Ju Ran, Emanuele Tirrito, Cheng Peng, Xi Chen, Luca Tagliacozzo, Gang Su, and Maciej Lewenstein. *Tensor network contractions: methods and applications to quantum many-body systems*. Springer Nature, 2020.

[38] Dominik Hangleiter and Jens Eisert. Computational advantage of quantum random sampling. *Reviews of Modern Physics*, 95(3):035001, 2023.

A Adversarial fault tolerance and polylog-weaker classical PCP

Here we discuss the connection between adversarial fault tolerance and polylog-weaker classical PCP. We take any NP-hard classical Constraint Satisfaction Problem (CSP) $C_1 = \frac{1}{m}\sum_i C_{1,i}$ on n bits in which each constraint acts on a constant number of bits and each bit participates in a constant number of constraints (for example, a 2D Ising model).

We can verify whether C_1 is satisfiable in logarithmic depth via a similar circuit as in Section 5. Specifically, the verifier (1) adds m ancilla bits (ith bit corresponding to the ith constraint) initialized to 0, (2) asks prover for the satisfying solution, and for each constraint, (3) flips the corresponding ancilla bit if the corresponding constraint was violated. Step (2) can be done in $O(1)$ depth since the constraints can be divided into $O(1)$ groups such that each group contains only non-intersecting constraints. Once this is done, we can run an OR function in $O(\log n)$ depth on the ancilla bits to accept or reject.

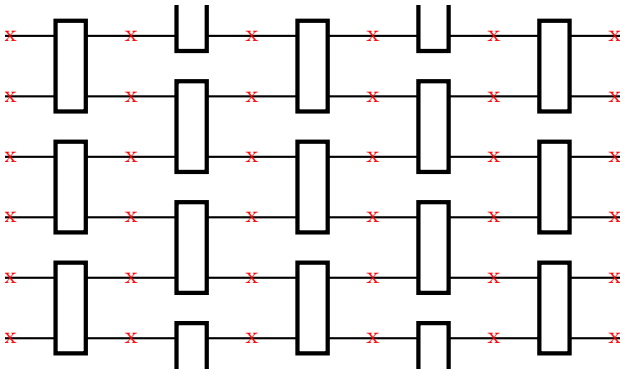


Figure 9: In Cook-Levin transformation from a classical circuit to a classical CSP, one places a binary variable on each wire (red 'x') and enforces a (local) consistency constraint on 4 variables that are input and output to a gate. Fix an assignment to the variables. If the assignment satisfies a local consistency constraint, then we can view it as a correct execution of the gate. On the other hand, if the assignment violates a local consistency constraint, then we can view it as an error in the computation.

If such $O(\log n)$ depth verification circuit can be transformed into a polylog(n) depth circuit sound against $\frac{1}{\text{polylog}(n)}$ fraction adversarial errors within the circuits, then applying the Cook-Levin transformation (Figure 9) on this fault-tolerant circuit gives us a CSP C_2 in which $\frac{1}{\text{polylog}(n)}$ energy density assignments still encode the accept/reject answer of C_1. The crucial observation here is that any violated constraint in C_2 can be viewed as an adversarial error on the circuit (Figure 9). Note that we only require adversarial fault tolerance for NP protocols, which could be different (possibly easier to achieve) from universal adversarially fault-tolerant computation. The fault-tolerant circuit will produce the output encoded in a length $\Omega\left(\frac{n}{\text{polylog}(n)}\right)$ repetition code.[9] We can fault-tolerantly verify the logical output bit by local checks. Let C_{out} be the CSP that realizes this check and penalizes the rejecting encoded output in C_2.

We claim that the CSP $\frac{1}{2}C_2 + \frac{1}{2}C_{\text{out}}$ has a promise gap of $\frac{1}{\text{polylog}(n)}$. In the yes case where C_1 is satisfiable, the above fault-tolerant computation accepts for some witness and hence C_2 as well as C_{out} are satisfiable. In the no case where C_1 is unsatisfiable, let x be an assignment to the variables in C_2 that have energy at most $\frac{1}{\text{polylog}(n)}$. This assignment encodes the fault-tolerant computation above with $\frac{1}{\text{polylog}(n)}$ fraction of adversarial errors. By assumption on fault-tolerant computation, the logical output bit of this computation should still be an encoding of logical 1. Hence the penalty from C_{out} is $\frac{1}{\text{polylog}(n)}$ due to the distance of the repetition code encoding the logical output bit. Thus, the energy of $\frac{1}{2}C_2 + \frac{1}{2}C_{\text{out}}$ is at least $\frac{1}{\text{polylog}(n)}$.

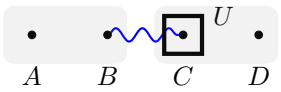
This concludes our claim that adversarial classical fault tolerance with polynomial depth overhead implies a 'polylog weaker' version of classical PCP theorem.

[9] Given the error budget with adversary, the output cannot be encoded in a code of length smaller than $O\left(\frac{n}{\text{polylog}(n)}\right)$.

B Locality of Clifford propagation terms

We show that the rotated propagation term $V^\dagger h_{i,j}^{(\ell)} V$ associated with a Clifford gate U (acting on qubits i,j in layer $\ell < D$) remains local. For simplicity, we first prove this for single-qubit gate. The generalization to the two-qubit case is straightforward. Below we refer to this term as h_U for brevity.

Recall $h_U = \Lambda_{AB,CD}^{\otimes 2}(\mathrm{I} - |\Phi_U\rangle\langle\Phi_U|_{BC})\Lambda_{AB,CD}^{\otimes 2}$ acting on qubits A, B, C, D as shown below.



As in the main text, let $|\Phi_{\vec{p}}\rangle_{AB,CD} \triangleq |\Phi_{p_1}\rangle_{AB} |\Phi_{p_2}\rangle_{CD}$ and we will omit the system labels when they are clear from the context. The rotated term is of the form

$$
\begin{aligned}
V^\dagger h_U V &= V^\dagger \Lambda_{AB,CD}^{\otimes 2}(\mathrm{I} - |\Phi_U\rangle\langle\Phi_U|_{BC})\Lambda_{AB,CD}^{\otimes 2} V \\
&= V^\dagger \left(\sum_{\vec{p},\vec{q}\in\mathcal{P}^{\otimes 2}} \delta^{4-|(\vec{p},\vec{q})|} |\Phi_{\vec{p}}\rangle\langle\Phi_{\vec{q}}|_{AB,CD} \otimes \langle\Phi_{\vec{p}}| (\mathrm{I} - |\Phi_U\rangle\langle\Phi_U|_{BC}) |\Phi_{\vec{q}}\rangle \right) V \\
&= V^\dagger \left(\sum_{\vec{p},\vec{q}\in\mathcal{P}^{\otimes 2}} \delta^{4-|(\vec{p},\vec{q})|} |\Phi_{\vec{p}}\rangle\langle\Phi_{\vec{q}}|_{AB,CD} \otimes \left(\mathbb{1}_{\vec{p},\vec{q}} - \frac{1}{8}\mathrm{Tr}(p_1^* q_1^\top U^\dagger q_2^\top p_2^* U) \right) \right) V \\
&= \sum_{\vec{p}} \delta^{4-2|\vec{p}|} |\Phi_{\vec{p}}\rangle\langle\Phi_{\vec{p}}| \\
&\quad - \frac{1}{8} \sum_{\vec{P}\in\mathcal{P}^{\otimes n(\ell-1)}} \sum_{\vec{p},\vec{q}} \delta^{4-|(\vec{p},\vec{q})|} |\Phi_{\vec{p}}\rangle\langle\Phi_{\vec{q}}| \mathrm{Tr}(p_1^* q_1^\top U^\dagger q_2^\top p_2^* U) \otimes |\Phi_{\vec{P}}\rangle\langle\Phi_{\vec{P}}| \otimes (\widetilde{W}_{\vec{P}}^{<\ell})^\dagger p_1^\dagger U^\dagger p_2^\dagger q_2 U q_1 (\widetilde{W}_{\vec{P}}^{<\ell}),
\end{aligned}
$$

where $\widetilde{W}_{\vec{P}}^{<\ell} \triangleq W_{\ell-1}\tilde{P}_{\ell-1}\ldots W_1\tilde{P}_1$. Above, $\mathbb{1}_{\vec{p},\vec{q}}$ denotes the Kronecker delta symbol. The sum $\sum_{\vec{P}\in\mathcal{P}^{\otimes n(\ell-1)}}$ is over the Pauli noise $\vec{P}$ in layers preceding the gate U. We can also drop the complex conjugate "$*$" because $\mathcal{P} = \{I, X, XZ, Z\}$ are real matrices.

Since U is a Clifford operator, the second term above is nonzero if and only if $U^\dagger q_2^\top p_2^* U = \alpha p_1^* q_1^\top$ for $\alpha \in \{\pm 1, \pm i\}$. We denote $\vec{p} \overset{U}{\sim} \vec{q}$ if this is the case, leaving the phase α implicit. This notation suggests that the phase α does not show up in $V^\dagger h_U V$. Indeed, it can be verified that

$$
\begin{aligned}
p_1^\top U^\dagger p_2^\top q_2 U q_1 \,\mathrm{Tr}(p_1 q_1^\top U^\dagger q_2^\top p_2 U) &= p_1^\top(\alpha^* q_1 p_1^\top) q_1 \,\mathrm{Tr}(\alpha p_1 q_1^\top p_1 q_1^\top) \\
&= \alpha^*\alpha p_1^\top q_1 p_1^\top q_1 \,\mathrm{Tr}(p_1^\top q_1 p_1^\top q_1) \\
&= 2\mathrm{I}.
\end{aligned}
$$

Thus,

$$
\begin{aligned}
V^\dagger h_U V &= \sum_{\vec{p}\in\mathcal{P}^{\otimes 2}} \delta^{4-2|\vec{p}|} |\Phi_{\vec{p}}\rangle\langle\Phi_{\vec{p}}| - \frac{1}{4} \sum_{\vec{p}\overset{U}{\sim}\vec{q}} \delta^{4-|(\vec{p},\vec{q})|} |\Phi_{\vec{p}}\rangle\langle\Phi_{\vec{q}}| \\
&= \Lambda^{\otimes 2} \left(\sum_{\vec{p}\in\mathcal{P}^{\otimes 2}} |\Phi_{\vec{p}}\rangle\langle\Phi_{\vec{p}}| - \frac{1}{4} \sum_{\vec{p}\overset{U}{\sim}\vec{q}} |\Phi_{\vec{p}}\rangle\langle\Phi_{\vec{q}}| \right) \Lambda^{\otimes 2}.
\end{aligned}
$$

A completely similar analysis for two-qubit gates gives

$$
V^\dagger h_U V = \Lambda^{\otimes 4} \left(\sum_{\vec{p}\in\mathcal{P}^{\otimes 4}} |\Phi_{\vec{p}}\rangle\langle\Phi_{\vec{p}}| - \frac{1}{16} \sum_{\vec{p}\overset{U}{\sim}\vec{q}} |\Phi_{\vec{p}}\rangle\langle\Phi_{\vec{q}}| \right) \Lambda^{\otimes 4},
$$

where in this case $\vec{p} \overset{U}{\sim} \vec{q}$ means $U^\dagger(q_2^\top p_2) \otimes (q_4^\top p_4)U \propto (q_1^\top p_1) \otimes (q_3^\top p_3)$.