Hardness of Approximation for Ground State Problems

Sevag Gharibian
Department of Computer Science and Institute for Photonic Quantum Systems, Paderborn University, Germany

Carsten Hecht
Department of Computer Science, Paderborn University, Germany

Abstract

After nearly two decades of research, the question of a quantum PCP theorem for quantum Constraint Satisfaction Problems (CSPs) remains wide open. As a result, proving QMA-hardness of approximation for ground state energy estimation, analogous to hardness of approximation for MAX-k-CSP, has remained elusive. (QMA is Quantum Merlin-Arthur, a quantum generalization of NP with a quantum proof and quantum verifier.) Recently, it was shown [Bittel, Gharibian, Kliesch, CCC 2023] that a natural problem involving variational quantum circuits is QCMA-hard to approximate within ratio $N^{1-\epsilon}$ for any $\epsilon > 0$ and N the input size. (Quantum Classical Merlin-Arthur is QMA, but with a classical proof.) Unfortunately, this problem was not related to quantum CSPs, leaving the question of hardness of approximation for quantum CSPs open.

In this work, we show that if instead of focusing on ground state *energies* (analogous to the optimal number of satisfied clauses), one considers computing *properties* of the ground space (analogous to computing properties of the MAX-k-CSP solution space), QCMA-hardness of computing ground space properties can be shown. In particular, we show that it is (1) QCMA-complete within ratio $N^{1-\epsilon}$ to approximate the Ground State Connectivity problem (GSCON), and (2) QCMA-hard within the same ratio to estimate the amount of entanglement of a local Hamiltonian's ground state, denoted Ground State Entanglement (GSE). As a bonus, a simplification of our construction yields NP-completeness of approximation for a natural k-SAT reconfiguration problem, to be contrasted with the recent PCP-based PSPACE-hardness of approximation results for a different definition of k-SAT reconfiguration [Karthik C.S. and Manurangsi, 2023, and Hirahara, Ohsaka, STOC 2024].

2012 ACM Subject Classification Theory of computation → Problems, reductions and completeness; Theory of computation → Quantum complexity theory

Keywords and phrases Quantum complexity, hardness of approximation, local Hamiltonians, ground state connectivity, reconfiguration

Digital Object Identifier 10.4230/LIPIcs.MFCS.2026.48

Related Version *Full Version*: https://arxiv.org/abs/2411.04874

1 Introduction

Whether a quantum PCP theorem holds remains an important open question [1]. In this paper, we study the question – can one nevertheless obtain hardness of approximation for quantum complexity classes *without* a quantum PCP theorem? The answer is known to be *yes*. In 1999, Umans showed how to obtain classical hardness of approximation for the second level of the Polynomial Hierarchy, Σ_2^p, without utilizing a PCP theorem; rather, the construction utilized dispersers [32]. By extending this approach to the quantum setting, the first quantum hardness of approximation result for a quantum complexity class was shown: The quantum SUCCINCT SET COVER problem is hard to approximate for a quantum generalization of the second level of the Polynomial Hierarchy, cq-Σ_2 [12]. (cq-Σ_2 is a bounded-error quantum generalization of Σ_2^p, where the existentially quantified proof is

classical, and the universally quantified proof is quantum.) Ideally, however, one would like hardness of approximation for a quantum proof system without alternating quantifiers, such as QMA or one of its many variants (e.g. QCMA, QMA(2), etc. . . ; see [11] for a survey).

Here, we focus on Quantum Classical Merlin-Arthur (QCMA) [4], defined as QMA but with a classical proof. In [12], it was observed that a simple modification to the cq-Σ_2-hardness results therein also yields QCMA-hardness of approximation for an artificial problem, Quantum Monotone Minimum Satisfying Assignment (QMSA, Definition 10). Building on this, the first natural hardness of approximation result for QCMA was given [6], namely for the problem of estimating the optimal depth of a variational quantum circuit (MIN-VQA).

Local Hamiltonians, ground spaces, and GSCON. Unfortunately, MIN-VQA is not related to a quantum CSP. So, can one show QCMA-hardness of approximation for a natural computational problem for quantum CSPs? To answer this, recall first that a *k-local Hamiltonian* is an n-qubit $2^n \times 2^n$ complex Hermitian matrix H with a succinct representation $H = \sum_i H_i$, where analogous to a MAX-k-SAT clause acting on k out of n bits, each H_i is a Hermitian matrix or *clause* acting non-trivially on some subset of k out of n qubits. The problem k-LH then asks, given H, to estimate the *ground state energy*, i.e. the smallest eigenvalue $\lambda_{\min}(H)$. The corresponding set of optimal quantum assignments then form the *ground space*, i.e. the span of eigenvectors $|\psi\rangle \in \mathbb{C}^{2^n}$ with eigenvalue $\lambda_{\min}(H)$. The quantum CSP formulation of the quantum PCP conjecture [1] then posits that it is QMA-hard to decide if for positive semidefinite H, $\lambda_{\min}(H) = 0$ or $\lambda_{\min} \geq c$ for some fixed constant $c > 0$, under the promise that one of these cases holds. As stated above, however, resolving this conjecture remains a difficult challenge. Here, we show that by instead focusing on the natural problem of computing properties of the *ground space* (i.e. the *space* of optimal solutions), as opposed to the ground state energy (i.e the *value* attained by all optimal solutions), QCMA-hardness of approximation can be achieved.

Results. We now introduce the two and a "half" computational problems we study, and state our results. The first of these, GSCON, also has a natural classical counterpart from the classical study of *reconfiguration problems* [16], for which we show an analogous NP-hardness of approximation result; this is the "half" we refer to above.

Result 1: Ground State Connectivity (GSCON). Introduced in [13], Ground State Connectivity (Definition 6) is the problem of deciding if two ground states of H have a small circuit connecting them through the ground space. The input is a k-local Hamiltonian H, two ground states $|\psi\rangle$ and $|\phi\rangle$ (specified via quantum circuits), and a natural number m. The output is whether there exists a sequence of at most m 2-local unitary gates $(U_1, U_2, \ldots U_m)$, such that two properties hold:

1. (the unitary sequence maps $|\psi\rangle$ to $|\phi\rangle$) $U_m \cdots U_2 U_1 |\psi\rangle \approx |\phi\rangle$, and
2. (all intermediate states are ground states) for all $i \in [m]$, $|\psi_i\rangle := U_i \cdots U_1 |\psi\rangle$ is a ground state of H.

Besides the connection to classical reconfiguration problems, GSCON is physically motivated: Given Hamiltonian H and two ground states $|\psi\rangle$ and $|\phi\rangle$, it asks whether there is an "energy barrier" between the states? This question is important in the study of *quantum memories* and *stabilizer codes*. In quantum memories, the aim is to preserve the state $|\psi\rangle$ of a quantum system, often achieved by logically encoding the state within the ground space of a (gapped) Hamiltonian H. Since H is gapped, it is unlikely for noise to take $|\psi\rangle$ *out* of the ground space. However, this does not *a priori* prevent a noise process from mapping

one codeword to another *through* the ground space. Since noise is usually modeled as local operations, a desirable requirement is thus for the ground space to have an energy barrier between codewords, i.e. no short sequence of local operations should map $|\psi\rangle$ to another codeword. This is precisely the principle behind Kitaev's toy chain model [23] and toric code [24]. The application to quantum stabilizer error-correcting codes is similar. GSCON is QCMA-complete for $k = 5$ [13], surprisingly even for commuting Hamiltonians [1] [17]. In the 1D translation invariant setting, GSCON remains hard, being QCMAEXP-complete [34].

We now state our first main result, which shows that GSCON is QCMA-hard to *approximate*. For this, we reformulate GSCON to have two thresholds $m \leq m'$. We add the promise that there always exists a unitary sequence of length $\text{poly}(N')$ satisfying conditions (1) and (2) above. In the YES case, this sequence has length at most m, whereas in the NO case, any such sequence has length at least m'. That such a sequence always exists even in the NO case is important to ensure the approximation ratio m'/m is well-defined, for if no such unitary sequence satisfying (1) and (2) existed[2], the ratio can trivially be set to $m'/m \approx \infty$ by choosing arbitrarily large m' for the NO case.

▶ **Theorem 1.** *For all $\varepsilon > 0$ and $k \geq 5$, GSCON with k-local Hamiltonians is QCMA-complete for $\frac{m'}{m} \in \Theta(N'^{1-\epsilon})$, for N' the encoding size of the GSCON instance.*

In words, the length of the minimal unitary sequence satisfying conditions (1) and (2) is QCMA-hard to approximate, even within large relative error scaling essentially linearly in the input size, N'. Three comments: (1) The local Hamiltonians in Theorem 1 do not obey any particular geometry constraints. (2) As in the classical work of Umans [32], we write all hardness ratios in this paper relative to the *encoding size* of the input of the problem being reduced to. This is because, in general, the number of clauses can be *superlinear* in n. Thus, mapping an approximation ratio given relative to n to one relative to N' can yield an asymptotically smaller ratio. This is why simple promise gap amplification techniques, such as taking parallel copies of all clauses, typically do not suffice to achieve hardness of approximation ratios such as in Theorem 1. We remark[3], however, that for classical PCP-based MAX-SAT hardness of approximation results with constant hardness gap (as opposed to near-linear as in Theorem 1), hardness of approximation is sometimes instead stated relative to the value gap, the number of variables, or number of constraints.

Result 1.5: Classical analogue of GSCON, Boolean reconfiguration. The origins of GSCON as a computational problem are rooted in the classical study of *reconfiguration problems*, to which we now make a detour. First studied in [16], classical reconfiguration problems have since evolved into an entire research area [26]. In the case of k-SAT, the reconfiguration problem is specified as follows: Given an input k-SAT formula ϕ and satisfying assignments x and y, does there exist a sequence of bit flips from x to y, so that each intermediate string obtained is also satisfying for ϕ? This is the classical analogue of GSCON, with the exception that the input does not specify a bound m on the number of bit flips allowed – in the worst case, one may require an *exponential* number of bit flips. As a result, its complexity is more difficult than NP, being PSPACE-complete [16]. Recently, the first PCP theorems for reconfiguration problems have been established [20, 27, 29, 18], allowing for the first PSPACE-hardness of approximation results for many reconfiguration problems [29, 18].

[1] This is in contrast to the fact that k-LH for commuting Hamiltonians is *not* known to remain QMA-hard, and in fact is in NP for certain cases [8, 2, 30, 3, 19].
[2] Existing GSCON QCMA-hardness constructions indeed have no such sequence in the NO case.
[3] We thank an anonymous referee for pointing this out.

By leveraging our proof technique for GSCON from Theorem 1, we are able to give our own hardness of approximation result for k-SAT reconfiguration, albeit with respect to a different parameter than [29, 18]. We will first state our result, followed by a discussion of how it differs from [29, 18]. For this, we define the approximation problem Boolean Reconfiguration (BR, Definition 22), which is the k-SAT reconfiguration problem above, but with thresholds h and h' on the number of bit flips allowed in the YES and NO cases (analogous to GSCON), respectively.

▶ **Theorem 2.** *BR is NP-hard to approximate for $h'/h \in \Theta(N^{1-\varepsilon})$, for any constant $\varepsilon > 0$ and N the size of the input BR instance.*

The main differences between Theorem 1 and [29, 18] are now as follows. First, our proof does not rely on a PCP as in [29, 18], but uses Umans' disperser-based hardness gap construction as a starting point. Second, our hardness of approximation is with respect to the *length* of the reconfiguration sequence. In [29, 18], in contrast, there is no length parameter m. Rather, hardness of approximation is shown in the following sense: $\exists$ constant $\epsilon > 0$ such that, in the YES case, there exists a satisfying reconfiguration sequence from x to y, and in the NO case, for any reconfiguration sequence, there exists an intermediate string which satisfies at most a $(1-\epsilon)$-fraction of the clauses of ϕ, for some fixed constant $\epsilon > 0$. In this sense, Theorem 2 is complementary to [29, 18] in both its proof technique and actual result.

Result 2. Ground State Entanglement (GSE). We now return to the quantum setting, and define the second natural quantum problem we study, which asks whether a given local Hamiltonian has a ground state of low entanglement. More formally, we define the Ground State Entanglement problem (GSE, Definition 18), for which the input is a local Hamiltonian H, prespecified cut A versus A' of the qubits H acts on, and inverse polynomially separated thresholds $\eta_4 > \eta_3$. The output is to decide whether H has a ground state with entanglement entropy at most η_3 across the A versus A' cut, or whether all ground states have entanglement entropy at least η_4 across this cut? Here, the *entanglement entropy* is a standard entanglement measure, defined for a bipartite pure state $|\psi\rangle_{AA'}$ as $S(\rho_A)$, for reduced state $\rho_A := \mathrm{Tr}_{A'}(\rho_{AA'})$ and $S(\rho) := -\mathrm{Tr}(\rho \log \rho)$ the von Neumann entropy.

Our second main result is as follows.

▶ **Theorem 3.** *For all $\varepsilon > 0$ and $k \geq 5$, GSE is QCMA-hard for $\frac{\eta_4}{\eta_3} \in \Theta(N'^{1-\epsilon})$, for N' the encoding size of the GSE instance.*

In words, it is QCMA-hard to estimate the minimal entanglement entropy over all states in the ground space, even within large relative error N'. Two comments: First, unlike Theorem 1, it is not clear that GSE is also *in* QCMA. This is because verifying the entropy of a given quantum state $|\psi\rangle$ is, in general, at least as hard as Quantum Statistical Zero Knowledge (QSZK) [33]. In particular, when one has access to a poly-size quantum circuit preparing $|\psi\rangle$, the problem is QSZK-complete [5]. However, in GSE one does *not* have access to such a circuit, since $|\psi\rangle$ is an unknown ground state of H!

Second, there are two previous works [15, 7] we are aware of which study questions similar to GSE, whose relationship to Theorem 3 we now clarify. The first [15] defines the Hamiltonian Quantum Entropy Difference (HQED) problem, which takes two local Hamiltonians as input and a pre-specified cut A versus A', and asks: Among all ground states $|\psi_1\rangle$ of H_1 and $|\psi_2\rangle$ of H_2 of minimal entanglement entropy, which of $|\psi_1\rangle$ or $|\psi_2\rangle$ has larger entanglement entropy across A versus A'? Here, the promise gap in entropy is an additive constant (i.e. it is not a relative-error hardness of approximation result), and the hardness obtained is for the Learning With Errors (LWE) problem [28], not QCMA. In

words, Reference [15] shows that if one could efficiently estimate the entanglement entropy difference between ground states of two given Hamiltonians within constant additive error, then one could solve LWE.

The second relevant work [7] introduces the Learning Ground State Entanglement Structure (LGSES) problem. The question is then, roughly, to determine whether the ground states of a geometrically constrained H have volume law or area law entanglement across a given set of cuts. First, the differences: Like [15], the hardness results obtained in [7] are for LWE, not QCMA. Second, the input model to LGSES is not the standard one used in defining complexity classes, but rather has a cryptographic flavor. Namely, there are two families of computationally indistinguishable Hamiltonians, $\mathcal{H}_n^{\text{low}}$ and $\mathcal{H}_n^{\text{high}}$ (which we call "YES" and "NO" cases for reference), so that when H is drawn from either $\mathcal{H}_n^{\text{low}}$ or $\mathcal{H}_n^{\text{high}}$ according to an appropriate random distribution, with overwhelming probability the entanglement entropy will be low or high, respectively. In contrast, our input to GSE is a single Hamiltonian H along with a cut A versus A' (i.e. the standard QCMA input model). Now, the similarity: The promise gap for [7] on the entanglement entropy between the "YES" and "NO" cases is $O(\text{polylog}(n))$ versus $\Omega(n)$, i.e. a ratio of $(n/\text{polylog}(n))$. However, this gap is relative to the number of qubits n, not the encoding size of the input, N', as in Theorem 3. Our understanding[4] is that $N' \in \Omega(n^2)$ in [7], yielding ratio at most $\sqrt{N'}/\text{polylog}(N')$. Thus, relative to the input model therein, this may be viewed as an LWE-hardness of approximation result, albeit with a quadratically weaker ratio than Theorem 3.

Techniques for GSCON. We show a gap-preserving reduction from the artificial QCMA-hard to approximate problem Quantum Monotone Minimum Satisfying Assignment (QMSA, Definition 10), roughly defined as follows: The input is a quantum circuit $V = V_T \cdots V_1$ accepting a non-empty monotone set[5] $S \subseteq \{0,1\}^n$, and integer thresholds $0 \leq g \leq g' \leq n$. The output is to decide whether there is a low Hamming weight string $x \in \{0,1\}^n$, meaning of Hamming weight at most g, accepted by V, or if every $x \in \{0,1\}^n$ of Hamming weight at most g' is rejected by V, with the promise that one of these two cases holds.

Our starting point is the QCMA-hardness construction of [13], which we must briefly sketch (more details in Section 2). Namely, one first applies a *circuit-to-Hamiltonian* construction [22] to V, obtaining a local Hamiltonian H_{kit} encoding the action of V as follows: (1) $\lambda_{\min}(H_{\text{kit}})$ is "small" if and only if V accepts the string x it is given, and (2) the quantum state achieving this low energy is the *history state* of V, $|\psi_{\text{hist}}(x)\rangle := \frac{1}{\sqrt{T+1}} \sum_{t=0}^{T} V_t \cdots V_1 |x\rangle_B \otimes |0\rangle_C \otimes |t\rangle_D$, where x is the string fed to V in register B, C is the ancilla, and D is the clock register tracking time, t. To turn this into a GSCON instance, one then appends a 3-qubit "GO" register E, and sets the final Hamiltonian to $H = (H_{\text{kit}})_{BCD} \otimes P_E$ for $P := (I - |000\rangle\langle 000| - |111\rangle\langle 111|)_E$. Finally, the initial and target states are $|\psi\rangle = |0\cdots 0\rangle_{BCD}|000\rangle_E$ and $|\phi\rangle = |0\cdots 0\rangle_{BCD}|111\rangle_E$, respectively.

The intuition is now as follows: In the YES case, an honest prover can first prepare the history state $|\psi_{\text{hist}}(x)\rangle$ on registers BCD based on a Hamming weight g string x, obtaining $|\psi_{\text{hist}}(x)\rangle_{BCD}|000\rangle_E$, which is still in the null space of H. The goal is now to flip the GO qubits, but since we are restricted to 2-local gates, this must be done in two steps, e.g. the first two bits of E are first flipped to $|11\rangle$. At this point, our state has high overlap with P_E, "switching on" the Hamiltonian H_{kit}, which checks that our history state indeed has low energy. The last bit of E can now be flipped, and the history state uncomputed in order to

[4] Due to the multiplication of an $n \times n$ matrix A with n-bit string x in Definition 3.2 of the arxiv version [7].

[5] A set $S \subseteq \{0,1\}^n$ is *monotone* if for any $x \in S$, any strings x' obtained by flipping 0's to 1's in x is also in S.

arrive at target state $|\phi\rangle$. Soundness follows via the Traversal Lemma (Lemma 7), which shows that it is impossible to map $|000\rangle_E$ to $|111\rangle_E$ via 2-local gates without preparing an intermediate state with non-trivial overlap with P. And when this overlap occurs, H_{kit} will administer a large energy penalty to any history state prepared in BCD based on a string of Hamming weight at most g'.

At first glance, this construction already seems to give hardness of approximation when applied to V from QMSA – in the YES case, the history state is based on a low Hamming weight g string x, whereas in the NO case, any low-energy history state must encode a high Hamming weight g' string. Since $g'/g \in \Theta(N^{1-\epsilon})$, this suggests preparing the history state in the YES case requires much fewer gates than in the NO case, which translates into a short unitary sequence for GSCON in the YES case versus a long one in the NO case. The catch is that, while it takes g (respectively, g') bits to prepare x in the YES (respectively, NO) case, the cost of preparing the superposition in $|\psi_{\text{hist}}(x)\rangle$ given x depends on the number of gates T in V, which we have no control over! Thus, the approximation ratio achieved scales roughly as $(g+T)/(g'+T)$, which approaches 1 if T grows superlinearly in the proof size, which is in general the case.

To overcome this challenge, we begin with the basic GSCON setup outlined above, but move to a *triple* clock construction, along with taking inspiration from [6] to artificially "amplify the cost" of flipping each bit while preparing string x. Briefly, in addition to register $BCDE$ where B stores the n-bit string input to V, we (1) add two additional clock registers K and L of sizes $4n$ and 2, respectively, (2) add two registers F and G each of size n and intended to hold copies of proof x, and (3) finally an "amplification" register M. The basic premise is now that an honest prover proceeds in two phases. In the first phase, it manipulates clocks KL to be able to first prepare its desired proof x in B. The Hamiltonian constraints we design are such that there is a unique time step t_i in which bits B_i, F_i, and G_i can be acted on. At time t_i, if desired, the prover flips bits B_i, F_i and G_i, thus creating triple redundancy. It is then forced to flip *all* (roughly) T bits in the amplification register M from all zeroes to all ones, because the next time step t_i+1 will administer an energy penalty if B_i was set to 1 but the qubits of M still 0. It is now the triple redundancy on BFG that ensures that, once we leave timestep t_i, it is impossible to change bit B_i, because any single 2-local gate will break the equality constraints we place between B_i, F_i, and G_i on all time steps other than t_i. Next, moving to time t_i+2, the prover is allowed to uncompute M, and subsequently goes to time t_i+3, activating another Hamiltonian check that register M is correctly reset to all zeroes. Finally, we move to time t_i+4, and the entire process repeats for the next proof bit B_{i+1}. This first phase continues until we arrive at the final time step on clocks KL, which de-activates a constraint preventing the GO register E from being acted upon. In the second phase, the prover can now build a history state on registers BCD, and subsequently flip the GO qubits as in the basic setup to activate H_{kit}, which checks the history state. For clarity, it is the flipping of all T bits in the amplification register, M, which allows us to achieve our desired approximation ratio.

To make this logic sound, it is imperative for the clocks KL to be carefully designed. The clock sequence we construct for the first phase is $|00000\ldots0\rangle_K|00\rangle_L, |10000\ldots0\rangle_K|10\rangle_L$, $|11000\ldots0\rangle_K|11\rangle_L, |11100\ldots0\rangle_K|01\rangle_L \ldots$, which has three important properties enabling the soundness analysis to work: (1) Moving from a timestep t_i to t_{i+1} requires application of a 2-qubit unitary on a *unique* pair of qubits $q_{i,1}$ and $q_{i,2}$. (2) It is impossible to jump from t_i to (say) t_i+2 via a single 2-local gate. (3) To satisfy the YES case conditions of GSCON, every intermediate state $|\psi_l\rangle$ computed must have essentially all its amplitude on some *single* timestep t_{i_l} – for if not, any 2-qubit unitary attempting to increment the time from t_i to t_{i+1} will necessarily lead to non-trivial weight being placed on an *invalid* clock state, which is penalized.

Techniques for GSE. We now discuss Theorem 3 for estimating ground state entanglement, which again proceeds via a gap-preserving reduction from QMSA. Thus, given a circuit V which either accepts a low Hamming weight g proof or only high Hamming weight g' proofs, our goal is to construct a Hamiltonian with a low-energy state of low entanglement across a pre-specified cut in the YES case, or only highly entangled low-energy states in the NO case. The main idea is to add additional gates to V, each of which is controlled on a different proof qubit. Then, for each proof qubit set to 1, the corresponding added gate creates a Bell pair on a new register. This ensures that when there is a low (respectively, high) Hamming proof in the proof register, V prepares a low- (respectively, high-) entanglement proof across a certain cut. Applying Kitaev's circuit-to-Hamiltonian construction [22] now yields a local Hamiltonian whose history state is correspondingly entangled. This works as desired for the YES case, when the history state is also a low energy state of H_{kit}. However, the main technical hurdle is that in the NO case, we must show that *all* low energy states are highly entangled, not just the honest prover history state. For example, if the proof register contains a *superposition* over multiple high Hamming weight proofs (each of whose amplitudes can differ, and whose 1's can be in different positions), it is not as straightforward to argue that the corresponding ground state generated is highly entangled.

A natural first idea is to try the standard QCMA trick [35] of having V immediately measure its proof register upon reading it to destroy such a superposition, or equivalently immediately copy its proof to a fresh ancilla and apply the principle of deferred measurement. However, this principle requires all qubit copies in the output to be traced out at the end of the computation, which is indeed the case for a general QCMA verifier interested in just measurement statistics on its designated output qubit. In our case, in contrast, we wish to quantify entanglement of the *full* output state across our cut. To bypass this, we instead coherently copy the proof to *two* new registers, and do not assume any qubits of our circuit's output are traced out. Crucially, these registers will be on different sides of the cut for the entanglement entropy. Thus, when we consider a Schmidt decomposition for computing the entropy across this cut, it will depend on amplitudes corresponding to standard basis states which contain *both* a classical proof and some standard basis state on the Bell pair register. This allows us to reduce the analysis to the case where the proof register just contains *one* classical state.

A soundness analysis can now be run when the prover sends an arbitrary history state. However, when the prover sends an arbitrary low energy state, two more tricks are required. First, we weight the $H_{\text{in}} + H_{\text{prop}} + H_{\text{stab}}$ terms of H_{kit} with a large penalty term and apply the (Extended) Projection Lemma [21, 14] to argue that any low-energy state must be close to a history state. We then apply the Fannes inequality [10], which roughly says that states close in trace distance are also close in entanglement.

Open questions. First, in terms of GSCON, a curious fact is that it remains QCMA-complete even on commuting Hamiltonians [17]. Does QCMA-hardness of approximation also hold in this setting? For GSE, we have studied the estimation of ground state entanglement across a specified cut. Can one also show QCMA-hardness of approximation for detecting other entanglement structures, such as area law versus volume law entanglement as in the LWE-hardness results of [7]? Finally, are there other natural QCMA-hard to approximate problems? An excellent candidate is one of the first QCMA-complete problems, approximating the length of the minimum circuit preparing a ground state of a given local Hamiltonian (MIN-CIRCUIT) [35], under the additional promise that a poly-length preparation circuit exists. Can one show that this minimum circuit size is QCMA-hard to approximate?

2 Preliminaries

Complexity classes

▶ **Definition 4** (Quantum-Classical Merlin Arthur (QCMA) [4]). *A promise problem $\mathcal{A} = (A_{\text{yes}}, A_{\text{no}})$ is in QCMA if there exists a poly-time uniform quantum circuit family $\{V_n\}$ and polynomials $p, q : \mathbb{N} \to \mathbb{N}$ satisfying the following properties. For any input $x \in \{0,1\}^n$, V_n takes in $n + p(n) + q(n)$ qubits as input, consisting of the input x on register A, $p(n)$ qubits initialized to a classical proof $|y\rangle \in \{0,1\}^{p(n)}$ on register B, and $q(n)$ ancilla qubits initialized to $|0\rangle$ on register C. The first qubit of register C, denoted C_1, is the designated output qubit, a measurement of which in the standard basis after applying V_n yields the following:*

- *(Completeness) If $x \in A_{\text{yes}}$, $\exists$ proof $y \in \{0,1\}^{p(n)}$ that V_n accepts with probability $\geq 2/3$.*
- *(Soundness) If $x \in A_{\text{no}}$, then $\forall$ proofs $y \in \{0,1\}^{p(n)}$, V_n accepts with probability $\leq 1/3$.*

As is standard, throughout this work, we will assume the input x is hard-coded into the circuit for simplicity, and thus concern ourselves with registers B, C, and D.

Circuit-to-Hamiltonian constructions

We use the 5-local circuit-to-Hamiltonian construction of [22] as a black box. The precise details are not important, only the following properties. The input to the construction is a quantum circuit $V = V_T \cdots V_1$ (where in this work, each V_i comprising a circuit is at most 2-local) acting on a proof register B and ancilla register C. The output is a 5-local Hamiltonian $H = H_{\text{in}} + H_{\text{out}} + H_{\text{prop}} + H_{\text{stab}}$ acting on $B \otimes C \otimes D$, for D a clock register (encoded in unary). Such constructions are quantum generalizations of the Cook-Levin construction [9, 25], and aim to force low-energy/eigenvalue states of H to look like the quantum analogue of a tableau, the *history state*:

$$|\psi_{\text{hist}}(\phi)\rangle := \frac{1}{\sqrt{T+1}} \sum_{t=0}^{T} V_i \cdots V_1 |\phi\rangle_B \otimes |0\rangle_C \otimes |t\rangle_D, \tag{1}$$

for $|\phi\rangle$ an arbitrary proof in register B. For this, H_{in} forces the initial state at time $t = 0$ to be encoded correctly, H_{out} penalizes computations which reject at $t = T$, H_{prop} enforces correct propagation from each timestep t to $t + 1$, and H_{stab} ensures the clock D is correctly encoded. Moreover, $H_{\text{in}}, H_{\text{out}}, H_{\text{prop}}, H_{\text{stab}} \succeq 0$.

▶ **Lemma 5** (Kitaev [22]). *The construction of [22] maps a quantum circuit V to a 5-local Hamiltonian H_{kit} with parameters α and β such that:*

- *If there exists a proof $|\phi\rangle$ accepted by V with probability at least $1 - \epsilon$, then*

$$\text{Tr}(H_{\text{kit}} |\psi_{\text{hist}}(\phi)\rangle\langle\psi_{\text{hist}}(\phi)|) \leq \alpha := \frac{\epsilon}{T+1} \tag{2}$$

- *If V rejects all proofs $|\psi\rangle$ with probability at least $1 - \epsilon$, then the smallest eigenvalue of H_{kit} is at least $\beta := \frac{\pi^2(1-\sqrt{\epsilon})}{2(T+1)^3}$.*

To ensure $\beta - \alpha \geq 1/\text{poly}$, without loss of generality one first applies standard parallel repetition to V to reduce ϵ to at most an inverse polynomial.

Computational problems

We now define two computational problems we utilize, beginning with GSCON.

▶ **Definition 6** (Ground State Connectivity (GSCON) [13]). *Fix inverse polynomial* $\Delta : \mathbb{N} \to \mathbb{R}^+$.

- *Input:*
  1. *k-local Hamiltonian* $H = \sum_i H_i$ *acting on n qubits with* $H_i \in \mathrm{Herm}\left((\mathbb{C}^2)^{\otimes k}\right)$ *satisfying* $\| H_i \|_\infty \leq 1$.
  2. *Thresholds* η_1, η_2, η_3 *and* $\eta_4 \in \mathbb{R}$ *such that* $\eta_2 - \eta_1 \geq \Delta$ *and* $\eta_4 - \eta_3 \geq \Delta$, *and* $m, m' \in \mathbb{N}$.
  3. *Poly-size quantum circuits* U_ψ *and* U_ϕ *generating "starting" and "target" states* $|\psi\rangle$ *and* $|\phi\rangle$ *(starting from* $|0\rangle^{\otimes n}$*), respectively, satisfying* $\langle\psi|H|\psi\rangle \leq \eta_1$ *and* $\langle\phi|H|\phi\rangle \leq \eta_1$.
- *Output:*
  1. *If there exists a sequence of 2-qubit unitaries* $(U_i)_{i=1}^m \in \mathrm{U}\left(\mathbb{C}^2\right)^{\times m}$ *such that:*
     - **a.** *(Intermediate states remain in low energy space) For all* $i \in [m]$ *and intermediate states* $|\psi_i\rangle := U_i \cdots U_2 U_1 |\psi\rangle$, *one has* $\langle\psi_i|H|\psi_i\rangle \leq \eta_1$ *and*
     - **b.** *(Final state close to target state)* $\| U_m \cdots U_1|\psi\rangle - |\phi\rangle \|_2 \leq \eta_3$,

     then output YES.
  2. *If for all sequences of 2-qubit unitaries* $(U_i)_{i=1}^{m'} \in \mathrm{U}\left(\mathbb{C}^2\right)^{\times m'}$ *either:*
     - **a.** *(An intermediate state has high energy) There exists* $i \in [m']$ *and an intermediate state* $|\psi_i\rangle := U_i \cdots U_2 U_1 |\psi\rangle$, *that has energy* $\langle\psi_i|H|\psi_i\rangle \geq \eta_2$ *or*
     - **b.** *(Final state is far from target state)* $\| U_{m'} \cdots U_1|\psi\rangle - |\phi\rangle \|_2 \geq \eta_4$,

     then output NO.

As mentioned in Section 1, GSCON was shown QCMA-hard for $m = m'$ [13]. This was shown using the following traversal lemma, which we will also need in this work. Intuitively, it says that any sequence of local gates moving between two sufficiently separated subspaces must pass through the orthogonal complement.

▶ **Lemma 7** (Traversal Lemma [13]). *Let* $S, T \subseteq (\mathbb{C}^d)^{\otimes n}$ *be k-orthogonal subspaces. Fix arbitrary states* $|v\rangle \in S$ *and* $|w\rangle \in T$, *and consider a sequence of k-qubit unitaries* $(U_i) =_{i=1}^m$ *such that*

$$\| |w\rangle - U_m \cdots U_1 |v\rangle \| \leq \varepsilon \tag{3}$$

for some $0 \leq \varepsilon \leq 1/2$. *Define* $|v_i\rangle := U_i \cdots U_1 |v\rangle$ *and* $P := I - \Pi_S - \Pi_T$. *Then, there exists an* $i \in [m]$ *such that*

$$\langle v_i|P|v_i\rangle \geq \left(\frac{1-2\varepsilon}{2m}\right)^2. \tag{4}$$

The second computational problem we will leverage in our proofs is QMSA. For this, we require the following definitions from [12].

▶ **Definition 8** (Monotone set). *A set* $S \subseteq \{0,1\}^n$ *is called* monotone *if for any* $x \in S$, *any string obtained from x by flipping one or more zeroes in x to one is also in S.*

▶ **Definition 9** (Quantum circuit accepting monotone set). *Let V be a quantum circuit consisting of* 1- *and* 2*-qubit gates, which takes in an n-bit classical proof register, m-qubit ancilla register initialized to all zeroes, and outputs a single qubit, q. For any input* $x \in \{0,1\}^n$, *we say V* accepts *(respectively,* rejects*) x if measuring q in the standard basis yields* 1 *(respectively,* 0*) with probability at least* $1 - \delta$. *(If not specified,* $\delta = 1/3$.*) We say V accepts a* monotone set *if the set* $S \subseteq \{0,1\}^n$ *of all strings accepted by V is monotone (Definition 8).*

▶ **Definition 10** (Quantum Monotone Minimum Satisfying Assignment (QMSA)). *Given a quantum circuit V accepting a non-empty monotone set $S \subseteq \{0,1\}^n$, integer thresholds $0 \leq g \leq g' \leq n$, output:*

- *YES if there exists an $x \in \{0,1\}^n$ of Hamming weight at most g accepted by V.*
- *NO if every $x \in \{0,1\}^n$ of Hamming weight at most g' is rejected by V.*

The starting point for our gap-preserving reductions of Section 3 and Section 4 is:

▶ **Theorem 11** ([12]). *For all $\varepsilon > 0$, QMSA is QCMA-hard for $\frac{g'}{g} \in \Theta(N^{1-\varepsilon})$, where N is the input size.*

In the 3 following sections, we will state our main theorems and the corresponding proof sketches. The formal proofs can be found in the arxiv version.

3 Hardness of approximation for GSCON

We now show our first main result, which we restate for convenience:

▶ **Theorem 1.** *For all $\varepsilon > 0$ and $k \geq 5$, GSCON with k-local Hamiltonians is QCMA-complete for $\frac{m'}{m} \in \Theta(N'^{1-\epsilon})$, for N' the encoding size of the GSCON instance.*

Proof. Containment in QCMA is known [13]. We show QCMA-hardness of approximation via reduction from QMSA (Definition 10). We divide the proof into four steps: Constructing the GSCON instance from the QMSA instance (Section 3.1), showing completeness (Section 3.2), showing soundness (Section 3.3), and analyzing the approximation ratio (Section 3.4).

3.1 Construction

Let (V, g, g') be an instance of QMSA of encoding size N (i.e. requiring N bits to encode), with V taking in an n-qubit proof, and let $T = |V|$ (the number of 1- and 2-qubit gates comprising V). Without loss of generality, the completeness and soundness parameters for V are $1 - \varepsilon$ and ε for $\varepsilon \in \Theta(1/2^N)$ [12]. We also assume that the circuit immediately copies the proof to a different register by applying CNOT gates to ancilla qubits, and does not touch the actual proof register afterwards. We first map (V, g, g') to an instance of GSCON in Section 3.1.1. We then prove properties of the construction in Section 3.1.2 which we will use for completeness and soundness.

3.1.1 Details of construction

Basic initial setup. We begin with a basic setup similar to [13], as outlined in Section 2. Let H_{kit} be the Kitaev Hamiltonian (Lemma 5) constructed from V, and acting on proof register B of size n, ancilla C of size $q(n)$, clock D of size T. For clarity, register B is where the string $x \in \{0,1\}^n$ from Definition 10 is fed. Let $P := I - |000\rangle\langle 000| - |111\rangle\langle 111|$ be a projector acting on a 3-qubit GO register E. Let W be a circuit preparing the history state of the Kitaev Hamiltonian H_{kit}, given that B contains a string x. We note that W can be chosen to have size linear in N, the encoding size of (V, g, g'). Next we define the maximum number of 2-qubit unitaries permitted to prepare the target state in the YES and NO cases, respectively:

$$m := 2(2g + g\,|W| + 8n + |W| + 1), \tag{5}$$

$$m' := \frac{1}{2}g' \cdot |W|\,. \tag{6}$$

Recalling from Lemma 5 definitions $\alpha := \frac{\varepsilon}{N+1}$ and $\beta := \frac{\pi^2(1-\sqrt{\varepsilon})}{2(N+1)^3}$, we define the step and energy bounds for the GSCON instance:

$$\eta_1 := \alpha, \qquad \eta_2 := \min\left\{\frac{1}{m'^{13}}, \frac{\beta}{6 \cdot 64 m'^2}\right\}, \qquad \eta_3 := 0, \qquad \eta_4 := \frac{1}{4}. \tag{7}$$

Amplifying the cost of preparing each proof bit. To achieve our claimed hardness of approximation result, we henceforth deviate significantly from [13] in our Hamiltonian construction. In particular, we need to "amplify" the cost of preparing proof bits so as to "penalize" proofs of large Hamming weight. This first requires the addition of five additional registers:

- Recalling that B has size n, add two registers F and G of size n each.
- Add a second clock register K of size $4n$, and a 2-qubit register L.
- Add an amplification register M of size $|W|$.

We define the starting and final states as

$$|\psi\rangle = |0\rangle_{BCD}^{\otimes(n+q(n)+T)} |0\rangle_{E}^{\otimes 3} |0\rangle_{FGKLM}^{\otimes(2n+4n+2+|W|)} \tag{8}$$

$$|\phi\rangle = |0\rangle_{BCD}^{\otimes(n+q(n)+T)} |1\rangle_{E}^{\otimes 3} |0\rangle_{FGKLM}^{\otimes(2n+4n+2+|W|)}. \tag{9}$$

Finally, we add local Hamiltonian constraints as follows. In the full construction, we will have two types of constraints: The first type is for checking if a low energy state for the Hamiltonian H_{kit} has been prepared (*energy* checks), and the second to force the prover to prepare a state in a manner which requires the application of many local gates for high Hamming weight proofs (*Hamming weight checks*). We now discuss both types in detail.

Full set of constraints. Our full set of Hamiltonian constraints comprising H is:

1. (*Energy checks*) The local constraints comprising $H_{\text{kit}} \otimes P$ check that the string x in proof register B, used as the basis for constructing the history state for H_{kit}, is an accepting proof for the QMSA circuit V. We also add a new check so that only when the second clock, K, is set to its final timestep, $4n$, can one flip qubits in the GO register E:

$$P_E \otimes |0\rangle\langle 0|_{K_{4n}}. \tag{10}$$

   where we use notation K_{4s} to indicate the $4s^{\text{th}}$ qubit in register K.
2. (*Hamming weight checks*) These constraints ensure that preparing high Hamming weight proofs in B takes a "large" number of local unitaries. At a high level, in order to flip any particular proof bit, we force the prover to simulate 4 phases (corresponding to the four settings of clock register L in Definition 12):
   a. (Phase 00) Depending on the current time step, the prover will be allowed to flip one particular qubit of B, along with the copies of this qubit in F and G. Then, the prover is allowed to flip all qubits in the amplification register, M, from 0 to 1. (On other timesteps, modifying this particular proof qubit will not be possible due to 3-local equality constraints we will place between registers B, F and G.)
   b. (Phase 10) Check if all qubits in M are either all 0's or all 1's, and whether they match the prepared proof qubit in B.
   c. (Phase 11) The prover is allowed to flip all qubits in M from 1 to 0.
   d. (Phase 01) Check if all qubits in M are 0.

Additionally, we define constraints that force the prover to go through all clock states in K in the order we desire, and that force K to be in a correct clock state, meaning the clock qubits will be in a state of the form

$$|1\rangle^{\otimes k}|0\rangle^{\otimes 4n-k}. \tag{11}$$

Formally we can write these constraints as follows, which we partition into "classes" (a) through (f):

a. Clock register K has the right form:

$$|01\rangle\langle 01|_{K_i,K_{i+1}} \quad \forall i \in \{1,\ldots,4n-1\} \tag{12}$$

b. Clock register L has the right form, i.e. progresses through the 4 phases outlined earlier: For all $i \in \{1,\ldots,n\}$,

$$|0\rangle\langle 0|_{K_1} \otimes (I - |00\rangle\langle 00|)_L, \tag{13}$$
$$|10\rangle\langle 10|_{K_{4i-3},K_{4i-2}} \otimes (I - |10\rangle\langle 10|)_L, \tag{14}$$
$$|10\rangle\langle 10|_{K_{4i-2},K_{4i-1}} \otimes (I - |11\rangle\langle 11|)_L, \tag{15}$$
$$|10\rangle\langle 10|_{K_{4i-1},K_{4i}} \otimes (I - |01\rangle\langle 01|)_L, \tag{16}$$
$$|10\rangle\langle 10|_{K_{4i},K_{4i+1}} \otimes (I - |00\rangle\langle 00|)_L \quad \text{for } i < n, \tag{17}$$
$$|1\rangle\langle 1|_{K_{4n}} \otimes (I - |00\rangle\langle 00|)_L. \tag{18}$$

c. In phases 10 and 01, the amplification register M is either all 0's or all 1's:

$$|10\rangle\langle 10|_L \otimes \sum_{i=1}^{|M|-1} (|01\rangle\langle 01| + |10\rangle\langle 10|)_{M_i,M_{i+1}} \tag{19}$$
$$|01\rangle\langle 01|_L \otimes \sum_{i=1}^{|M|-1} (|01\rangle\langle 01| + |10\rangle\langle 10|)_{M_i,M_{i+1}}. \tag{20}$$

d. In phase 01, all amplification qubits are 0: For all $i \in \{1,\ldots,n\}$,

$$|10\rangle\langle 10|_{K_{4i-1},K_{4i}} \otimes |1\rangle\langle 1|_{M_1}. \tag{21}$$

e. The three proofs cannot be changed outside of exactly one time step (Definition 12): $\forall i \in \{1,\ldots,n-1\}$,

$$(I - |000\rangle\langle 000|_{B_1,F_1,G_1} - |111\rangle\langle 111|_{B_1,F_1,G_1}) \otimes (I - |00\rangle\langle 00|_{K_1,K_2}) \tag{22}$$
$$(I - |000\rangle\langle 000|_{B_{i+1},F_{i+1},G_{i+1}} - |111\rangle\langle 111|_{B_{i+1},F_{i+1},G_{i+1}}) \otimes (I - |10\rangle\langle 10|_{K_{4i},K_{4i+1}}). \tag{23}$$

f. After changing the proof qubits, they should be equal to the amplification qubits, i.e. $\forall i \in \{0,\ldots,n-1\}$,

$$|10\rangle\langle 10|_{K_{4i+1},K_{4i+2}} \otimes |1\rangle\langle 1|_{B_{i+1}} \otimes |0\rangle\langle 0|_{M_1} \tag{24}$$
$$|10\rangle\langle 10|_{K_{4i+1},K_{4i+2}} \otimes |0\rangle\langle 0|_{B_{i+1}} \otimes |1\rangle\langle 1|_{M_1}. \tag{25}$$

The Final Hamiltonian. Let H_{amp} denote the sum over all local terms corresponding to the energy checks and Hamming weight checks introduced above, except $H_{\text{kit}} \otimes P$. (Here, "amp" stands for amplification.) Recall that $H_{\text{kit}} = H_{\text{in}} + H_{\text{prop}} + H_{\text{out}} + H_{\text{stab}}$. The final Hamiltonian for our construction is

$$H := (\mu(H_{\text{in}} + H_{\text{prop}} + H_{\text{stab}}) + H_{\text{out}})_{BCD} \otimes P_E + (H_{\text{amp}})_{BEFGKLM} \tag{26}$$

for $\mu := 32\pi^4/\beta^2$. This factor μ is only used in a projection-lemma argument later (see the full arxiv version for that).

3.1.2 Properties of the construction

The constraints above allow certain operations, depending on how much weight we have on certain configurations of the clock registers. In particular, they only allow a strict subset of strings on clock registers K and L, which we call *timesteps*:

▶ **Definition 12** (Timestep t_i). *For $i \in \{0, \ldots, 4n\}$, define unary state*

$$|\widetilde{i}\rangle := |1\rangle^{\otimes i}|0\rangle^{\otimes 4n-i}. \tag{27}$$

Then, we define timestep t_i as standard basis state $|t_i\rangle := |\widetilde{i}\rangle_K|ab\rangle_L$, where

$$a = \begin{cases} 0, & \text{for } i = 0,3 \bmod 4 \\ 1, & \text{for } i = 1,2 \bmod 4 \end{cases} \qquad \text{and} \qquad b = \begin{cases} 0, & \text{for } i = 0,1 \bmod 4 \\ 1, & \text{for } i = 2,3 \bmod 4 \end{cases} \tag{28}$$

In words, this clock follows the sequence:

$$\begin{aligned} &|00000\ldots0\rangle_K|00\rangle_L = |\widetilde{0}\rangle|00\rangle_L, \\ &|10000\ldots0\rangle_K|10\rangle_L = |\widetilde{1}\rangle|10\rangle_L, \\ &|11000\ldots0\rangle_K|11\rangle_L = |\widetilde{2}\rangle|11\rangle_L, \\ &|11100\ldots0\rangle_K|01\rangle_L = |\widetilde{3}\rangle|01\rangle_L, \\ &|11110\ldots0\rangle_K|00\rangle_L = |\widetilde{4}\rangle|00\rangle_L, \ \ldots \end{aligned}$$

▶ **Observation 13** (k-orthogonality of time steps). *Any pair of timesteps $|t_i\rangle$ and $|t_j\rangle$ for $i < j$ are at least $|j - i + 1|$-orthogonal, due to the monotonically increasing Hamming weight in K with increasing i. Moreover, for any i, $|t_i\rangle$ and $|t_{i+1}\rangle$ are (precisely) 2-orthogonal.*

One last definition. In Section 3.3, we will need to argue that intermediate states (in the sense of Definition 6) have overlap with certain timesteps. We call this overlap *weight*:

▶ **Definition 14** (Weight on a state). *For a register S and string $x \in \{0,1\}^{|S|}$, we say that a state $|\psi\rangle$ has weight $w_\psi(x) \geq k$ on $|x\rangle$, if $|\psi\rangle$ can be written:*

$$|\psi\rangle = \alpha|x\rangle_S|\psi_x\rangle + \sum_{\substack{y \in \{0,1\}^{|S|} \\ y \neq x}} (\beta_y \cdot |y\rangle_S|\psi_y\rangle) \qquad \text{for} \quad |\alpha|^2 \geq k \tag{29}$$

for some unit vectors $|\psi_z\rangle$ with $z \in \{0,1\}^{|S|}$, and $|\alpha|^2 + \sum_y |\beta_y|^2 = 1$.

This completes the construction.

3.2 Completeness

Let us now formalize the intuition of the previous section. In the YES case, there exists a classical proof $x^* \in \{0,1\}^n$ of Hamming weight $\leq g$ that is accepted by V. When we talk about flipping a qubit, we will mean applying Pauli X to the corresponding qubit. We now construct a sequence of 2-local unitaries mapping $|\psi\rangle$ to $|\phi\rangle$ through the ground space of H, as per Definition 6:

In the YES case, let x^* be an accepted proof of Hamming weight at most g. Starting from $|\psi\rangle$, advance the KL clock through all timesteps. At the unique timestep for bit i, if $x_i^* = 1$, flip B_i, F_i, G_i, flip M to $|1\rangle^{|W|}$, and two timesteps later flip M back to $0^{|W|}$; otherwise only advance the clock. At the final timestep, prepare the history state on BCD using W, flip the GO register from $|000\rangle$ to $|111\rangle$ in two local moves, uncompute W, and reverse the clock/proof/amplification procedure. The only positive-energy intermediate state is the one activating the energy check, which has energy at most $\alpha = \eta_1$ by Lemma 5. All other intermediate states have 0 energy. The number of gates is

$$2(2g + g|W| + 8n + |W| + 1) = m.$$

3.3 Soundness

In the NO case, no classical proof of Hamming weight $\leq g'$ is accepted by V. The proof strategy will be to show that (1) at some point the Kitaev Hamiltonian H_{kit} must be "switched on", and (2) at this point the proof register will mostly be a superposition of low Hamming weight states. The main technical part of the proof will be to show that there exists a subset S of qubits on register B of size $\geq n - g'/2$, such that measuring any intermediate state (in the sense of Definition 6) on register S in the standard basis yields all zeros with probability at least $1/2$.

Intuitively, (1) follows from the Traversal Lemma. For (2), consider first the X/I-gate case: a 1 in B_i can only be introduced near timestep t_{4i-4}. To later reach the final clock region, the path must pass through the two amplification-check timesteps for bit i, forcing M from $0^{|W|}$ to $1^{|W|}$ and back. Thus each reliably prepared proof bit costs $\Omega(|W|)$ local operations. Lemma 15 makes this counting argument robust to arbitrary 2-qubit unitaries; Lemma 16 and Corollary 17 are the key clock claims used to prove it.

▶ **Lemma 15** (Low Hamming Weight Lemma). *Let $|\psi_j\rangle$ be a state with weight $\geq 1 - 1/m'^5$ on timesteps t_{4n} and t_{4n-1}. If we allow only m' many operations, all but $g'/2$ many qubits in B have overlap at most $1/m'^3$ with $|1\rangle$. Formally, there exists a subset S of at least $n - g'/2$ many qubits in B such that*

$$\forall\, i \in S, \qquad \mathrm{Tr}(|0\rangle\langle 0|_{B_i} |\psi_j\rangle\langle\psi_j|) \geq 1 - 1/(m')^3. \tag{30}$$

The main difficulty is that the prover may use arbitrary 2-qubit unitaries, and hence may place the clock registers in superposition. The next lemma shows that, nevertheless, any low-energy intermediate state is essentially supported on at most two consecutive legal timesteps. Moreover, transferring noticeable weight to a timestep t_{i+2} while retaining noticeable weight on t_i, creates illegal-clock weight and is therefore energetically penalized.

▶ **Lemma 16.** *If all intermediate states $|\psi_j\rangle$ are low energy, i.e. satisfy $\langle\psi_j|H|\psi_j\rangle \leq \eta_2$, then there is a constant c, such that for all $0 \leq j \leq m'$ there exists i with*

$$|\psi_j\rangle = \zeta|t_i\rangle_{KL} \otimes |\phi\rangle + \chi|t_{i+1}\rangle_{KL} \otimes |\phi'\rangle + \sum_{\substack{x \in \{0,1\}^{|K|+|L|} \\ x \neq t_i, t_{i+1}}} \gamma_x |x\rangle_{KL} \otimes |\phi_x\rangle \tag{31}$$

such that

$$\sum_{\substack{x\in\{0,1\}^{|K|+|L|}\\ x\neq t_i,t_{i+1}}} |\gamma_x|^2 < \frac{1}{(m')^c}. \tag{32}$$

Additionally, if $|\psi_j\rangle$ *has this form, and* $|\psi_{j+1}\rangle$ *has* $\Delta \in \mathbb{R}$ *more weight on* t_{i+2} *than* $|\psi_j\rangle$*, i.e.* $w_{\psi_{j+1}}(t_{i+2}) \geq w_{\psi_j}(t_{i+2}) + \Delta$ *then*

$$|\zeta|^2 \Delta \in O(\eta_2). \tag{33}$$

By setting η_2 small enough, we can make c an arbitrarily high constant. We note that for our analysis, $c = 5$ does suffice. A consequence of this is that whenever we need to move weight from one timestep t_i to a later timestep t_j, then we will have to have weight almost 1 on all intermediate timesteps:

▶ **Corollary 17.** *If an intermediate state* $|\psi_k\rangle$ *has weight* $\geq 1/3m'^4$ *on timestep* t_i *and we want to move weight* $\geq 1/3m'^4 - 1/m'^5$ *from* t_i *to* t_j *for* $j > i$*, then for all* $i < l < j$*, there exists intermediate state* $|\psi_{k_l}\rangle$ *with* $k_l > k$ *weight at least* $1 - 1/3m'^4$ *on timestep* t_l*. Let* k_l *be the first integer bigger than* k *where this happens. Then for* $i < l < l' < j$ *with corresponding states* $|\psi_{k_l}\rangle$ *and* $|\psi_{k_{l'}}\rangle$*, it holds that* $k < k_l < k_{l'}$*.*

Lemma 15 follows by applying Corollary 17 to each proof bit whose value is changed. To reach the final clock region, the path must pass through the amplification-check timesteps for that bit; during these timesteps the register M must be flipped from $|0\rangle^{|W|}$ to $|1\rangle^{|W|}$ and back. Thus each reliably prepared 1 in the proof costs $\Omega(|W|)$ local operations, so within m' operations at most $g'/2$ proof qubits can have noticeable overlap with $|1\rangle$.

Now we can finish the NO case. We are given the starting and final states

$$|\psi\rangle = |0\rangle^{\otimes(n+q(n)+N)}|0\rangle^{\otimes 3}|0\rangle^{\otimes(2n+4n+2+M)} \tag{34}$$

$$|\phi\rangle = |0\rangle^{\otimes(n+q(n)+N)}|1\rangle^{\otimes 3}|0\rangle^{\otimes(2n+4n+2+M)}. \tag{35}$$

By the traversal lemma 7, some intermediate state has inverse-polynomial weight on the active subspace of the GO register. The constraint $P_E \otimes |0\rangle\langle 0|_{K_{4n}}$, together with Lemma 16, implies that this state is essentially supported on the final timestep. Lemma 15 then implies that the proof register is supported with non-negligible weight on strings of Hamming weight at most $g'/2$. Since all such strings are rejected in the QMSA NO case, the H_{out} term gives energy at least η_2, contradicting the assumption that all intermediate states are low energy. ◀

3.4 Approximation ratio

Fix any $\epsilon > 0$. Recall that N is the encoding size of the QMSA instance with ratio $g'/g \in \Theta(N^{1-\epsilon})$, and let N' denote the encoding size of our constructed GSCON instance. We wish to show that $m'/m \in \Theta(N'^{1-\epsilon})$. For this, note first that since $|W| \in \Theta(N)$,

$$\frac{m'}{m} \in \Theta\left(\frac{\frac{1}{2}g'N}{4g + 2g|W| + 16n + 2N + 2}\right) \in \Theta\left(\frac{\frac{1}{2}g'}{2g + \Theta(1)}\right) \in \Theta(N^{1-\epsilon}). \tag{36}$$

Thus, the claim will follow if $N' \in \Theta(N)$. This, in turn, is true because the total number of constraints (each of constant size) present in H is $O(N)$. (Specifically, the H_{kit} terms contribute $O(N)$ terms, and each class from (a)-(f) contributes $\Theta(n) \in O(N)$ terms.)

4 Hardness of approximation for GSE

We next state Theorem 3, i.e. hardness of approximation for GSE. We begin by formally defining the latter.

▶ **Definition 18** (Ground State Entanglement(GSE)). *Fix an inverse polynomial $\Delta : \mathbb{N} \mapsto \mathbb{R}^+$. Denote the function for the von Neumann entropy as $\mathcal{S}$.*

- *Input:*
  1. *k-local Hamiltonian $H = \sum_i H_i$ acting on n qubits with $H_i \in \mathrm{Herm}\left((\mathbb{C}^2)^{\otimes k}\right)$ satisfying $\| H_i \| \leq 1$.*
  2. *Thresholds η_1, η_2, η_3 and $\eta_4 \in \mathbb{R}$ such that $\eta_2 - \eta_1 \geq \Delta$, $\eta_4 - \eta_3 \geq \Delta$ and $0 \leq \eta_3 \leq \eta_4 \leq \min\{|A|, |A'|\}$*
  3. *A partition (A, A') dividing the n qubits into two sets.*
- *Output:*
  1. *If there exists a state $|\psi\rangle \in (\mathbb{C}^2)^{\otimes n}$ with (low energy) $\langle\psi|H|\psi\rangle \leq \eta_1$ and (low entanglement) $\mathcal{S}(|\psi\rangle\langle\psi|_A) \leq \eta_3$, then output YES.*
  2. *If for all states $|\psi\rangle \in (\mathbb{C}^2)^{\otimes n}$, either (high energy) $\langle\psi|H|\psi\rangle \geq \eta_2$ or (high entanglement) $\mathcal{S}(|\psi\rangle\langle\psi|_A) \geq \eta_4$, then output NO.*

We restate Theorem 3 for convenience.

▶ **Theorem 3.** *For all $\varepsilon > 0$ and $k \geq 5$, GSE is QCMA-hard for $\frac{\eta_4}{\eta_3} \in \Theta(N'^{1-\epsilon})$, for N' the encoding size of the GSE instance.*

This is again shown by reduction from QMSA. We are given a QMSA circuit with an input x, and there either exists a low Hamming weight proof that is accepted, or only high hamming proofs could be accepted. The main idea is to extend the circuit given to us with gates that, controlled by the proof register, create Bell pairs on an additional register. Now we apply the Kitaev Hamiltonian construction. In the YES case, when there is a low Hamming weight proof in the proof register, these controlled gates will only create a small amount of entanglement, where as in the NO case all accepting proof will create states which are highly entangled. Using a few more technical tools like Fannes inequality [10], we can deduce our Theorem. The approximation ratio is derived analogously to the GSCON case.

5 NP-hardness of approximation

In this section, we state our hardness of approximation result for the classical analogue of GSCON, Boolean Reconfiguration. Again, this can be proven by reduction from a Hamming weight problem that is known to be hard to approximate:

▶ **Definition 19** (Monotone Minimum Satisfying Assignment (MMSA) [32]). *Given a monotone classical circuit C and two thresholds g, g', decide whether there exists an input of Hamming weight $\leq g$ that is accepted, or whether all accepted inputs have Hamming weight $\geq g'$.*

▶ **Theorem 20** ([32]). *MMSA is NP hard for $\frac{g'}{g} \in \Theta(N^{1/5-\epsilon})$.*

Combining this with the improved dispersers constructed in [31] immediately yields [32] the stronger result:

▶ **Theorem 21.** *MMSA is NP hard for $\frac{g'}{g} \in \Theta(N^{1-\epsilon})$.*

▶ **Definition 22** (Boolean Reconfiguration (BR)). *Given a k-SAT formula ϕ with n variables, two solutions $s, t \in \{0,1\}^n$ and thresholds h and h', does there exist a path of satisfying assignments $s = s_1, \ldots, s_h = t$, such that each s_i can be obtained from s_{i-1} by flipping a single bit, or does every such path have length $\geq h'$.*

▶ **Theorem 2.** *BR is NP-hard to approximate for $h'/h \in \Theta(N^{1-\varepsilon})$, for any constant $\varepsilon > 0$ and N the size of the input BR instance.*

Construction and proof are analogous to the quantum setting.

References

1 Dorit Aharonov, Itai Arad, and Thomas Vidick. Guest Column: The Quantum PCP Conjecture. *SIGACT News*, 44(2):47–79, 2013. doi:10.1145/2491533.2491549.

2 Dorit Aharonov and Lior Eldar. On the Complexity of Commuting Local Hamiltonians, and Tight Conditions for Topological Order in Such Systems. In *Proceedings of the 2011 IEEE 52nd Annual Symposium on Foundations of Computer Science*, FOCS '11, pages 334–343, USA, 2011. IEEE Computer Society. doi:10.1109/FOCS.2011.58.

3 Dorit Aharonov, Oded Kenneth, and Itamar Vigdorovich. On the Complexity of Two Dimensional Commuting Local Hamiltonians. In Stacey Jeffery, editor, *13th Conference on the Theory of Quantum Computation, Communication and Cryptography (TQC 2018)*, volume 111 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 2:1–2:21, Dagstuhl, Germany, 2018. Schloss Dagstuhl – Leibniz-Zentrum für Informatik. doi:10.4230/LIPIcs.TQC.2018.2.

4 Dorit Aharonov and Tomer Naveh. Quantum NP - A Survey, 2002. doi:10.48550/arXiv.quant-ph/0210077.

5 Avraham Ben-Aroya, Oded Schwartz, and Amnon Ta-Shma. Quantum Expanders: Motivation and Constructions. In *2008 23rd Annual IEEE Conference on Computational Complexity*, pages 292–303, 2008. doi:10.1109/CCC.2008.23.

6 Lennart Bittel, Sevag Gharibian, and Martin Kliesch. The Optimal Depth of Variational Quantum Algorithms Is QCMA-Hard to Approximate. In Amnon Ta-Shma, editor, *38th Computational Complexity Conference (CCC 2023)*, volume 264 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 34:1–34:24, Dagstuhl, Germany, 2023. Schloss Dagstuhl – Leibniz-Zentrum für Informatik. doi:10.4230/LIPIcs.CCC.2023.34.

7 Adam Bouland, Bill Fefferman, Soumik Ghosh, Tony Metger, Umesh Vazirani, Chenyi Zhang, and Zixin Zhou. Public-Key Pseudoentanglement and the Hardness of Learning Ground State Entanglement Structure. In *39th Computational Complexity Conference (CCC 2024)*. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2024. doi:10.4230/LIPIcs.CCC.2024.21.

8 Sergey Bravyi and Mikhail Vyalyi. Commutative version of the local Hamiltonian problem and common eigenspace problem. *Quantum Information & Computation*, 5(3):187–215, 2005. doi:10.26421/QIC5.3-2.

9 Stephen A. Cook. The complexity of theorem-proving procedures. In *Proceedings of the Third Annual ACM Symposium on Theory of Computing*, STOC '71, pages 151–158, New York, NY, USA, 1971. Association for Computing Machinery. doi:10.1145/800157.805047.

10 M. Fannes. A continuity property of the entropy density for spin lattice systems. *Commun. Math. Phys.*, 31(4):291–294, 1973. doi:10.1007/BF01646490.

11 Sevag Gharibian. Guest Column: The 7 faces of quantum NP. *ACM SIGACT News*, 54(4):54–91, 2024. doi:10.1145/3639528.3639535.

12 Sevag Gharibian and Julia Kempe. Hardness of Approximation for Quantum Problems. In Artur Czumaj, Kurt Mehlhorn, Andrew Pitts, and Roger Wattenhofer, editors, *Automata, Languages, and Programming*, Lecture Notes in Computer Science, pages 387–398, Berlin, Heidelberg, 2012. Springer. doi:10.1007/978-3-642-31594-7_33.

13 Sevag Gharibian and Jamie Sikora. Ground State Connectivity of Local Hamiltonians. In Magnús M. Halldórsson, Kazuo Iwama, Naoki Kobayashi, and Bettina Speckmann, editors, *Automata, Languages, and Programming (ICALP)*, pages 617–628, Berlin, Heidelberg, 2015. Springer. doi:10.1007/978-3-662-47672-7_50.

14 Sevag Gharibian and Justin Yirka. The complexity of simulating local measurements on quantum systems. *Quantum*, 3:189, 2019. doi:10.22331/q-2019-09-30-189.

15 Alexandru Gheorghiu and Matty J. Hoban. On estimating the entropy of shallow circuit outputs, 2024. doi:10.48550/arXiv.2002.12814.

16 Parikshit Gopalan, Phokion G. Kolaitis, Elitza Maneva, and Christos H. Papadimitriou. The Connectivity of Boolean Satisfiability: Computational and Structural Dichotomies. *SIAM Journal on Computing*, 38(6):2330–2355, 2009. doi:10.1137/07070440X.

17 David Gosset, Jenish C. Mehta, and Thomas Vidick. QCMA hardness of ground space connectivity for commuting Hamiltonians. *Quantum*, 1:16, 2017. doi:10.22331/q-2017-07-14-16.

18 Shuichi Hirahara and Naoto Ohsaka. Probabilistically Checkable Reconfiguration Proofs and Inapproximability of Reconfiguration Problems. In *Proceedings of the 56th Annual ACM Symposium on Theory of Computing*, STOC 2024, pages 1435–1445, New York, NY, USA, 2024. Association for Computing Machinery. doi:10.1145/3618260.3649667.

19 Sandy Irani and Jiaqing Jiang. Commuting Local Hamiltonian Problem on 2D beyond qubits, 2023. doi:10.48550/arXiv.2309.04910.

20 Takehiro Ito, Erik D. Demaine, Nicholas J. A. Harvey, Christos H. Papadimitriou, Martha Sideri, Ryuhei Uehara, and Yushi Uno. On the complexity of reconfiguration problems. *Theoretical Computer Science*, 412(12):1054–1065, 2011. doi:10.1016/j.tcs.2010.12.005.

21 Julia Kempe and Oded Regev. 3-local Hamiltonian is QMA-complete. *Quantum Information & Computation*, 3(3):258–264, 2003. doi:10.26421/QIC3.3-7.

22 A. Kitaev, A. Shen, and M. Vyalyi. *Classical and Quantum Computation*, volume 47 of *Graduate Studies in Mathematics*. American Mathematical Society, 2002. doi:10.1090/gsm/047.

23 A. Yu Kitaev. Unpaired Majorana fermions in quantum wires. *Physics-Uspekhi*, 44(10S):131, 2001. doi:10.1070/1063-7869/44/10S/S29.

24 A. Yu. Kitaev. Fault-tolerant quantum computation by anyons. *Annals of Physics*, 303(1):2–30, 2003. doi:10.1016/S0003-4916(02)00018-0.

25 Leonid Levin. Universal search problems. *Problems of Information Transmission*, 9(3):265–266, 1973.

26 Naomi Nishimura. Introduction to Reconfiguration. *Algorithms*, 11(4):52, 2018. doi:10.3390/a11040052.

27 Naoto Ohsaka. Gap Preserving Reductions Between Reconfiguration Problems. In *40th International Symposium on Theoretical Aspects of Computer Science (STACS 2023)*. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2023. doi:10.4230/LIPIcs.STACS.2023.49.

28 Oded Regev. On lattices, learning with errors, random linear codes, and cryptography. *J. ACM*, 56(6):34:1–34:40, 2009. doi:10.1145/1568318.1568324.

29 Karthik C. S. and Pasin Manurangsi. On Inapproximability of Reconfiguration Problems: PSPACE-Hardness and some Tight NP-Hardness Results, 2024. doi:10.48550/arXiv.2312.17140.

30 Norbert Schuch. Complexity of commuting Hamiltonians on a square lattice of qubits. *Quantum Information & Computation*, 11(11-12):901–912, 2011. doi:10.26421/QIC11.11-12-1.

31 A. Ta-Shma, C. Umans, and D. Zuckerman. Lossless Condensers, Unbalanced Expanders, and Extractors. *Combinatorica*, 27(2):213–240, 2007. doi:10.1007/S00493-007-0053-2.

32 C. Umans. Hardness of approximating /spl Sigma//sub 2//sup p/ minimization problems. In *40th Annual Symposium on Foundations of Computer Science (Cat. No.99CB37039)*, pages 465–474, 1999. doi:10.1109/SFFCS.1999.814619.

33 J. Watrous. Limits on the power of quantum statistical zero-knowledge. In *The 43rd Annual IEEE Symposium on Foundations of Computer Science, 2002. Proceedings.*, pages 459–468, 2002. doi:10.1109/SFCS.2002.1181970.

34 James D. Watson, Johannes Bausch, and Sevag Gharibian. The Complexity of Translationally Invariant Problems Beyond Ground State Energies. In Petra Berenbrink, Patricia Bouyer, Anuj Dawar, and Mamadou Moustapha Kanté, editors, *40th International Symposium on Theoretical Aspects of Computer Science (STACS 2023)*, volume 254 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 54:1–54:21, Dagstuhl, Germany, 2023. Schloss Dagstuhl – Leibniz-Zentrum für Informatik. doi:10.4230/LIPIcs.STACS.2023.54.

35 Pawel Wocjan and Shengyu Zhang. Several natural BQP-Complete problems, 2006. doi:10.48550/arXiv.quant-ph/0606179.