The Complexity of Stoquastic Local Hamiltonian Problems

Sergey Bravyi * David P. DiVincenzo † Roberto Oliveira ‡ Barbara M. Terhal §

January 8, 2007

Abstract

We study the complexity of the Local Hamiltonian Problem (denoted as LH-MIN) in the special case when a Hamiltonian obeys the condition that all off-diagonal matrix elements in the standard basis are real and non-positive. We will call such Hamiltonians, which are common in the natural world, *stoquastic*. An equivalent characterization of stoquastic Hamiltonians is that they have an entry-wise non-negative Gibbs density matrix for any temperature. We prove that LH-MIN for stoquastic Hamiltonians belongs to the complexity class AM — a probabilistic version of NP with two rounds of communication between the prover and the verifier. We also show that 2-local stoquastic LH-MIN is hard for the class MA. With the additional promise of having a polynomial spectral gap, we show that stoquastic LH-MIN belongs to the class PostBPP=$\mathrm{BPP}_{\mathrm{path}}$— a generalization of BPP in which a post-selective readout is allowed. This last result also shows that any problem solved by adiabatic quantum computation using stoquastic Hamiltonians is in PostBPP.

1 Introduction

For the last few years significant progress has been made in understanding the computational complexity of spin Hamiltonian problems. This area of research is of great importance for physics, since most strongly interacting quantum many-body systems can not be fully analyzed by analytical methods; thus, we can only hope to understand their properties from numerical simulations. A system is efficiently simulatable if the computational resources one needs for simulation grow only polynomially with the number of spins in the system. For example, one-dimensional spin chains with a small amount of entanglement can be simulated by the DMRG method and its recent generalizations to matrix product states [1, 2, 3]. It has been proposed that systems of interacting bosons, like those described by the bosonic Hubbard model, can be simulated using the Green's function Monte-Carlo technique, see [4, 5]. It is believed that a quantum computer will offer more possibilities to simulate quantum systems. Understanding the computational complexity of spin Hamiltonian problems might help to identify classes of Hamiltonians for which efficient classical or quantum simulation algorithms could be developed.

We shall consider the Local Hamiltonian Problem defined in [6, 7]. A k-local n-qubit Hamiltonian is a Hermitian operator H acting on $(\mathbb{C}^2)^{\otimes n}$ that can be expressed as a sum of k-qubit interactions: $H=\sum_S H_S$. Here $S\subseteq\{1,\ldots,n\}$ runs over all subsets of qubits of cardinality k and H_S may be an arbitrary Hermitian operator on S tensored with the identity on all qubits from $\{1,\ldots,n\}\backslash S$. A natural unit of energy set by H is given by the maximum operator norm of the interactions, $J=\max_S ||H_S||$. Let $\lambda(H)$ be the smallest eigenvalue of H, i.e. the ground-state energy. Suppose we are promised that either $\lambda(H)\le 0$ or $\lambda(H)\ge\delta$, where δ is at least $\frac{J}{\mathrm{poly}(n)}$. The Local Hamiltonian Problem is formulated as a decision problem: given the data $(n,\{H_S\},\delta)$, one has to decide whether $\lambda(H)\le 0$. A more formal definition is given in Section 2.1. We will refer to the Local Hamiltonian problem as LH-MIN indicating that it is the problem of estimating the minimum eigenvalue of H.

*IBM Watson Research Center, Yorktown Heights, NY, USA 10598. sbravyi@us.ibm.com
†IBM Watson Research Center, Yorktown Heights, NY, USA 10598. divince@watson.ibm.com
‡IBM Watson Research Center, Yorktown Heights, NY, USA 10598. rob.oliv@gmail.com
§IBM Watson Research Center, Yorktown Heights, NY, USA 10598. bterhal@gmail.com

If one considers a generic spin Hamiltonian H that lacks any additional structure except for the locality of interactions, it is extremely unlikely that LH-MIN can be solved in polynomial time (even on a quantum computer). Indeed, it was shown by Kitaev [6] that LH-MIN is a complete problem in the complexity class QMA — the quantum analogue of NP. This QMA-completeness result applies even to Hamiltonians with 2-qubit nearest-neighbor interactions on 2D square lattice [7, 8]. Therefore, instead of looking for efficient algorithms for evaluating the ground-state energy, we have to focus on efficient *proving protocols* by which the prover (a party with unlimited computational power) can prove an upper bound on the ground-state energy to the verifier (a party that has polynomial resources).

By definition, the inclusion LH-MIN $\in$ QMA means that the upper bound $\lambda(H) \le 0$ has an efficient quantum proving protocol with one round of communication between the prover and the verifier, see [10]. One of the goals of the present paper is to argue that there exists a large subclass of quantum local Hamiltonians for which LH-MIN has an efficient *classical* proving protocol with a constant number of communication rounds. This subclass involves all local spin Hamiltonians whose matrix elements in the standard basis of n qubits the condition that *all off-diagonal matrix elements are real and non-positive.* A nice property of such Hamiltonians is that the corresponding Gibbs density matrix $\rho = e^{-\beta H}/\,\mathrm{Tr}\left(e^{-\beta H}\right)$ has non-negative matrix elements in the standard basis for any $\beta \ge 0$. From this non-negativity property of the Gibbs matrix it follows by simple linear algebra arguments that the ground-state $|\Psi_0\rangle$ of H has non-negative real coefficients, i.e. $|\Psi_0\rangle = \sum_i \alpha_i |i\rangle$ where $\alpha_i \ge 0$. Thus one can associate a probability distribution with the ground-state, $\mathbb{P}(i) = \frac{\alpha_i}{\sum_i \alpha_i}$. If one is able to sample efficiently from this distribution one can determine $\lambda(H)$ (for details see Section 6). Because of the relation to stochastic processes we have adopted the term *stoquastic* to refer to these Hamiltonians. In the 'standard' basis, these Hamiltonians have non-positive off-diagonal matrix elements. This standard basis for local Hamiltonians is typically the local spin-z basis, but one can of course allow for local unitary basis changes without changing the complexity of the problem.

Clearly, any classical spin Hamiltonian, i.e. a Hamiltonian which is diagonal in the standard basis, falls into the stoquastic class. Here are some 1-local and 2-local stoquastic operator on qubits:

$$-X, \quad -X \otimes |z\rangle\langle z| \quad \text{for} \quad z \in \{0,1\}, \quad -p\, X \otimes X - q\, Y \otimes Y \quad \text{for any} \quad 0 \le q \le p.$$

It can be shown that all 2-local stoquastic Hamiltonians on qubits can be generated by taking convex linear combinations of these stoquastic 2-local interactions and all classical 2-local interactions (composed solely from tensor products of Z) [1].

Stoquastic Hamiltonians are very common in physics. Among spin-1/2 models, the well-studied ferromagnetic Heisenberg models and the quantum transverse Ising model (considered for example by Farhi [11] in the context of adiabatic quantum computation) are stoquastic. Another example is a Heisenberg antiferromagnet on a cubic lattice (or more generally, on a bipartite graph):

$$H = \sum_{(j,k)} X_j \otimes X_k + Y_j \otimes Y_k + Z_j \otimes Z_k.$$

Here the qubits live at vertices of the lattice and the interactions couple nearest-neighbors on the lattice. Although H is not directly stoquastic, it can be simply made so by a local change of basis. Indeed, if a lattice admits a bi-coloring, one can apply Z to every white vertex to flip the sign of $X \otimes X$ and $Y \otimes Y$ [2]. This produces a stoquastic Hamiltonian.

Although in this paper we focus only on spin-1/2 Hamiltonians, the stoquastic class naturally extends to systems of qudits, or even infinite-dimensional particles (e.g. harmonic oscillators). For example, a system of spin-less interacting bosons is described (in the first quantization formalism) by a Hamiltonian $H = K + U$, where $K = -\frac{1}{2m}\sum_a \Delta_a$ is a kinetic energy (when the vector potential is zero) and U is a potential energy. Off-diagonal matrix elements of H come only from K. The discretized version of the Laplacian, $\Delta_a = \frac{d^2}{dx_a^2} = \sum_j |j+1\rangle\langle j| + |j\rangle\langle j+1| - 2|j\rangle\langle j|$ shows that all off-diagonal matrix elements of K are non-positive. Outstanding examples in this category are bosonic Bose-Einstein condensates and Helium-4 [12]; there is a general belief in the computational physics community that the ground-state properties of such systems are "easy" to simulate, although no rigorous basis for this opinion seems to exist presently.

[1] It can be shown that there are 3-local Hamiltonians on qubits which are stoquastic, but not *termwise* stoquastic, i.e. they cannot be written as a sum over stoquastic terms that acts on 3 qubits at the time.

[2] The new basis coincides with the original one up to phases of the basis vectors.

All Josephson-junction qubit systems of the 'flux'-type are stoquastic. The quantum-mechanics of any such system is that of a collection of *distinguishable* (rather than bosonic or fermionic) particles with a Hamiltonian $K+U$ as just discussed [13]. It was this observation that initiated the present investigation, and indicated that flux qubits would not be the most general choice for implementing adiabatic quantum computation.

Other stoquastic Hamiltonians are identified by noting that bosonic creation/annihilation operators $\hat{a}\,|j\rangle = \sqrt{j}\,|j-1\rangle$ and $\hat{a}^\dagger\,|j\rangle = \sqrt{j+1}\,|j+1\rangle$ have non-negative matrix elements in the occupation number basis. Therefore a hopping operator $-\hat{a}_j^\dagger\,\hat{a}_k - \hat{a}_k^\dagger\,\hat{a}_j$, and the entire class of bosonic Hubbard models, belongs to the stoquastic class. Among systems involving both spin-1/2 and bosonic degrees of freedom, the Jaynes-Cummings model [14], and the spin-boson model [15], are also stoquastic when suitable phases are associated with the vectors in the standard basis.

Naturally, not *all* Hamiltonians in physics are stoquastic. Many fermionic systems are non-stoquastic; the antisymmetry of the (first-quantized) wavefunction causes it to have sign changes in the position basis. In the occupation-number (second-quantized) basis, terms of both signs typically occur as off-diagonal matrix elements on account of the anticommutation relations of the creation and annihilation operators. Special fermionic systems, like the spin systems mentioned above, can avoid this 'sign problem' but generic fermionic systems do not. Hamiltonians of charged (bosonic or fermionic) particles in the presence of a magnetic field will also not be stoquastic (the Hamiltonian, and the ground-state are typically complex).

Stoquastic Hamiltonians have also featured in recent work in quantum information theory. In Ref. [16] they are used to define an adiabatic path algorithm that is derived from a classical reversible Markov chain and in Ref. [9] they are similarly defined on the basis of a Monte-Carlo process that generates the equilibrium distribution of some classical Hamiltonian. In these constructions, there is a direct connection between the rapid convergence of the Markov chain and the gap of the resulting stoquastic Hamiltonian. In some sense these constructions, and our results, are rigorous expressions and examples of the physics folklore theorem which says that one can map ground-state problems of d-dimensional Hamiltonians onto classical statistical problems in $d+1$-dimensions [18]. In this paper we show in fact that if some rigorously defined version of this folklore statement were true than it would have the complexity-theoretic consequence that QMA $\subseteq$ AM, which we consider unlikely. Thus as it stands, it is only the class of stoquastic Hamiltonians that allow for this quantum-to-classical mapping.

2 Summary of Main Results

Let us review our main results. Obviously, restricting ourselves to a subclass of local Hamiltonians can only reduce the complexity of LH-MIN which means that stoquastic LH-MIN belongs to the class QMA. On the other hand, stoquastic LH-MIN is NP-hard, since it includes all classical local Hamiltonians. Indeed, it was proved by Barahona [19] that finding the ground-state energy of the Ising model on the 3D cubic lattice with couplings $J \in \{-1,0,+1\}$ is a NP-complete problem.

Firstly, we prove that stoquastic LH-MIN belongs to the complexity class AM. AM is a probabilistic analogue of NP with two rounds of communication between the prover and the verifier, see Section 3. The proof proceeds by mapping stoquastic LH-MIN to the Approximate Set Size problem. We consider a "partition function" $Z = \mathrm{Tr}\,(G^L)$, where $G = I - \beta\,H$ is a non-negative matrix whose largest eigenvalue is $\mu = 1 - \beta\,\lambda(H)$. If L is a sufficiently large, $Z \approx \mu^L$ and thus Z provides enough information about $\lambda(H)$. Then we convert G into a sum of $0,1$-matrices thus expressing Z as a sum of a Boolean function over all input arguments. Evaluating this sum is equivalent to the Approximate Set Size problem. The latter problem admits a two-round interactive proof based on Carter-Wegman universal hashing, see [20, 21]. It should be noted that AM also contains a generalization of stoquastic LH-MIN in which G may be an arbitrary non-negative matrix specified by a black box. In a sequel to this paper [22] we will strengthen this result and prove that stoquastic LH-MIN is in a class called SBP.

Secondly, we show that the 6-local stoquastic Hamiltonian problem is hard for the class MA — the probabilistic analogue of NP(see Section 4 for details). The main idea of the proof is that any classical probabilistic machine can be simulated by a classical circuit C with *reversible* gates whose input include ancillary random bits. Such a circuit can be transformed into a coherent form U_C by replacing each gate with a unitary operator (which just permutes basis vectors) and replacing each random bit with a coherent superposition $(|0\rangle + |1\rangle)/\sqrt{2}$. Making use of the standard clock Hamiltonian construction [6] we can define a local Hamiltonian H whose ground-state energy is related to the maximum acceptance probability of the quantum circuit U_C. The condition that U_C is composed only of classical gates guarantees that H is an stoquastic Hamiltonian. We then prove that allowing Merlin to feed quantum states into the verifying circuit does not give him any additional cheating power as compared to the classical case.

Thirdly, we prove that for any constant k k-local stoquastic LH-MIN can be reduced in polynomial time to 2-local stoquastic LH-MIN. The proof is based on perturbation theory gadgets introduced in [7]. We construct a new three-qubit gadget that involves only stoquastic interactions, see Section 5 for details. A corollary of this result is that 2-local stoquastic LH-MIN is hard for MA. The fact that the complexity of k-local stoquastic LH-MIN does not depend upon k indicates that this problem might be complete for some well-defined computational class, even though the nature of this class remains elusive to us.

Finally, we consider a special case of stoquastic LH-MIN in which the Hamiltonian has a polynomial spectral gap (the difference between the smallest and the second smallest eigenvalue is $1/\mathrm{poly}(n)$ for some polynomial in n), see Section 6 for details. In this case we prove that stoquastic LH-MIN belongs to the class PostBPP — the class of languages recognizable by poly-time probabilistic Turing machines which produce the correct answer (with constant error probability) conditioned on the value of a 'success flag' bit (the success probability may be exponentially small though). The proof relies on the ideas borrowed from the Green's Function Quantum Monte Carlo method, see [5] and we show that post-selected classical computation gives us the power to sample from the ground-state distribution. This last result also implies that any decision problem solved by an adiabatic quantum algorithm that uses only stoquastic Hamiltonians is contained in PostBPP.

2.1 Definition of the Local Hamiltonian Problem

We shall denote the smallest eigenvalue of a Hamiltonian H by $\lambda(H)$.

Definition 1 *For any integer k, polynomials $p_1(n)$ and $p_2(n)$ define a set $\Omega(k, p_1, p_2)$ involving all k-local n-qubit Hamiltonians $H = \sum_S H_S$ such that for any fixed $k \le n < \infty$ one has*

- $||H_S|| \le p_1(n)$ *for all subsets* $S \subseteq \{1, \ldots, n\}$, $|S| = k$
- *Either* $\lambda(H) \le 0$ *or* $\lambda(H) \ge 1/p_2(n)$

Suppose we are given a Hamiltonian $H \in \Omega(k, p_1, p_2)$ and our goal is to decide whether $\lambda(H) \le 0$. Clearly, the correct decision can be made even if the interactions H_S are specified only up to some precision δ polynomial in $1/n$. Indeed, if Hamiltonians H and H' are ϵ-close in the operator norm, $||H - H'|| < \epsilon$, then their ground-state energies are also ϵ-close, $|\lambda(H) - \lambda(H')| < \epsilon$ (see for example [23]). Thus, although $\Omega(k, p_1, p_2)$ is a continuum set, we can safely assume that any $H \in \Omega(k, p_1, p_2)$ is described by $\mathrm{poly}(n)$ bits. In that sense we can regard $\Omega(k, p_1, p_2)$ as a set of finite binary strings.

Definition 2 (Local Hamiltonian Problem (LH-MIN)) *Given a description of a Hamiltonian $H \in \Omega(k, p_1, p_2)$, decide whether $\lambda(H) \le 0$.*

3 Stoquastic LH-MIN in AM

The complexity class AM was introduced originally by Babai [24] as a class of decision problems that possess a randomized interactive proof with two-way communication between the prover (Merlin) having unlimited computational resources and the verifier (Arthur) capable of doing only polynomial-time computation. It is a remarkable property of the class AM that any proving protocol with constant number of communication rounds[3] can be simulated by a protocol with just two rounds [24], such that the first message is sent from Arthur to Merlin, and the second one backwards.

[3] A communication round involves a single message sent from one party to the other.

We shall mostly consider promise problems. Let $\Sigma=\{0,1\}$ and let Σ^n be a set of n-bit strings and Σ^* be a set of all finite binary strings. A promise problem can be regarded as a set of binary strings $L\subseteq\Sigma^*$ partitioned into two non-overlapping subsets, $L=L_{\rm yes}\cup L_{\rm no}$, corresponding to positive and negative instances. An Arthur-Merlin proving protocol for a membership $x\in L_{\rm yes}$ involves Arthur's question $q\in\Sigma^{p(|x|)}$ and Merlin's response $r\in\Sigma^{p(|x|)}$, where p is a fixed polynomial and $|x|$ is the number of bits in x. Arthur's question is just a random bit string drawn from the uniform distribution. The response r may be an arbitrary function of x and q. Once the communication is completed, Arthur has at his disposal all the data x,q,r. Then he runs a BPP test $V(x,q,r)$ that outputs either 1 (accept the proof) or 0 (reject the proof).

A proving protocol must obey soundness and completeness properties. Completeness means that for positive instances, $x\in L_{yes}$, Merlin has a strategy (i.e. a response functions $r(x,q)$) for which Arthur's acceptance probability is close to 1. Soundness means that for negative instances, $x\in L_{\rm no}$, Arthur's acceptance probability is close to 0 for all possible Merlin's strategies. Here is a formal definition[4]:

Definition 3 *A promise problem $L=L_{\rm yes}\cup L_{\rm no}\subseteq\Sigma^*$ belongs to the class* AM *iff there exists a polynomial p and a* BPP *predicate $V(x,q,r)$ defined for any $q,r\in\Sigma^{p(|x|)}$, such that*

$$\begin{aligned} x\in L_{\rm yes} &\quad\Longrightarrow\quad \exists\, r(x,q)\ \mathbb{P}[V(x,q,r(x,q))=1]\ge 2/3 \\ x\in L_{\rm no} &\quad\Longrightarrow\quad \forall\, r(x,q)\ \mathbb{P}[V(x,q,r(x,q))=1]\le 1/3 \end{aligned}\tag{1}$$

where $q\in\Sigma^{p(|x|)}$ is a uniformly distributed random bit string.

The main goal of this section is to show that LH-MIN for stoquastic Hamiltonians belongs to the class AM. Moreover, we will prove that evaluation of the largest eigenvalue of any n-qubit non-negative matrix whose matrix elements are efficiently computable is a problem that naturally sits in AM. This result applies even to matrices that lack any additional structure like locality or sparseness. To emphasize this point, we will formulate all results in terms of *black box matrices.* A black box matrix G of size $2^n\times 2^n$ is an oracle that takes as input two binary strings $x,y\in\Sigma^n$ and returns a matrix element $G_{x,y}=\langle x|G|y\rangle$ written in the binary form. We shall always assume that any matrix element $G_{x,y}$ has at most poly(n) binary digits (see the remark after Definition 1). In the case when G is specified by a local Hamiltonian, there is no need to query the oracle, since G has a concise representation and we can compute $G_{x,y}$ in a time poly(n).

Let G be a black box non-negative matrix and let $\mu(G)$ be the largest eigenvalue of G. To cast the evaluation of $\mu(G)$ into a decision problem we shall introduce two thresholds: an *upper threshold* μ_+ and a *lower threshold* μ_-, such that $0<\mu_-<\mu_+$ and the separation between μ_- and μ_+ is large enough.

Definition 4 *For any polynomial $p(n)$ define a set $\Lambda(p)$ involving all 4-tuples (n,G,μ_+,μ_-) such that n is an integer $1\le n<\infty$, $\mu_\pm$ are positive numbers such that $\log(\mu_+)-\log(\mu_-)\ge 1/p(n)$, and G is a $2^n\times 2^n$ real symmetric matrix such that*

- $0\le G_{x,y}\le 1$ *for all $x,y\in\Sigma^n$.*
- *Either $\mu(G)\ge\mu_+$ or $\mu(G)\le\mu_-$.*

Suppose we are given a 4-tuple $(n,G,\mu_+,\mu_-)\in\Lambda(p)$ and our goal is to decide whether $\mu(G)\ge\mu_+$. According to the Weyl perturbation theorem (see the remark after Definition 1), the correct decision can be made even if the matrix elements $G_{x,y}$ and the numbers $\mu_\pm$ are specified only up to some precision δ polynomial in 2^{-n}. Indeed, if G and G' are two $2^n\times 2^n$ matrices such that matrix elements of G and G' are ϵ-close, then $|\mu(G)-\mu(G')|\le||G-G'||\le 2^n\epsilon$. Thus, although $\Lambda(p)$ is a continuum set, we can safely assume that the numbers $\mu_\pm$ and any matrix element $G_{x,y}$ are described by poly(n) bits.

Definition 5 (Stoquastic Largest Eigenvalue Problem) *Given is a 4-tuple $(n,G,\mu_+,\mu_-)\in\Lambda(p)$ where G is specified by a black box. Decide whether $\mu(G)\ge\mu_+$.*

[4] It is known that completeness with a constant error probability is equivalent to perfect completeness, see [25].

Remark: One can easily see that stoquastic LH-MIN is a special case of the problem above. Indeed, if $H \in \Omega(k, p_1, p_2)$ is a k-local stoquastic Hamiltonian on n-qubits, see Definitions 1,2, one can define a non-negative matrix $G = (1/2)(I - H/p_1(n))$, so that $0 \le G_{x,y} \le 1$. One can also define the thresholds $\mu_+ = 1/2$ and $\mu_- = (1/2)(1 - 1/p_1(n)p_2(n))$. Clearly, the resulting 4-tuple $(n, G, \mu_+, \mu_-) \in \Lambda(p)$ for a proper choice of the polynomial p.

Theorem 6 *Stoquastic Largest Eigenvalue Problem belongs to the class* AM.

Proof: Consider any 4-tuple $(n, G, \mu_+, \mu_-) \in \Lambda(p_1)$ where p_1 is a fixed polynomial. Instead of proving the lower bound $\mu(G) \ge \mu_+$ Merlin will actually try to prove a lower bound $\mathrm{Tr}(G^L) \ge (\mu_+)^L$ where L is a large even integer. Note that

$$\mu(G) \ge \mu_+ \quad \Longrightarrow \quad \mathrm{Tr}(G^L) \ge \mu_+^L$$

$$\mu(G) \le \mu_- \quad \Longrightarrow \quad \mathrm{Tr}(G^L) \le 2^n \, \mu_-^L.$$

The separation between the value of the trace for positive and negative instances is thus given by

$$\frac{\mathrm{Tr}(G^L)_{\mathrm{yes}}}{\mathrm{Tr}(G^L)_{\mathrm{no}}} \ge 2^{\frac{L}{p_1(n)} - n}.$$

If one chooses $L = 2np_1(n)$, the separation is 2^n.

The next step is to represent the evaluation of the trace $\mathrm{Tr}(G^L)$ as a counting problem. As was mentioned above, we can assume that the matrix elements $G_{x,y}$ have at most $p_2(n)$ digits, where $p_2(n)$ is a polynomial. In order to define the counting problem, we shall represent G as an average over an ensemble of binary stochastic matrices $G(t)$, where t is a random uniformly distributed binary string $t \in \Sigma^{p_2(n)}$, that is

$$G = \frac{1}{2^m} \sum_{t \in \Sigma^m} G(t), \quad m \equiv p_2(n). \tag{2}$$

Any member of the ensemble $G(t)$ is a binary matrix, that is, matrix elements of $G(t)$ take only values 0 and 1. This representation is efficient in the sense that for any fixed strings x, y, t one can find a matrix element $\langle x|G(t)|y\rangle$ by making one query to the black box for G and performing a polynomial-time computation. Details of the representation Eq. (2) are not essential for the analysis of the proving protocol, so we postpone its proof until Lemma 1. Now we have

$$\mathrm{Tr}(G^L) = \frac{1}{2^{m\,L}} \sum_{t_1, \ldots, t_L} \mathrm{Tr}\left(G(t_1) \cdots G(t_L)\right) \equiv \frac{1}{2^{m\,L}} \sum_s F(s),$$

where $s = (t_1, \ldots, t_L, x_1, \ldots, x_L)$ is a binary string of length $(m+n)L$ and $F(s)$ is a Boolean function

$$F(s) = \langle x_1|G(t_1)|x_2\rangle \, \langle x_2|G(t_2)|x_3\rangle \cdots \langle x_L|G(t_L)|x_1\rangle \in \{0, 1\}.$$

Evaluation of $F(s)$ requires L black box queries and polynomial-time computation. Summarizing, the value of $\mathrm{Tr}(G^L)$ is proportional to a cardinality of a set $\Omega \subseteq \Sigma^{(m+n)L}$ supporting the function F,

$$\mathrm{Tr}\left(G^L\right) = \frac{1}{2^{mL}} \, |\Omega|, \quad \Omega = \{s \in \Sigma^{(m+n)L} \, : \, F(s) = 1\},$$

and membership $s \in \Omega$ can be efficiently verified. Note that there is large enough separation between the cardinality of Ω for positive and negative instances:

$$\mu(G) \ge \mu_+ \quad \Longrightarrow \quad |\Omega| \ge \mathrm{LARGE}$$
$$\mu(G) \le \mu_- \quad \Longrightarrow \quad |\Omega| < \mathrm{SMALL},$$

where

$$\mathrm{LARGE} = 2^{L(p_2(n) + \log \mu_+)} \quad \text{and} \quad \mathrm{SMALL} = 2^{L(p_2(n) + \log \mu_- + \frac{n}{L})}, \tag{3}$$

such that

$$\text{LARGE} = 2^n \cdot \text{SMALL} \quad \text{if} \quad L = 2np_1(n). \tag{4}$$

Thus it suffices for Merlin to prove a lower bound $|\Omega| \geq \text{LARGE}$.

We can now invoke the Goldwasser and Sipser approximate counting protocol [20] based on Carter-Wegman universal hashing functions [21]. Recall that Ω is a set of k-bit strings, where $k = L(n + p_2(n))$. The main idea of [20] is that Arthur can compress k-bit strings to shorter b-bit strings using randomly chosen linear hash functions. One can choose parameters of the hashing such that the image $h(\Omega) \subseteq \Sigma^b$ is sufficiently dense (for positive instances). Arthur evaluates a volume of $h(\Omega)$ using the standard Monte-Carlo method: he generates a large list of random b-bit strings and evaluates the fraction of strings that belong to $h(\Omega)$. At this stage he needs Merlin's help, since a membership in the set $h(\Omega)$ is no longer efficiently verifiable (each string in Σ^b may have exponentially large number of pre-images). In Appendix A we give some details of the parameters of the hash functions.

Now we prove the Lemma underlying Eq. (2)

Lemma 1 *Let $I_m = \{2^{-m}\, p\}_{p=0,\ldots,2^m-1}$ be the set of all real numbers between 0 and 1 having at most m binary digits. Let $g : \Sigma^n \to I_m$ be a function specified by a black box. Then there exists a Boolean function $f : \Sigma^n \times \Sigma^m \to \Sigma$ such that*

$$g(x) = \frac{1}{2^m} \sum_{t \in \Sigma^m} f(x,t) \quad \textit{for all} \quad x \in \Sigma^n.$$

Besides, $f(x,t)$ can be represented by a circuit of length poly$(n+m)$ making one query to the black box.

Proof Let $d_j(x)$ be the j-th binary digit of $g(x)$, that is

$$g(x) = \sum_{j=1}^m \frac{1}{2^j}\, d_j(x).$$

Define m auxiliary Boolean functions

$$\begin{aligned}
f_1(x,t) &= d_1(x) \wedge t_1, \\
f_2(x,t) &= d_2(x) \wedge (\neg\, t_1) \wedge t_2, \\
f_3(x,t) &= d_3(x) \wedge (\neg\, t_1) \wedge (\neg\, t_2) \wedge t_3, \\
&\ldots \\
f_m(x,t) &= d_m(x) \wedge (\neg\, t_1) \wedge \ldots \wedge (\neg\, t_{m-1}) \wedge t_m.
\end{aligned}$$

Here t_j is the j-th bit of t. Clearly,

$$\frac{1}{2^m} \sum_{t \in \Sigma^m} f_j(x,t) = \frac{1}{2^j}\, d_j(x), \quad j = 1, \ldots, m.$$

By definition, the functions f_j and f_k are mutually exclusive for $j \neq k$. Therefore

$$\sum_{j=1}^m f_j = f_1 \vee f_2 \vee \ldots \vee f_m.$$

Thus we can define the desired function $f(x,t)$ as $f = f_1 \vee f_2 \vee \ldots \vee f_m$. ■

Comment: The representation Eq. (2) corresponds to choosing $g(x) = \langle y|G|z\rangle$, where x is a concatenation of the strings y and z.

4 Stoquastic LH-MIN is MA-hard

In order to show that stoquastic LH-MIN is MA-hard we will view Arthur's BPP circuit as a quantum circuit. This quantum circuit will take as input: a quantum state $|\xi\rangle$ from Merlin, a set of $|+\rangle$ states (to simulate randomness) and some ancillas set to $|0\rangle$. The quantum circuit consists only of classical reversible gates and at the end Arthur measures a single qubit q_{out} in the z-basis. He obtains 1 with high probability if the answer to his decision problem is yes; otherwise he obtains 0 with high probability. If Merlin can only provide a classical state it is clear that the class of decision problems that can be solved this way is equal to MA. Before we argue that this new class of decision problems is equal to MA, let us give the proper definition.

Definition 7 ($\mathrm{MA_q}$) *A promise problem $L = L_{\text{yes}} \cup L_{\text{no}} \subseteq \Sigma^*$ belongs to the class $\mathrm{MA_q}$ iff there exists a polynomial p and a classical reversible circuit V_x that takes an input in $(\mathbb{C}^2)^{\otimes p(|x|)}$ and is followed by a single qubit measurement, such that*

$$
\begin{aligned}
x \in L_{\text{yes}} &\implies \exists\, |\xi\rangle\ \mathbb{P}\big[V_x(|00\ldots0\rangle, |+\rangle^{\otimes r}, |\xi\rangle) = 1\big] \geq 2/3 \\
x \in L_{\text{no}} &\implies \forall\, |\xi\rangle\ \mathbb{P}\big[V_x(|00\ldots0\rangle, |+\rangle^{\otimes r}, |\xi\rangle) = 1\big] \leq 1/3.
\end{aligned}
\tag{5}
$$

Lemma 2 $\mathrm{MA} = \mathrm{MA}_q$.

Proof $\mathrm{MA}_q \subseteq \mathrm{MA}$: Let $L \in \mathrm{MA}_q$. If $x \in L$ we have $\mathbb{P}(V_x(|+\rangle^{\otimes r}, |00\ldots0\rangle, |\xi\rangle) = 1) \geq 2/3$. Let $\Pi_1 = |1\rangle\langle 1|_{q_{\text{out}}}$. We can write the success probability as

$$
\mathbb{P}(1) = \langle \xi | M | \xi \rangle \geq 2/3, \tag{6}
$$

where $M = \langle 00\ldots0, +^{\otimes r}| V_x^T \Pi_1 V_x |00\ldots0, +^{\otimes r}\rangle$. We note that the observable M is diagonal in the standard basis, i.e. $M = \frac{1}{2^r}\sum_z a_z |z\rangle\langle z|$ where a_z is a non-negative integer. This implies that $\lambda_{\max}(M) = \max_\xi \langle \xi|M|\xi\rangle$ is achieved for some bit string $|\xi\rangle = z_{\max}$. Thus there exists a bit-string for which $\mathbb{P}(1) \geq 2/3$ and this bit-string will be the input for the MA-verifier. If $x \notin L$, we have that $\forall \xi\ \mathbb{P}(1) = \langle \xi|M|\xi\rangle \leq 1/3$, thus this also holds for the subset of all classical inputs from Merlin.
$\mathrm{MA} \subseteq \mathrm{MA}_q$: let a decision problem be in MA. If $x \in L$, the classical witness can be used as input to the $\mathrm{MA_q}$-verifier and gives $\mathbb{P}(1) \geq 2/3$. If $x \notin L$, we need to argue that Merlin cannot cheat by giving Arthur a quantum state. Since the problem is in MA, we have that $\forall z\ \mathbb{P}(1) = \langle z|M|z\rangle \leq 1/3$. Since M is diagonal in the z-basis, this implies that $\lambda_{\max}(M) \leq 1/3$ and thus there is no quantum state with expectation value higher than $1/3$ with respect to M. ■

Since Arthur's verifying circuit in MA_q is a quantum circuit, one can apply Kitaev's circuit-to-Hamiltonian construction to MA_q and prove that the ground-state energy problem for a 6-local stoquastic Hamiltonian is MA_q=MA-hard.

Lemma 3 *6-local stoquastic* LH-MIN *is* MA*-hard.*

Proof Let V_x be Arthur's verifying circuit that has an input of r qubits in the state $|+\rangle$ (labeled as coin-qubits), k ancilla qubits in the state $|00\ldots0\rangle$ (labeled as anc-qubits) and a quantum state $|\xi\rangle$ with s qubits. Let V_x have a total of T reversible classical gates, denoted as $R_T \ldots R_2 R_1$. W.l.o.g. we can assume that each gate is a Toffoli gate, since these gates are universal for classical reversible computation. We follow the Hamiltonian construction in [6] (see also [26]). Let $H^{(5)} = H_{\text{in}} + H_{\text{out}} + H_{\text{prop}} + H_{\text{clock}}$ be a Hamiltonian acting on T clock-qubits labeled by $t = 1\ldots T$ and $n = r + k + s$ computational qubits. We have

$$
\begin{aligned}
H_{\text{in}} &= \sum_{i=1}^{r} |-\rangle\langle -|_{\text{coin},i} \otimes |0\rangle\langle 0|_{t=1} + \sum_{j=1}^{k} |1\rangle\langle 1|_{\text{anc},j} \otimes |0\rangle\langle 0|_{t=1}, \\
H_{\text{out}} &= |0\rangle\langle 0|_{q_{\text{out}}} \otimes |1\rangle\langle 1|_{t=T}, \\
H_{\text{clock}} &= \sum_{t=1}^{T} |01\rangle\langle 01|_{t-1,t}.
\end{aligned}
\tag{7}
$$

Furthermore, $H_{\rm prop} = \sum_{t=1}^T H_{\rm prop}(t)$ with

$$\begin{aligned}
H_{\rm evolv}(1) &= |00\rangle\langle 00|_{1,2} + |10\rangle\langle 10|_{1,2} - R_1 \otimes (|10\rangle\langle 00|_{1,2} + |00\rangle\langle 10|_{1,2}), \\
H_{\rm evolv}(t) &= |100\rangle\langle 100|_{t-1,t,t+1} + |110\rangle\langle 110|_{t-1,t,t+1} \\
&\quad - R_t \otimes (|110\rangle\langle 100|_{t-1,t,t+1} + |100\rangle\langle 110|_{t-1,t,t+1}), \quad 1 < t < T \\
H_{\rm evolv}(T) &= |10\rangle\langle 10|_{T-1,T} + |11\rangle\langle 11|_{T-1,T} - R_T \otimes (|11\rangle\langle 10|_{T-1,T} + |10\rangle\langle 11|_{T-1,T}).
\end{aligned} \tag{8}$$

It was proved in [6] that if there exists a $|\xi\rangle$ such that V_x outputs 1 with probability larger than or equal to $1-\epsilon$ then $\lambda(H^{(5)}) \leq \epsilon$. If on the other hand for all $|\xi\rangle$ V_x outputs 1 with probability smaller or equal to ϵ, then $\lambda(H^{(5)}) \geq \frac{c(1-\epsilon)}{T^3}$ for some constant c. Thus the ground-state energy problem of this Hamiltonian is MA_q-hard. We need only to verify that this Hamiltonian $H^{(5)}$ is of the stoquastic-type. The only terms that are off-diagonal in the computational basis can be found in $H_{\rm prop}$ and $H_{\rm in}$. Inspection of these terms confirms that the Hamiltonian is stoquastic. ■

Remarks: One may wonder whether one can extend the class MA_q to a class in which Arthur's verification circuit is more quantum, while the corresponding Hamiltonian is still stoquastic. One possibility is to allow for a measurement in the x-basis (instead of the z-basis) at the end, see [22].

5 Perturbation Theory Gadgets for Stoquastic Hamiltonians

The goal of this section is to understand whether the complexity of stoquastic k-local LH-MIN depends upon k — the number of qubits involved in the interactions. We will answer this question for Hamiltonians that are termwise-stoquastic, i.e., those having a decomposition $H = \sum_S H_S$, where S runs over subsets of k qubits and H_S is a stoquastic Hamiltonian acting on the subset S. Direct inspection shows that all examples of stoquastic Hamiltonians encountered in the paper are also termwise-stoquastic and for 2-local Hamiltonians these notions coincide.

Theorem 8 *Let k be any constant. Any instance of k-local termwise stoquastic LH-MIN can be reduced in polynomial time to 2-local stoquastic LH-MIN.*

Throughout this section we will use the word stoquastic to refer to Hamiltonians that are termwise-stoquastic. Our main technical tool is the perturbation theory gadgets developed in [7] and extended in [8]. The proof can be organized in three parts. Firstly we reduce k-local interactions to 3-local interactions using a variant of the subdivision gadget from [8]. This gadget only requires perturbation theory to second-order. The second step is to bring a stoquastic 3-local Hamiltonian into a special form

$$H = H_{\rm else} - \sum_{(j,k,l)} h_{jkl}\, X_j\, X_k\, X_l, \tag{9}$$

where $H_{\rm else}$ is a 2-local stoquastic Hamiltonian, (j,k,l) labels triples of qubits, and h_{jkl} are non-negative constants. We shall refer to Hamiltonians having a decomposition as in Eq. (9) as *triple-X 3-local Hamiltonians*. In order to implement the second step a new three-qubit gadget will be constructed. The final step is to reduce 3-qubit interactions $-h_{jkl}\, X_j\, X_k\, X_l$ to 2-local interactions. This can be done using the three-qubit gadget of [7]. Throughout this section we follow the notation of [7] and [8].

5.1 Reduction to 3-local interactions: the subdivision gadget

Using the standard operator algebra basis of n qubits, any stoquastic k-local Hamiltonian $H_{\rm target}$ can be written as

$$H_{\rm target} = \Omega\, I - \sum_{(j_1,\ldots,j_k)} \sum_{a_1,\ldots,a_k} h^{\alpha_1,\ldots,\alpha_k}_{j_1,\ldots,j_k}\, E^{\alpha_1}_{j_1} E^{\alpha_2}_{j_2} \cdots E^{\alpha_k}_{j_k} + \text{h.c.}$$

Here $(j_1,\ldots,j_k)$ labels subsets of k qubits, α labels one-qubit matrices $E^0 = |0\rangle\langle 0|$, $E^1 = |0\rangle\langle 1|$, $E^2 = |1\rangle\langle 0|$, $E^3 = |1\rangle\langle 1|$, and $h^{\alpha_1,\ldots,\alpha_k}_{j_1,\ldots,j_k}$ are non-negative constants. The energy shift $\Omega\, I$ is introduced in order to make all diagonal matrix elements of $H_{\rm target}$ non-positive. Let us partition each subset $(j_1,\ldots,j_k)$ into two non-overlapping subsets of nearly equal size. Then we can rewrite $H_{\rm target}$ as

$$H_{\rm target} = \Omega\, I - \sum_{a=1}^M (C_a\otimes D_a + C_a^\dagger \otimes D_a^\dagger), \quad M = 4^k \binom{n}{k}$$

where C_a and D_a are operators having the following properties:
(1) All C_a and D_a have non-negative matrix elements,
(2) C_a and D_a act on non-overlapping subsets of at most $\lceil k/2\rceil$ qubits,
(3) $C_a^\dagger C_a$ and $D_a D_a^\dagger$ are diagonal.
Since we regard k as a constant, the number of terms in the sum is polynomial, $M=\mathrm{poly}(n)$.

Let us introduce M mediator qubits and consider a Hamiltonian $\tilde{H}$ acting on n data qubits and M mediator qubits:

$$\tilde{H} = H+V, \quad H=\Delta \sum_{a=1}^M I_{\rm data}\otimes |1\rangle\langle 1|_a, \quad V=-\sqrt{\Delta}\sum_{a=1}^M (C_a+D_a^\dagger)\otimes \sigma_a^+ + (C_a^\dagger + D_a)\otimes \sigma_a^- + Q\otimes I_M,$$

where $Q=\sum_{a=1}^M (C_a^\dagger C_a + D_a D_a^\dagger)$, $\sigma^+=|1\rangle\langle 0|$, $\sigma^-=|0\rangle\langle 1|$. As for Δ, it must be chosen such that $||V||\ll \Delta$. Note that all terms in H and V are stoquastic. Denote the Hilbert space of n data qubits as $\mathcal{H}_{\rm data}$. Then H has zero-energy levels defining the eigen-subspace $\mathcal{L}_- = \mathcal{H}_{\rm data}\otimes |0^{\otimes M}\rangle$ separated from the rest of the spectrum by a gap Δ. Considering V as a perturbation, we compute the self-energy operator

$$\Sigma_- = V_{--} + V_{-+}G_+V_{+-} + V_{-+}G_+V_{++}G_+V_{+-} + V_{-+}G_+V_{++}G_+V_{++}G_+V_{+-} + \cdots \tag{10}$$

up to second-order of the perturbation theory one gets[5]

$$\Sigma_-(z) = -\sum_{a=1}^M (C_a\otimes D_a + C_a^\dagger\otimes D_a^\dagger) + O(\Delta^{-1/2}) \quad \text{for any} \quad z=O(1).$$

Accordingly, the ground-state energy of $\tilde{H}$ approximates the ground-state energy of $H_{\rm target}-\Omega I$ with precision $\delta = O(\Delta^{-1/2})$. This reduces k-local stoquastic LH-MIN to $\lceil k/2\rceil + 1$-local stoquastic LH-MIN. By repeating this reduction $O(\log(k))$ times[6] we end up with a 3-local stoquastic Hamiltonian.

For obvious reasons the subdivision gadget cannot transform 3-local terms into 2-local terms. However, we can use it to reduce the variety of 3-local terms which have to be dealt with using different (and more complicated) methods.

If one considers all possible 3-local terms proportional to $-E_1^{\alpha_1}E_2^{\alpha_2}E_2^{\alpha_2}$, there are essentially four different types of terms (up to permutations of qubits and bit flips $0\leftrightarrow 1$) shown in the first column of Table 1. By choosing the operators C_a and D_a from the second and the third column, one can reduce interactions of type (a) to type (b), type (b) to type (c), and finally type (c) to type (d). This requires at most three repetitions of the subdivision gadget. Now we can assume that a Hamiltonian has the form

$$H_{\rm target} = H_{\rm else} - \sum_{(j,k,l)}\sum_{\alpha,\beta,\gamma=\pm} h_{j,k,l}^{\alpha,\beta,\gamma}\, \sigma_j^\alpha\otimes\sigma_k^\beta\otimes\sigma_l^\gamma, \tag{11}$$

where $H_{\rm else}$ is a stoquastic 2-local Hamiltonian, (j,k,l) labels triples of qubits, and $h_{j,k,l}^{\alpha,\beta,\gamma}\ge 0$.

5.2 Reduction from 3-local to special 3-local stoquastic Hamiltonians

Our next goal is to construct a gadget reducing the stoquastic Hamiltonian of Eq. (11) to a special 3-local Hamiltonian, see Eq. (9). To simplify the discussion, let us first consider a Hamiltonian Eq. (11) with a

[5] To avoid a proliferation of poly(n) bounds, we treat all terms proportional to $||C_a||$, $||D_a||$, or M as $O(1)$. In general all these terms can be bounded by poly(n)). Since we are free to choose Δ polynomially large, the bounds $O(\Delta^{-1/2})$ and $O(\mathrm{poly}(n)\Delta^{-1/2})$ are equally good.

[6] After each iteration we have to introduce an energy shift $\Omega\, I$, since the terms $C_a^\dagger C_a + D_a D_a^\dagger$ may produce positive matrix elements on the diagonal.

	3-local term	choice of C_a	choice of D_a
(a)	$-\|000\rangle\langle 000\|_{jkl}$	$\|00\rangle\langle 00\|_{jk}$	$\|0\rangle\langle 0\|_l$
(b)	$-\|000\rangle\langle 100\|_{jkl}$	$\|00\rangle\langle 10\|_{jk}$	$\|0\rangle\langle 0\|_l$
(c)	$-\|000\rangle\langle 110\|_{jkl}$	$\|00\rangle\langle 11\|_{jk}$	$\|0\rangle\langle 0\|_l$
(d)	$-\|000\rangle\langle 111\|$		

Table 1: Successive application of the subdivision gadget with the choice of C_a and D_a as above reduces any 3-local term to a term of type (d).

single 3-local term:

$$H_{\rm target} = H_{\rm else} - 3(B_1 \otimes B_2 \otimes B_3 + B_1^\dagger \otimes B_2^\dagger \otimes B_3^\dagger), \quad B_j = \sigma^+ \quad \text{or} \quad B_j = \sigma^-.$$

where $H_{\rm else}$ is a 2-local stoquastic Hamiltonian, and the factor 3 is introduced for convenience. We shall need three mediator qubits which will be labeled by $1,2,3$. Consider a Hamiltonian $\tilde{H}$ acting on n data qubits and three mediator qubits:

$$\begin{aligned} \tilde{H} &= H + V, \quad H = I_{\rm data} \otimes H_M \\ H_M &= -\frac{1}{2}\Delta_x\,(X_1 \otimes X_2 \otimes X_3 - I) - \frac{1}{4}\Delta_z\,(Z_1 \otimes Z_2 + Z_2 \otimes Z_3 + Z_1 \otimes Z_3 - 3\,I), \\ V &= -\omega \sum_{j=1}^{3} B_j \otimes \sigma_j^+ + B_j^\dagger \otimes \sigma_j^- + H_{\rm else} \otimes I_M. \end{aligned} \tag{12}$$

The parameters $\omega, \Delta_x, \Delta_z$ must be chosen as

$$\omega = \delta^{-4}, \quad \Delta_x = \delta^{-5}, \quad \Delta_z = \delta^{-6}, \quad 0 < \delta \ll 1. \tag{13}$$

It will be shown later that δ is the precision up to which the ground-state energy of $\tilde{H}$ approximates the ground-state energy of $H_{\rm target}$ (as before, we assume for simplicity that $||B_j||$ and $||H_{\rm else}||$ are of order $O(1)$). Note that Eq. (13) implies $\omega \ll \Delta_x \ll \Delta_z$. Also note that all local terms in H and V are stoquastic. The only 3-local term in $\tilde{H}$ is the one proportional to $-X_1X_2X_3$, so that $\tilde{H}$ is a special 3-local stoquastic Hamiltonian.

The Hamiltonian H_M is diagonal in the basis of states

$$|\Psi^\pm\rangle = \frac{1}{\sqrt{2}}\,|000\rangle \pm \frac{1}{\sqrt{2}}\,|111\rangle, \quad \text{and} \quad |\phi_j^\pm\rangle = X_j\,|\Psi^\pm\rangle, \quad j = 1,2,3. \tag{14}$$

The spectrum of H_M is illustrated in Figure 1. By construction, H_M has a unique ground-state $|\Psi^+\rangle$ having zero energy[7], while the first excited state $|\Psi^-\rangle$ has energy Δ_x. The top part of the spectrum involves six nearly-degenerate (as long as $\Delta_x \ll \Delta_z$) states $\phi_j^\pm$. Since $||V|| = O(\omega) \ll \Delta_x$, we can treat V as a perturbation and compute the self-energy operator on the zero-energy subspace of H, that is $\mathcal{L}_- = \mathcal{H}_{\rm data} \otimes |\Psi^+\rangle$.

We can use the expansion of Eq. (10). The perturbation V is designed such that $V_{--} = \langle \Psi^+|V|\Psi^+\rangle = H_{\rm else}$, see Eq. (12). The contribution of the second-order term is proportional to the identity operator (see Appendix C for details of the calculation):

$$V_{-+}G_+V_{+-} = -(3/4)\omega^2\left[\Delta_z^{-1} + (\Delta_z + \Delta_x)^{-1}\right]\, I \equiv \Omega\, I.$$

We can regard it as a shift of energy. Therefore

$$\Sigma_-(z) = \Omega\, I + H_{\rm else} + V_{-+}G_+V_{++}G_+V_{+-} + [\text{higher order terms}]. \tag{15}$$

The key feature of the gadget is that the perturbation V cannot cause a direct transition from the ground-state Ψ^+ to the first excited state Ψ^- (or vice versa). Any direct transition maps Ψ^+ into the high-energy

[7] One should not confuse labels $\pm$ of the states $\Psi^\pm$ and $\phi^\pm$ with the labels $\pm$ referring to the low-energy and high-energy subspaces that appear in the perturbative series Eq. (10).

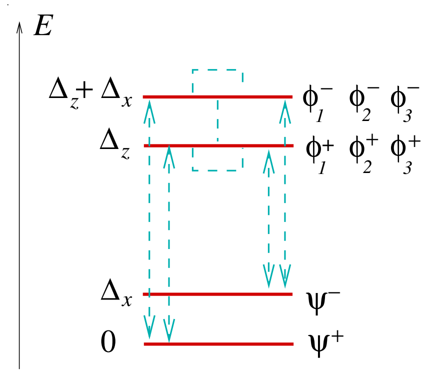

Figure 1: Allowed transitions induced by perturbation V acting on the eigenstates of H are indicated by dashed lines. Direct transitions between Ψ^+ and Ψ^- levels are impossible.

band $\phi^\pm$ spanned by six states ϕ_j^+ and ϕ_j^- having energy of order $\Delta_z \gg \Delta_x$, see Figure 1. Thus any third-order process follows the following scheme:

$$\Psi^+ \to \phi^\pm \to \phi^\pm \to \Psi^+.$$

Since the energy splitting Δ_x of the $\phi^\pm$ band is much smaller than its absolute energy Δ_z, one can use an approximation in which the two intermediate Green's functions G_+ in Eq. (15) are proportional to the identity operator, $G_+(z) = (zI - H_+)^{-1} \approx -I/\Delta_z$ for any $z = O(1)$. Within this approximation one has

$$\Sigma_-(z) - \Omega\, I \approx H_{\rm else} + \frac{1}{\Delta_z^2} V_{-+}V_{++}V_{+-} = H_{\rm else} + \frac{1}{\Delta_z^2}\langle \Psi^+|V^3|\Psi^+\rangle \approx H_{\rm else} - \frac{3\omega^3}{\Delta_z^2}(B_1\otimes B_2 \otimes B_3 + B_1^\dagger \otimes B_2^\dagger \otimes B_3^\dagger),$$

which approximates $H_{\rm target}$ since $\omega^3 = \Delta_z^2$. An accurate calculation of $\Sigma_-(z)$, performed in Appendix C, shows that the error in the approximation is of order $O(\delta)$. Contributions from transitions involving the Ψ^- level appear only in the fourth-order term in Eq. (10) according to the following scenario:

$$\Psi^+ \to \phi^\pm \to \Psi^- \to \phi^\pm \to \Psi^+, \tag{16}$$

see Figure 1. In Appendix C we show that the fourth-order term is of order $O(\delta)$. Therefore $\Sigma_-(z) = \Omega\, I + H_{\rm target} + O(\delta)$ for any $z = O(1)$, and thus the ground-state energy of $H_{\rm target}$ is δ-close to the ground-state energy of $\tilde{H} - \Omega\, I$.

One can applying this gadget in parallel to each term in the Hamiltonian Eq. (11) and obtain the desired reduction to a special 3-local stoquastic Hamiltonian.

5.3 Reduction from special 3-local to 2-local Hamiltonians

To simplify the discussion let us consider a special 3-local stoquastic Hamiltonian with a single 3-qubit interaction:

$$H_{\rm target} = H_{\rm else} - 6 B_1 \otimes B_2 \otimes B_3,$$

where B_j are non-negative operators proportional to X_j and $H_{\rm else}$ is a 2-local stoquastic Hamiltonian. The 3-qubit interaction can be treated using the original three-qubit gadget in [7]. This original gadget coincides with the gadget defined in Eq. (12) if one chooses $\Delta_x = 0$. In this case the zero-energy subspace of H is $\mathcal{L}_- = \mathcal{H}_{\rm data} \otimes \mathcal{L}_-$, where $\mathcal{L}_-$ is spanned by the mediator qubit states $|000\rangle$ and $|111\rangle$. Note that $\tilde{H}$ is now a 2-local stoquastic Hamiltonian.

We can choose $\Delta_z = \delta^{-3}$ and $\omega = \delta^{-2}$. The analysis performed in [7] implies that the ground-state energy of $\tilde{H} = H + V$, see Eq. (12), is δ-close to the ground-state energy of an effective Hamiltonian

$$H_{\rm eff} = \Omega\, I + H_{\rm else} \otimes I_m - 6 B_1 B_2 B_3 \otimes X_m,$$

where I_m and X_m act on the two dimensional subspace of the mediator qubits spanned by $|000\rangle$ and $|111\rangle$ (regarded as logical $|0\rangle$ and $|1\rangle$ states). The energy shift is $\Omega\, I = -\delta^{-1}(B_1^2 + B_2^2 + B_3^2)$. Since $H_{\rm eff}$ is a stoquastic Hamiltonian, the Perron-Frobenius theorem implies that its ground-state $|\Psi_0\rangle$ can be chosen as a non-negative vector. Then a state

$$|\Psi_0'\rangle = |\Psi_0\rangle + (I \otimes X_m)\,|\Psi_0\rangle$$

is also a non-negative ground-state of H_{eff}. In addition, we have $(I\otimes X_m)|\Psi_0'\rangle = |\Psi_0'\rangle$. Therefore $H_{\rm eff} - \Omega\, I$ has the same ground-state energy as $H_{\rm target}$. This proves that the ground-state energies of $\tilde{H} - \Omega\, I$ and $H_{\rm target}$ are δ-close. To deal with multiple 3-qubit terms in Eq. (9) one applies this three-qubit gadget in parallel to every individual 3-qubit term.

Remark: In the original three-qubit gadget the operators B_j are required to be positive semi-definite in order to guarantee that the ground-state of $H_{\rm eff}$ belongs to the sector where X_m has eigenvalue $+1$.

6 Stoquastic LH-MIN and Classical Post-Selected Computation

The main goal of this section is to examine the complexity of stoquastic LH-MIN in the special case when the Hamiltonian possesses a polynomial spectral gap (i.e., the spectral gap scales as $1/p(n)$, where n is the number of qubits and p is a fixed polynomial). We shall prove that this problem can be placed in the complexity class PostBPP — a class of languages recognizable by a probabilistic polynomial time classical circuits with a post-selective readout of the answer. Speaking informally, any problem in the class PostBPP can be solved by a classical probabilistic circuit that outputs two random bits: a (the answer bit) and b (the success flag). The answer bit a contains the correct answer of the problem provided that $b=1$ (if $b=0$ the value of a may be arbitrary). The success probability $\mathbb{P}[b=1]$ must be positive for all input strings (however it may be exponentially small). Here is a more formal definition:

Definition 9 (PostBPP) *A promise problem $L = L_{\rm yes} \cup L_{\rm no}$ belongs to the class* PostBPP *iff there exist a polynomial p, predicates $a(x,y)$ and $b(x,y)$ from the class* P *defined for any $y \in \Sigma^{p(|x|)}$, such that*

$$\begin{aligned} x \in L &\implies \mathbb{P}[b(x,y)=1] > 0, \\ x \in L_{\rm yes} &\implies \mathbb{P}[a(x,y)=1\,|\,b(x,y)=1] \ge 2/3, \\ x \in L_{\rm no} &\implies \mathbb{P}[a(x,y)=1\,|\,b(x,y)=1] \le 1/3. \end{aligned}$$

where $y \in \Sigma^{p(|x|)}$ is a random uniformly distributed bit string, and $\mathbb{P}[a\,|\,b]$ is the conditional probability.

The quantum version of this class, PostBQP, was defined in Ref. [17] and in that paper it was shown that PP=PostBQP. The following lemma provides a characterization of PostBPP in terms of the standard complexity classes.

Lemma 4 $\mathrm{MA} \subseteq \mathrm{NP}^{\mathrm{BPP}} \subseteq \mathrm{PostBPP} = \mathrm{BPP}_{\mathrm{path}} \subseteq \mathrm{BPP}^{\mathrm{NP}} \subseteq \Sigma_3^p$.

Here $\mathrm{BPP}_{\mathrm{path}}$ is a class of problems solvable in polynomial time with a bounded error probability by a non-deterministic Turing machine that chooses its computational path randomly from the uniform distribution on a set of all possible paths, see [27]. The class $\mathrm{BPP}_{\mathrm{path}}$ is more powerful than BPP, since it offers the possibility to amplify the total probability of successful computational paths by adding 'idle' computational branches to a non-deterministic algorithm. In Appendix B we give a proof of the equality $\mathrm{PostBPP}=\mathrm{BPP}_{\mathrm{path}}$. All other statements made in the previous lemma follow directly from [27].

Theorem 10 *k-local stoquastic* LH $-$ MIN *with the promise that the spectral gap $\Delta = 1/\mathrm{poly}(n)$ belongs to* PostBPP.

Proof Let $H = \sum_S H_S \in \Omega(k, p_1, p_2)$ be k-local stoquastic Hamiltonian on n qubits, see Definitions 1 and 2. The first step is to transform H into a doubly-substochastic[8] matrix G. This is achieved by choosing

$$G = \frac{1}{2}\,(I - H/q(n)), \quad q(n) = 2\max(1, 2^k \binom{n}{k} p_1(n)).$$

[8] By definition, a non-negative matrix is doubly-substochastic iff it is symmetric and the sum of the elements in every row is smaller or equal to 1, see [28]

The coefficient $q(n)$ is chosen such that all eigenvalues of G are between 0 and 1, while the matrix elements $G_{x,y} = \langle x|G|y\rangle$ obey the inequalities

$$G_{x,y} \geq 0 \quad \text{and} \quad \frac{1}{4} \leq \sum_{z\in\Sigma^n} G_{x,z} \leq 1 \quad \text{for all} \quad x, y \in \Sigma^n. \tag{17}$$

Obviously, $q(n)$ is a fixed polynomial. Let $\mu(G)$ be the largest eigenvalue of G. The correct decision for LH-MIN with the Hamiltonian H can be made if we can evaluate $\mu(G)$ with polynomial precision:

$$\begin{aligned} \lambda(H) \leq 0 \quad &\Rightarrow \quad \mu(G) \geq \mu_+ = \frac{1}{2}, \\ \lambda(H) \geq 1/p_2(n) \quad &\Rightarrow \quad \mu(G) \leq \mu_- = \frac{1}{2}\left(1 - \frac{1}{q(n)p_2(n)}\right). \end{aligned}$$

We shall present a polynomial-time probabilistic algorithm that evaluates $\mu(G)$ with a precision $1/\mathrm{poly}(n)$ using a post-selective readout of the answer.

Define a matrix B which is diagonal in the standard basis such that

$$B_x \equiv \langle x|B|x\rangle = \sum_{y\in\Sigma^n} G_{x,y}. \tag{18}$$

We can transform G into a doubly-stochastic matrix F as follows:

$$F = G \otimes I + (I - B) \otimes X = \begin{pmatrix} G & I - B \\ I - B & G \end{pmatrix}.$$

F acts on $n + 1$ qubits, and, thus, specifies a random walk on a space Σ^{n+1}. The fact that H is a k-local Hamiltonian implies that F is a sparse matrix — it has at most $\binom{n}{k}2^k + 1$ non-zero elements in each column (row). Moreover, for any fixed column (row) positions of the non-zero matrix elements and their values can be computed in $\mathrm{poly}(n)$ time. This means that the random walk defined by F can be efficiently simulated on a BPP machine, provided that the number of steps is at most $\mathrm{poly}(n)$.

Our algorithm requires the simulation of w independent random walks $(X_t^{(i)})_{t=0,\ldots,L,i=1,\ldots,w}$ whose transition probabilities are given by F. Here $0 \leq t \leq L$ is the (discrete) time parameter, i is the index of the random walk and L, w will be specified later. Let us start each random walk $X_t^{(i)}$ from a point $X_0^{(i)} = (x_0^{(i)}, 0) \in \Sigma^{n+1}$, such that the ancillary bit (the last one) is set to 0, and the n bits constituting the original system are initialized by a random string $x_0^{(i)} \in \Sigma^n$ drawn from the uniform distribution with independent choices of $x_0^{(i)}$ for different i. Suppose that after t steps the ith random walk arrives at a point $X_t^{(i)} = (x_t^{(i)}, b_t^{(i)})$ ($0 \leq t \leq L$, $1 \leq i \leq w$). Let us postselect only those samples where the ancillary bits remain in the state 0 for the whole duration of each of the w walks. In terms of the formal definition of PostBPP we have to define the success flag bit as $b = \neg(\vee_{i=1}^w \vee_{t=0}^L b_t^{(i)})$. The probability for the ancillary bit to stay in 0 is

$$\mathbb{P}[b = 1] = \left(\frac{1}{2^n}\sum_{x_0,x_L\in\Sigma^n} \langle x_0|G^L|x_L\rangle\right)^w \geq \frac{1}{4^{wL}} > 0,$$

where we have used the inequality Eq. (17).

Conditioned on $b = 1$, the random variables $(x_L^{(i)})_{i=1}^w$ are independent samples from the probability distribution $P_L(\cdot)$ given by

$$P_L(y) = \frac{\sum_{x\in\Sigma^n}\langle x|G^L|y\rangle}{\sum_{x,y\in\Sigma^n}\langle x|G^L|y\rangle}, \; y \in \Sigma^n.$$

Therefore, the quantity

$$\mu_{\mathrm{est}}(G) \equiv \frac{\sum_{i=1}^w B_{x_L^{(i)}}}{w} = \frac{\sum_{i=1}^w \left(\sum_{x\in\Sigma^n}\langle x|G|x_L^{(i)}\rangle\right)}{w} \tag{19}$$

has an average value equal to

$$\mathbb{E}(\mu_{\text{est}}(G)) = \frac{\sum_{x,y\in\Sigma^n}\langle x|G^{L+1}|y\rangle}{\sum_{x,y\in\Sigma^n}\langle x|G^L|y\rangle}.$$

Since $1/4 \le B_x \le 1$ for all $x \in \Sigma^n$, Azuma's inequality implies that

$$\forall \delta > 0,\ \ \mathbb{P}(|\,\mu_{\text{est}}(G) - \mathbb{E}(\mu_{\text{est}}(G))| > \delta \mid b = 1) \le 2e^{-\frac{\delta^2 w}{2}}. \tag{20}$$

We now *claim* that for L chosen sufficiently large,

$$\mathbb{E}(\mu_{\text{est}}(G)) = \frac{\sum_{x,y\in\Sigma^n}\langle x|G^{L+1}|y\rangle}{\sum_{x,y\in\Sigma^n}\langle x|G^L|y\rangle}$$

is close to the largest eigenvalue $\mu(G)$ of G. More precisely, we will show that

Lemma 5 *Let $\mu_0 \ge \mu_1$ be the largest eigenvalue and second largest eigenvalue of G. Suppose that* $\log(\mu(G)) - \log(\mu_1) \ge \frac{1}{r(n)}$. *If one chooses $L = \frac{5nr(n)}{2}$ then*

$$|\mu_0 - \mathbb{E}(\mu_{\text{est}}(G))| = O(2^{-n}).$$

This is the only step where the spectral gap assumption is used. Let us postpone the proof of the lemma until the end of the section. We choose $L = \frac{5nr(n)}{2}$ (clearly, the spectral gaps of H and G are related by a polynomial factor, so that $r(n)$ is a fixed polynomial). Then by Eq. (20)

$$\forall \delta > 0,\ \ \mathbb{P}\big(|\mu_{\text{est}}(G) - \mu(G)| > \delta + O(2^{-n})\,\big|\, b = 1\big) \le 2e^{-\frac{\delta^2 w}{2}}.$$

For some constant $c > 0$, taking $w = 2n^{2c}\ln(6)$ ensures that $|\mu_{\text{est}}(G) - \mu(G)| = O(n^{-c})$ with probability at least $2/3$. Since the coefficients B_y can be computed efficiently for any bit-string y and vary within a constant range we can evaluate $\mu_{\text{est}}(G)$ (as in Eq. (19)) with a precision $1/\text{poly}(n)$ using $w = \text{poly}(n)$ random walks of length $L = \text{poly(n)}$. The complexity of simulating the random walks is polynomial in L, w, and n; it follows that we can solve our decision problem in PostBPP. ■

Proof of Lemma 5:
Define an operator

$$\hat{\Delta}_\ell = \frac{1}{\mu_0^\ell}\left(G^\ell - \mu_0^\ell\,|\Psi_0\rangle\langle\Psi_0|\right).$$

After simple algebra one gets

$$\mathbb{E}(\mu_{\text{est}}(G)) = \mu_0\left(\frac{1 + \epsilon^2\sum_{x,y\in\Sigma^n}\langle x|\hat{\Delta}_{L+1}|y\rangle}{1 + \epsilon^2\sum_{x,y\in\Sigma^n}\langle x|\hat{\Delta}_L|y\rangle}\right), \quad \epsilon \equiv \frac{1}{\sum_x\langle x|\Psi_0\rangle}.$$

Let $\mu_0 \ge \mu_1 \ge \ldots, \mu_{2^n-1}$ be the eigenvalues of G. Note that G is chosen such that $\mu_j \ge 0$. Therefore

$$||\hat{\Delta}_L|| = \left(\frac{\mu_1}{\mu_0}\right)^L \le 2^{-\frac{L}{r(n)}}.$$

Let us choose $L = \frac{5nr(n)}{2}$. Then $||\hat{\Delta}_L|| \le 2^{-5n/2}$ and therefore

$$\left|\sum_{x,z}\langle x|\hat{\Delta}_L|z\rangle\right| \le 2^n\left|\left(\frac{\sum_x\langle x|}{2^{n/2}}\right)\hat{\Delta}_L\left(\frac{\sum_z|z\rangle}{2^{n/2}}\right)\right| \le 2^n||\Delta_L|| \le 2^{-3n/2}.$$

Clearly, the same inequalities hold with $L+1$ replacing L. On the other hand, $\epsilon \le 1$ since

$$\sum_{x\in\Sigma^n}\langle x|\Psi_0\rangle \ge \sqrt{\sum_{x\in\Sigma^n}(\langle x|\Psi_0\rangle)^2} = 1.$$

It follows that $|\mu_0 - \mathbb{E}(\mu_{\rm est}(G))| \le O(\mu_0 2^{-3n/2}) = O(2^{-n})$ as $\mu_0 \le \max_x B_x \le 1$ under our assumptions.■

Our result has a simple implication for adiabatic quantum computation using stoquastic Hamiltonians. It is known that the power of efficient adiabatic quantum computation with general 2-local Hamiltonians is equal to that of polynomial-time quantum circuits [29]. All Hamiltonians on the adiabatic path are required to have a polynomial gap in order for the adiabatic theorem to apply. Now let us restrict ourselves to stoquastic Hamiltonians with a polynomial gap. By the MA-hardness construction and analogous to the arguments in [29], one can argue that any polynomial-time probabilistic computation can be simulated by an efficient adiabatic path using stoquastic Hamiltonians only. It is a more interesting but open question whether every efficient adiabatic path using stoquastic Hamiltonians can be simulated by a polynomial-time probabilistic machine. The proof of Theorem 10 shows that post-selected classical computation allows one to efficiently sample from the ground-state distribution of a stoquastic Hamiltonian. Note that this may be potentially stronger than merely estimating the lowest-lying eigenvalue. In the proof we use the ability to sample from the ground-state to estimate the lowest-lying eigenvalue. A adiabatic path with stoquastic Hamiltonians, each of which has a $1/\mathrm{poly}(n)$ gap, can thus be simulated by post-selected classical computation and the decision problem that can be solved by these means is contained in PostBPP.

7 Acknowledgements

We acknowledge support by the NSA and the ARDA through ARO contract number W911NF-04-C-0098.

A The Approximate Counting Problem and Hash Functions

For the sake of completeness we explain how to choose the parameters of the hash functions in the proof of Theorem 6, see the original paper [20] for more details. Define

$$b = \lceil \log \mathrm{LARGE} \rceil + 3.$$

Without loss of generality $b \le k$ (otherwise Arthur has to verify that Ω contains a finite fraction of k-bit strings, which can be done by the standard Monte-Carlo method without compression). Let $h_1, \ldots, h_k$ be $k \times b$ binary matrices chosen uniformly at random. Each matrix h_j defines a linear hash function $h_j : \Sigma^k \to \Sigma^b$. Denote

$$h(\Omega) = \bigcup_{j=1}^{k} h_j(\Omega) \subseteq \Sigma^b.$$

We need the following technical lemma from [20] (a proof is given at the end of this appendix).

Lemma 6 *For any set $\Omega \subseteq \Sigma^k$ and for any $b \le k$ such that $|\Omega| \le 2^{b-2}$ one has*

$$\mathbb{P}\left[|h(\Omega)| \ge \frac{|\Omega|}{k} \right] \ge 1 - \frac{1}{2^k}.$$

Neglecting the exponentially small error probability 2^{-k} one gets

$$\begin{aligned} |\Omega| \ge \mathrm{LARGE} \quad &\Longrightarrow \quad |h(\Omega)| \ge \frac{\mathrm{LARGE}}{k} \ge \left(\frac{1}{8k}\right) 2^b, \\ |\Omega| \le \mathrm{SMALL} \quad &\Longrightarrow \quad |h(\Omega)| \le k \cdot \mathrm{SMALL} \le k 2^{-n}\, \mathrm{LARGE} \le \left(\frac{k}{2^{n+2}}\right) 2^b \end{aligned}$$

For the second line we have used the trivial bound $|h(\Omega)| \le k|\Omega|$ and Eq. (4). If n is sufficiently large, $h(\Omega)$ contains a polynomially large fraction of b-bit strings for positive instances and an exponentially small fraction for negative instances. Arthur can distinguish the two case by the Monte-Carlo method using Merlin's advice to verify membership in $h(\Omega)$. This completes the proof of Theorem 6. ■

Proof of Lemma 6:
Let us say that a function h_j is *invertible* at the point $x \in \Omega$ if $h_j(x) \neq h_j(y)$ for all $y \in \Omega \backslash \{x\}$. Define a set

$$\Omega_j = \{x \in \Omega \, : \, h_j \text{ is invertible at } x\}.$$

Clearly,

$$|h(\Omega)| \geq |h_j(\Omega)| \geq |\Omega_j| \quad \text{for any} \quad j = 1, \ldots, k.$$

Thus

$$\mathbb{P}\left[|h(\Omega)| \geq \frac{|\Omega|}{k}\right] \geq \mathbb{P}\left[\bigcup_{j=1}^{k} \Omega_j = \Omega\right]. \tag{21}$$

Since the probability of collisions for h_j is 2^{-b}, we have

$$\mathbb{P}[h_j \text{ is not invertible at } x] \leq \frac{|\Omega|}{2^b}.$$

Therefore

$$\mathbb{P}\left[\bigcup_{j=1}^{k} \Omega_j \neq \Omega\right] = \mathbb{P}[\exists\, x \in \Omega \, : \, \forall j \; h_j \text{ is not invertible at } x] \leq |\Omega| \left(\frac{|\Omega|}{2^b}\right)^k. \tag{22}$$

Combining Eqs. (21) and (22) and taking into account the conditions on b, k, and $|\Omega|$ finishes the proof. ■

B PostBPP = $\mathrm{BPP}_{\mathrm{path}}$

The class $\mathrm{BPP}_{\mathrm{path}}$ is defined most conveniently in terms of non-deterministic Turing machines. Let M be a non-deterministic Turing machine (TM). We shall assume that at each step M chooses one of two computational paths. Given an input string $x \in \Sigma^*$, a polynomial-time non-deterministic TM makes at most $q(|x|)$ steps before it stops, where q is a fixed polynomial. Whenever M stops, it outputs an answer bit $a = 1$ (accept), or $a = 0$ (reject).

Let $\mathrm{path}(M, x)$ and $\mathrm{acc}(M, x) \subseteq \mathrm{path}(M, x)$ be a set of all computational paths and a set of accepting paths for a machine M running on input string x. By definition, $|\mathrm{path}(M, x)| \leq 2^{q(|x|)}$. One can visualize $\mathrm{path}(M, x)$ as a subtree of a binary branching tree of a height $q(|x|)$. Some paths make it all the way from the root to a leaf of the tree and some paths end before making $q(|x|)$ steps. Let us introduce a branching variable $y \in \Sigma^{q(|x|)}$, such that a bit y_j specifies what path M chooses at step j (if a computational path ends before making $q(|x|)$ steps, the remaining bits of y can be ignored). For any $x \in \Sigma^*$ and $y \in \Sigma^{q(|x|)}$ let $l(x, y)$ be the number of steps that M does on input x before it stops and $a(x, y)$ be the value of the answer bit. By definition, $1 \leq l(x, y) \leq q(|x|)$ for any x, y and

$$|\mathrm{path}(M, x)| = \frac{1}{2^{q(|x|)}} \sum_{y \in \Sigma^{q(|x|)}} 2^{l(x,y)}, \quad |\mathrm{acc}(M, x)| = \frac{1}{2^{q(|x|)}} \sum_{y \, : \, a(x,y)=1} 2^{l(x,y)}. \tag{23}$$

Now we can define the class $\mathrm{BPP}_{\mathrm{path}}$ more formally.

Definition 11 *A promise problem* $L = L_{\mathrm{yes}} \cup L_{\mathrm{no}}$ *belongs to the class* $\mathrm{BPP}_{\mathrm{path}}$ *iff there exist a non-deterministic polynomial-time Turing machine* M *such that*

$$x \in L_{\mathrm{yes}} \quad \Longrightarrow \quad |\mathrm{acc}(M, x)| \geq \frac{2}{3} \, |\mathrm{path}(M, x)|$$

$$x \in L_{\mathrm{no}} \quad \Longrightarrow \quad |\mathrm{acc}(M, x)| \leq \frac{1}{3} \, |\mathrm{path}(M, x)|$$

Let us first prove $\mathrm{BPP}_{\mathrm{path}} \subseteq \mathrm{PostBPP}$. Indeed, consider a non-deterministic polynomial-time Turing machine M as above. Let C be a classical circuit (more strictly, a uniform family of circuits) that takes as input a pair (x, y) with $y \in \Sigma^{q(|x|)}$, and simulates M for $q(|x|)$ steps according to the computational path y.

The circuit C outputs the answer bit $a(x,y)$ and the number of steps $l(x,y)$ in the path y. The idea is that we can simulate M by choosing y randomly from the uniform distribution and use post-selection to balance the resulting distribution on $\mathrm{path}(M,x)$. Indeed, define a random success flag bit b, such that we have a probability distribution of b conditioned on x and y

$$\mathbb{P}[b=1\,|\,x,y] = \frac{1}{2^{q(|x|)-l(x,y)}}.$$

Since the circuit C outputs $l(x,y)$, one can easily generate a bit with the desired distribution using a polynomial number of ancillary random bits. Making use of the formulas in Eq. (23) one can easily get

$$\mathbb{P}[a=1\,|\,b=1] = \frac{\mathbb{P}[a=1,b=1]}{\mathbb{P}[b=1]} = \frac{2^{-2\,q(|x|)}\sum_{y\,:\,a(x,y)=1} 2^{l(x,y)}}{2^{-2\,q(|x|)}\sum_y 2^{l(x,y)}} = \frac{|\mathrm{acc}(M,x)|}{|\mathrm{path}(M,x)|}.$$

Comparing it with Def. 9, we conclude that a language recognized by M belongs to PostBPP.

Now let us prove PostBPP $\subseteq$ $\mathrm{BPP_{path}}$. Indeed, let $L = L_{\mathrm{yes}} \cup L_{\mathrm{no}}$ be a language from PostBPP. One can use the standard majority voting procedure to reduce the error probability from 1/3 to 1/4, i.e., we can assume that the predicates $a(x,y)$ and $b(x,y)$ from Def. 9 satisfy

$$\begin{aligned}
x \in L \quad &\Longrightarrow \quad \mathbb{P}[b(x,y)=1] > 0,\\
x \in L_{\mathrm{yes}} \quad &\Longrightarrow \quad \mathbb{P}[a(x,y)=1\,|\,b(x,y)=1] \ge 3/4,\\
x \in L_{\mathrm{no}} \quad &\Longrightarrow \quad \mathbb{P}[a(x,y)=1\,|\,b(x,y)=1] \le 1/4.
\end{aligned}$$

Here $y \in \Sigma^{p(|x|)}$ is a uniformly random bitstring and p is a polynomial. The inequality $\mathbb{P}[b=1]>0$ implies that there exists at least one $y \in \Sigma^{p(|x|)}$ such that $b(x,y)=1$. Therefore we can bound the probability of successful computation from below as

$$\mathbb{P}[b=1] \ge \frac{1}{2^{p(|x|)}}.$$

Construct a non-deterministic Turing machine M that takes x as input and does the following:
(1) Perform $p(|x|)$ branchings to initialize a string $y \in \Sigma^{p(|x|)}$,
(2) Compute predicates $a = a(x,y)$ and $b = b(x,y)$,
(3) If $b=0$, output a,
(4) If $b=1$, perform $p(|x|)+4$ idle branchings and output a.
Let us verify that M recognizes the language L in the sense of Def. 11. Indeed, one can easily check that

$$|\mathrm{path}(M,x)| = 2^{p(|x|)}\left[\mathbb{P}[b=0] + 2^{p(|x|)+4}\,\mathbb{P}[b=1]\right]$$

and

$$|\mathrm{acc}(M,x)| = 2^{p(|x|)}\left[\mathbb{P}[a=1,b=0] + 2^{p(|x|)+4}\,\mathbb{P}[a=1,b=1]\right].$$

Consider first the case $x \in L_{\mathrm{yes}}$. Then

$$\frac{|\mathrm{acc}(M,x)|}{|\mathrm{path}(M,x)|} \ge \frac{\mathbb{P}[a=1,b=1]}{2^{-p(|x|)-4}+\mathbb{P}[b=1]} \ge \frac{\mathbb{P}[a=1,b=1]}{\mathbb{P}[b=1](1+2^{-4})} \ge \frac{3}{4(1+2^{-4})} > \frac{2}{3}.$$

Here we have used the fact that $\mathbb{P}[b=1] \ge 2^{-p(|x|)}$. Consider now the case $x \in L_{\mathrm{no}}$. Then

$$\frac{|\mathrm{acc}(M,x)|}{|\mathrm{path}(M,x)|} \le \frac{2^{-p(|x|)-4}+\mathbb{P}[a=1,b=1]}{\mathbb{P}[b=1]} \le \frac{1}{4} + 2^{-4} < \frac{1}{3}.$$

Thus M indeed recognizes L. ■

C The three-qubit gadget

In this appendix we will explicitly calculate the self-energy operator $\Sigma_-(z)$ for the perturbed Hamiltonian in Eq. (12) up to third order in the perturbative series of Eq. (10). We shall also evaluate the norm of the fourth-order term. It follows directly from Eq. (12) that

$$V_{--} = \langle \Psi^+ | V | \Psi^+ \rangle = H_{\rm else}.$$

A straightforward calculation yields

$$V_{+-} = -\frac{\omega}{2} \sum_{j=1}^{3} \sum_{\alpha=\pm 1} (B_j + \alpha B_j^\dagger) \otimes |\phi_j^\alpha\rangle,$$

where $|\phi_j^\alpha\rangle = \sigma_j^x |\Psi^\alpha\rangle$, see Eq. (14). Now we can compute the second-order term for the self-energy operator:

$$\Sigma_-^{(2)}(z) = V_{-+} G_+ V_{+-} = \left(-\frac{\omega}{2}\right)^2 \sum_{j=1}^{3} \sum_{\alpha=\pm 1} \frac{(B_j^\dagger + \alpha B_j)(B_j + \alpha B_j^\dagger)}{z - \Delta_\alpha}, \tag{24}$$

Here we denote $\Delta_+ = \Delta_z$ and $\Delta_- = \Delta_z + \Delta_x$. Substituting $z = O(1)$ and taking into account that $B_j^2 = 0$, $B_j B_j^\dagger + B_j^\dagger B_j = I$, we come to

$$\Sigma_-^{(2)}(z) = \Omega\, I + O(\delta^4), \quad \Omega = -(3/4)\omega^2 \left[\Delta_z^{-1} + (\Delta_z + \Delta_x)^{-1}\right]. \tag{25}$$

To compute the third-order term we need to know V_{++}. It is enough to find the matrix elements of V between the $\phi^\pm$ states (since transitions between Ψ^- and $\phi^\pm$ do not appear in the third order). A straightforward calculation yields

$$\langle \phi_j^\alpha | V | \phi_l^\beta \rangle = H_{\rm else} \delta_{j,l} \delta_{\alpha,\beta} - \frac{\omega}{2} \sum_{k=1}^{3} \epsilon(j,k,l) \left[\alpha B_k + \beta B_k^\dagger\right], \quad \text{where} \quad \epsilon(j,k,l) = \begin{cases} 1 & \text{if} \quad j \neq k \neq l \\ 0 & \text{otherwise} \end{cases} \tag{26}$$

Therefore

$$\Sigma_-^{(3)}(z) = V_{-+} G_+ V_{++} G_+ V_{+-} = \left(-\frac{\omega}{2}\right)^3 \sum_{j,k,l=1}^{3} \sum_{\alpha,\beta=\pm 1} \frac{(B_j^\dagger + \alpha B_j)(\alpha B_k + \beta B_k^\dagger)(B_l + \beta B_l^\dagger)\epsilon(j,k,l)}{(z - \Delta_\alpha)(z - \Delta_\beta)} + O(\delta^4).$$

Taking into account Eq. (13) one easily gets (for any $z = O(1)$)

$$\Sigma_-^{(3)}(z) = -3(B_1 \otimes B_2 \otimes B_3 + B_1^\dagger \otimes B_2^\dagger \otimes B_3^\dagger) + O(\delta). \tag{27}$$

Although we have not calculated the fourth-order correction

$$\Sigma_-^{(4)} = V_{-+} G_+ V_{++} G_+ V_{++} G_+ V_{+-}$$

exactly, we have to get an upper bound on its norm. The fourth-order processes may involve the low-lying level Ψ^-, see Eq. (16), and potentially these processes can give a non-negligible contribution to Σ_- as $\Sigma_-^{(3)}$. Keeping in mind Eq. (16) one can easily get (for any $z = O(1)$)

$$||\Sigma_-^{(4)}(z)|| = ||V_{-+} G_+ V_{++} G_+ V_{++} G_+ V_{+-}|| = O\left(\frac{\omega^4}{\Delta_z \Delta_x \Delta_z}\right) = O(\delta^{-16+12+5}) = O(\delta).$$

As for the higher-order corrections to Σ^- (from the fifth- order onwards) their contribution contains an additional factor ω/Δ_z, or ω/Δ_x which is at most δ. Therefore we arrive at

$$\Sigma_-(z) = \Omega\, I + H_{\rm else} - 3(B_1 \otimes B_2 \otimes B_3 + B_1^\dagger \otimes B_2^\dagger \otimes B_3^\dagger) + O(\delta)$$

for any $z = O(1)$. Here Ω is the energy shift given by Eq. (25).

References

[1] S.R. White. Density matrix formulation for quantum renormalization groups. *Phys. Rev. Lett.*, 69:2863, 1992.

[2] S. Rommer and S. Ostlund. A class of Ansatz wave functions for 1D spin systems and their relation to DMRG. 55:2164, 1997.

[3] F. Verstraete, D. Porras, and J.I. Cirac. DMRG and periodic boundary conditions: a quantum information perspective. *Phys. Rev. Lett.*, 93:227205, 2004.

[4] N. Trivedi and D. Ceperley. Ground-state correlations of quantum antiferromagnets: A Green's function Monte-Carlo study. 41:4552, 1990.

[5] M. Buonaura and S. Sorella. Numerical study of the two-dimensional Heisenberg model using a Green's function Monte-Carlo technique with a fixed number of walkers. 57:11446, 1998.

[6] A. Yu. Kitaev, A.H. Shen, and M.N. Vyalyi. *Classical and Quantum Computation. Vol. 47 of Graduate Studies in Mathematics.* American Mathematical Society, Providence, RI, 2002.

[7] J. Kempe, A. Kitaev, and O. Regev. The Complexity of the Local Hamiltonian Problem. *SIAM Journal of Computing*, 35(5):1070–1097, 2006. Earlier version in Proc. of 24th FSTTCS.

[8] R.I. Oliveira and B.M. Terhal. The complexity of quantum spin systems on a two-dimensional square lattice. 2005, `http://arxiv.org/abs/quant-ph/0504050`.

[9] F. Verstraete, M. M. Wolf, D. Perez-Garcia, and J. I. Cirac. Criticality, the area law, and the computational power of PEPS. `http://arxiv.org/abs/quant-ph/0601075`.

[10] J. Watrous. Succinct quantum proofs for properties of finite groups. In *Proceedings of 41st FOCS*, pages 537–546, 2000, `http://arxiv.org/abs/cs.CC/0009002`.

[11] E. Farhi, J. Goldstone, S. Gutmann, J. Lapan, A. Lundgren, and D. Preda. A quantum adiabatic evolution algorithm applied to random instances of an NP-complete problem. *Science*, 292:472—-476, 2001.

[12] D. M. Ceperley. Path integrals in the theory of condensed Helium. *Rev. Mod. Phys.* 67, 279-356, 1995.

[13] G. Burkard, R. H. Koch, and D. P. DiVincenzo. Multi-level quantum description of decoherence in superconducting flux qubits. *Phys. Rev. B* 69, 064503, 2004. `http://arxiv.org/abs/cond-mat/0308025`.

[14] D. F. Walls and G. J. Milburn. *Quantum Optics.* Springer, New York, 1995.

[15] See, e.g., A. J. Leggett, S. Chakravarty, A. T. Dorsey, M. P. A. Fisher, A. Garg, and W. Zwerger. Dynamics of the dissipative two-state system. *Rev. Mod. Phys.* 59:2–86, 1987.

[16] D. Aharonov and A. Ta-Shma. Adiabatic quantum state generation and statistical zero knowledge. `http://arxiv.org/abs/quant-ph/quant-ph/0301023`.

[17] S. Aaronson. Quantum Computing, Postselection, and Probabilistic Polynomial-Time. `http://arxiv.org/abs/quant-ph/quant-ph/0412187`.

[18] M. Suzuki. Relationship between d-dimensional quantum spin systems and (d+1)-dimensional Ising systems. *Prog. Theor. Phys.*, 56(5):1454–1469, 1976.

[19] F. Barahona. On the Computational Complexity of Ising Spin Glass Models. *Jour. of Phys. A: Math. and Gen.*, 15:3241–3253, 1982.

[20] S. Goldwasser and M. Sipser. Private coins versus public coins in interactive proof systems. In *Proceedings of 18th STOC*, pages 59–68, 1986.

[21] J. L. Carter and M. N. Wegman. Universal classes of hash functions. *Journal of Computer and Systems Sciences*, 18(2):143–154, 1979.

[22] A. Bessen, S. Bravyi and B.M. Terhal. In preparation.

[23] R. A. Horn and C. R. Johnson. *Matrix Analysis.* Cambridge University Press, 1985.

[24] L. Babai. Trading group theory for randomness. In *Proceedings of 17th STOC*, pages 421–429, 1985.

[25] M. Furer, O. Goldreich, Y. Mansour, M. Sipser, and S. Zachos. On completeness and soundness in Interactive Proof Systems. *Advances in Computing Research*, 5:429–442, 1989.

[26] D. Aharonov and T. Naveh. Quantum NP—A Survey. 2002, `http://arxiv.org/abs/quant-ph/0210077`.

[27] Y. Han, L. Hemaspaandra, and T. Thierauf. Threshold computation and cryptographic security. *SIAM J. Comput.*, 26(1):59–78, 1997.

[28] R. Bhatia, editor. *Matrix analysis.* Springer, New York, 1997.

[29] D. Aharonov, W. van Dam, Z. Landau, S. Lloyd, J. Kempe, and O. Regev. Universality of Adiabatic Quantum Computation. In *Proceedings of 45th FOCS*, 2004, `http://arxiv.org/abs/quant-ph/0405098`.

[30] L.L. Ng. Heisenberg Model, Bethe Ansatz and Random Walks. Harvard University, 1996.