Quantum Information and Computation, Vol. 15, No. 13&14 (2015) 1122–1140

MONTE CARLO SIMULATION OF STOQUASTIC HAMILTONIANS

SERGEY BRAVYI
IBM T.J. Watson Research Center, 1101 Kitchawan Road
Yorktown Heights, NY 10598, USA

Received January 9, 2015
Revised May 18, 2015

Stoquastic Hamiltonians are characterized by the property that their off-diagonal matrix elements in the standard product basis are real and non-positive. Many interesting quantum models fall into this class including the Transverse field Ising Model (TIM), the Heisenberg model on bipartite graphs, and the bosonic Hubbard model. Here we consider the problem of estimating the ground state energy of a local stoquastic Hamiltonian H with a promise that the ground state of H has a non-negligible correlation with some 'guiding' state that admits a concise classical description. A formalized version of this problem called Guided Stoquastic Hamiltonian is shown to be complete for the complexity class MA (a probabilistic analogue of NP). To prove this result we employ the Projection Monte Carlo algorithm with a variable number of walkers. Secondly, we show that the ground state and thermal equilibrium properties of the ferromagnetic TIM can be simulated in polynomial time on a classical probabilistic computer. This result is based on the approximation algorithm for the classical ferromagnetic Ising model due to Jerrum and Sinclair (1993).

Keywords: Quantum Monte Carlo, fermionic sign problem, transverse Ising model
Communicated by: B Kane & J Eisert

1 Introduction

Calculating the ground state energy and thermal equilibrium properties of interacting quantum many-body systems is one of the central problems in quantum chemistry and condensed matter physics. It was realized early on that the computational complexity of this problem is strongly affected by statistics of the constituent particles. For systems composed of bosons and for certain special classes of spin Hamiltonians the quantum partition function can be mapped to the one of a classical system occupying one extra spatial dimension [1] which often enables efficient Monte Carlo simulation [2, 3, 4, 5, 6, 7]. On the other hand, for systems composed of fermions and for the vast majority of spin Hamiltonians the quantum-to-classical mapping produces a partition function with unphysical Boltzmann weights taking both positive and negative (or complex) values — a phenomenon known as the "sign problem". In certain special cases, such as the Heisenberg antiferromagnetic model on bipartite graphs, the sign problem can be avoided by a suitable basis change [5]. In general however the sign problem appears to be an intrinsic feature of the quantum mechanics that limits applicability of Monte Carlo simulation algorithms [8].

The present paper attempts to provide a rigorous basis for the common belief that Hamiltonians avoiding the sign problem are "easy" to simulate. We focus on systems composed of

qubits (spins-1/2) and study the class of so-called *stoquastic*[a] Hamiltonians [9]. The defining property of stoquastic Hamiltonians is that their off-diagonal matrix elements in the computational basis must be real and non-positive. More formally, let n be the number of qubits and $k = O(1)$ be a small constant. Define a k-local stoquastic Hamiltonian as

$$H = \sum_{\alpha=1}^{M} H_\alpha, \tag{1}$$

where each term H_α is a hermitian operator on n qubits that acts non-trivially on a subset of at most k qubits and satisfies

$$\langle x|H_\alpha|y\rangle \le 0 \quad \text{for all } x, y \in \{0,1\}^n \text{ with } x \ne y. \tag{2}$$

From Eq. (2) one can easily infer that the Boltzmann exponential operator $e^{-\beta H_\alpha}$ has non-negative matrix elements for any inverse temperature $\beta \ge 0$. Furthermore, the matrix element $\langle x|e^{-\beta H_\alpha}|y\rangle$ depends non-trivially only on $O(1)$ bits of x and y. Combining these two properties one can approximate the quantum partition function $\mathrm{Tr}\, e^{-\beta H}$ by a classical partition function that involves local non-negative Boltzmann weights, for example, using the Suzuki-Trotter formula. In that sense any stoquastic Hamiltonian avoids the sign problem.

Below we assume that the Hamiltonian is normalized[b] such that $\|H_\alpha\| \le poly(n)$ for all α and $M \le poly(n)$. The quantity we are interested in is the ground state energy

$$\lambda = \min_{\psi} \langle\psi|H|\psi\rangle, \quad \langle\psi|\psi\rangle = 1. \tag{3}$$

We shall consider a problem Stoquastic Local Hamiltonian (Stoq-LH) where the goal is to estimate the ground state energy with a small additive error. More formally, an instance of Stoq-LH includes the number of qubits n, a list of interactions H_α as above, and real numbers $\lambda_{yes} < \lambda_{no}$ such that $\lambda_{no} - \lambda_{yes} \ge poly(1/n)$. The problem is to decide whether $\lambda \le \lambda_{yes}$ (yes-instance) or $\lambda \ge \lambda_{no}$ (no-instance) given a promise that λ does not belong to the interval $(\lambda_{yes}, \lambda_{no})$.

Since any Hamiltonian diagonal in the standard basis is stoquastic, Stoq-LH encompasses hard classical optimization problems such as k-SAT or MAX-CUT. This shows that Stoq-LH is at least NP-hard and essentially rules out a possibility that Stoq-LH admits an efficient algorithm. A natural next question is whether Stoq-LH is contained in the class NP or its probabilistic analogue called MA (Merlin-Arthur games) [10]. Loosely speaking, proving the containment in NP or MA would imply that certain ground state properties of stoquastic Hamiltonians can be efficiently *verified* even though they cannot be efficiently computed. Some progress along these lines has been made in Refs. [9, 11] by proving that Stoq-LH is contained in MA for the special case of frustration-free Hamiltonians (for a detailed discussion of the previous work see Section 3). The present paper extends these results by identifying two new classes of stoquastic Hamiltonians whose ground state properties can be efficiently verified and, in certain cases, efficiently computed.

[a] The term stoquastic introduced in Ref. [9] conveys the fact that the studied problems lie on the border between quantum mechanics and classical theory of stochastic matrices. For non-technical purposes 'stoquastic' is equivalent to "avoiding the sign problem".

[b] Here $poly(n)$ and $poly(1/n)$ denote functions of n with the asymptotic scaling $n^{O(1)}$ and $n^{-O(1)}$.

2 Summary of results

The first class of Hamiltonians that we study is motivated by the Quantum Phase Estimation (QPE) algorithm [12] and the question of whether application of QPE to stoquastic Hamiltonians can be efficiently simulated classically. Recall that the ground state energy of a local Hamiltonian H can be efficiently estimated on a quantum computer via QPE only if one is able to prepare some initial state ϕ that has a non-negligible (at least $poly(1/n)$) overlap with the exact ground state of H. Following Ref. [13] we shall refer to such initial state ϕ as a *guiding state* since its purpose is to guide the algorithm towards the ground state of H. It is usually assumed that a good choice of the guiding state can be made if the system under consideration is sufficiently well understood and some physical theory describing, at least approximately, its ground state properties is available. This motivates the study of the Stoquastic Local Hamiltonian problem with an extra promise that the Hamiltonian admits a guiding state. A natural question is whether in this case QPE can be replaced by some classical algorithm. We show that the answer is YES provided that the guiding state has efficiently computable amplitudes and a non-negligible pointwise correlation with the exact ground state as formally defined below.

Definition 1 *Let H be a stoquastic Hamiltonian. We will say that H admits a guiding state iff there exists a pair of normalized n-qubit states ψ, ϕ with non-negative amplitudes in the standard basis such that ψ is a ground state of H, the function $x \to \langle x|\phi\rangle$ is computable by a classical circuit of size $poly(n)$, and*

$$\langle x|\phi\rangle \geq \frac{\langle x|\psi\rangle}{poly(n)} \quad \textit{for all } x \in \{0,1\}^n \tag{4}$$

A state ϕ satisfying the above conditions will be referred to as a guiding state. Note that any stoquastic Hamiltonian has a ground state with non-negative amplitudes due to the Perron-Frobenius theorem. However, it is rather unlikely that any stoquastic Hamiltonian admits a guiding state. It should be emphasized that our classical algorithm and QPE need guiding states with different properties. The pointwise correlation condition in Eq. (4) is much stronger that the non-negligible overlap condition needed for QPE. Note that Eq. (4) implies $\langle\psi|\phi\rangle \geq poly(1/n)$, but the converse is not true. On the other hand, given a short classical circuit that computes amplitudes of ϕ, generally one cannot convert it to a short quantum circuit that prepares ϕ.

Define a problem Guided Stoquastic Local Hamiltonian (Guided Stoq-LH) as a special case of Stoq-LH where any yes-instance must satisfy two promises: *Promise 1:* $\lambda \leq \lambda_{yes}$ and *Promise 2:* H admits a guiding state. In the case of no-instances the only promise is that $\lambda \geq \lambda_{no}$. Note that the guiding state is not regarded as a part of the input. Our main result is the following.

Theorem 1 *Guided Stoq-LH is contained in the class promise-*MA *for any constant k. Guided Stoq-LH is complete for promise-*MA *for $k \geq 6$.*

Less formally, Theorem 1 asserts that a verifier (usually called Arthur) with polynomial classical computational resources and a random number generator can reliably distinguish between yes- and no-instances of the problem by consulting an untrusted prover (usually called Merlin) which has unlimited computational power. Merlin's goal is to convince Arthur that a given instance of the problem is positive, that is, $\lambda \leq \lambda_{yes}$. To this end Merlin sends

Arthur a witness — a classical bit string which, if Merlin is honest, includes the description of a guiding state and certain additional information. In the case of yes-instances Merlin can always find a witness convincing Arthur that $\lambda \leq \lambda_{yes}$ with probability close to one. Meanwhile, for no-instances Arthur decides that $\lambda \geq \lambda_{no}$ with probability close to one for any Merlin's witness.

To prove the containment in MA we employ a version of the Projection Monte Carlo algorithm with a variable number of walkers. This algorithm has been previously proposed in the context of quantum Monte Carlo simulations by Cerf and Martin [13] and by Oliveira [14]. The key idea of the algorithm is to convert a stoquastic Hamiltonian into a random walk using Poisson-distributed random variables. A state of the walk is a function that assigns a non-negative integer to each n-bit binary string. Such function can be visualized as a population of walkers distributed over the Boolean cube. A typical step of the walk involves moving one or several walkers to a new location, creating new walkers, and eliminating some existing walkers. We show that for any yes-instance Merlin can choose a witness such the the total population size is confined to the interval $[1, poly(n)]$ during all steps of the walk with a non-negligible probability. Meanwhile, for any no-instance and for any Merlin's witness the average population size decreases exponentially with time. In this case the population either becomes extinct after $poly(n)$ steps or becomes too large at some intermediate step due to statistical fluctuations. By implementing the random walk and monitoring the population size Arthur can therefore distinguish between yes- and no-instances. Our rigorous analysis of the algorithm based on the second moment method appears to be new. The last statement of Theorem 1 (MA-completeness) follows trivially from the results of Ref. [9]. We discuss some open problem and potential improvements of Theorem 1 in Section 4.5.

The second class of Hamiltonians that we study is the Transverse field Ising Model (TIM):

$$H = -\sum_{1\leq u<v\leq n} J_{u,v} Z_u Z_v - \sum_{1\leq u\leq n} h_u X_u. \tag{5}$$

Here X_u and Z_u are the Pauli operators acting on a qubit u, while $J_{u,v}$ and h_u are real coefficients. A direct inspection shows that H is stoquastic iff $h_u \geq 0$ for all u. Any TIM Hamiltonian can be made stoquastic by a transformation $H \to Z_u H Z_u$ that flips the sign of h_u without changing any other terms. Define a partition function

$$\mathcal{Z} = \mathrm{Tr}\, e^{-H}. \tag{6}$$

Our second result shows that $\mathcal{Z}$ can be efficiently approximated with a small multiplicative error δ in the special case of the *ferromagnetic* TIM, that is, when $J_{u,v} \geq 0$ for all u, v. More precisely, let

$$J = \max\{J_{u,v}, |h_u|\} \tag{7}$$

be the maximum norm of the interactions and $0 < \delta < 1$ be the desired precision. We shall say that $\mathcal{Z}$ admits a *fully polynomial randomized approximation scheme* (FPRAS) if there exists a classical randomized algorithm with the running time $poly(n, J, \delta^{-1})$ that takes as input a pair (H, δ) and outputs a random variable $\tilde{\mathcal{Z}}$ such that

$$\Pr\Big[\,(1-\delta)\mathcal{Z} \leq \tilde{\mathcal{Z}} \leq (1+\delta)\mathcal{Z}\,\Big] \geq 2/3. \tag{8}$$

Theorem 2 *The partition function of the ferromagnetic TIM admits FPRAS.*
This immediately implies that the free energy $F(T) = -T \log \left(\operatorname{Tr} e^{-H/T} \right)$ can be approximated with an additive error δ in time $poly(n, J, \delta^{-1}, T^{-1})$. Furthermore, since $F(0) - F(T) = \int_0^T dT' S(T') \leq nT$, where $S(T')$ is the entropy of the Gibbs state, we conclude that the ground state energy $\lambda = F(0)$ can be approximated with an additive error δ in time $poly(n, J, \delta^{-1})$. Theorem 2 is a simple application of the seminal result by Jerrum and Sinclair [15] who showed that thermal equilibrium properties of the ferromagnetic *classical* Ising model ($h_u = 0$ for all u) can be simulated efficiently.

Theorem 3 (Jerrum and Sinclair [15]) *The partition function of the ferromagnetic classical Ising model admits FPRAS.*
More precisely, the algorithm proposed in Ref. [15] has running time $poly(n, \delta^{-1})$ which is independent on J, as long as the cost of arithmetic operations with $J_{u,v}$ can be neglected[c]. To go from Theorem 3 to Theorem 2 we employ the standard quantum-to-classical mapping based on the Suzuki-Trotter formula. The only new ingredient that we add is a proof that the Suzuki-Trotter approximation leads to a small multiplicative error (as opposed to the additive error usually studied in the literature). We emphasize that although the proposed FPRAS for TIM is efficient in the complexity theory sense, it can hardly be used in practice. A rigorous upper bound on the running time of the FPRAS obtained in the proof of Theorem 2 is $O(n^{59} J^{21} \delta^{-9})$. Clearly, this leaves a lot of room for improvements.

3 Discussion and previous work

The class of stoquastic Hamiltonians encompasses many interesting quantum spin models originated both from the condensed matter physics and the quantum computing field. Well-known examples include TIM, the Heisenberg ferromagnetic and antiferromagnetic models (the latter can be made stoquastic on any bipartite graph), quantum annealing Hamiltonians [16, 17], the toric code Hamiltonian [18], and Hamiltonians derived from reversible Markov chains [19, 20, 21]. The definition of stoquastic Hamiltonians can be naturally extended to higher-dimensional spins and to bosonic systems. Notable examples of models in this category are the quantum double Hamiltonian [18], bosonic Hubbard model [7], and Hamiltonians describing "flux-type" Josephson junction qubits [22]. Identifying "easy" and "hard" instances of stoquastic Hamiltonians is therefore important as it could give insights on the power and limitations of quantum Monte Carlo algorithms [23, 24] and contribute to our understanding of speedups in quantum annealing algorithms [25, 26].

Complexity of stoquastic Hamiltonians has been partially characterized in Ref. [9] by proving that Stoq-LH is hard for the complexity class MA and contained in the class AM. Here MA and AM are probabilistic analogues of NP with one and two rounds of communication between the prover and the verifier respectively [10]. It was shown [9] that the complexity of Stoq-LH does not depend on the locality parameter k as long as $2 \leq k \leq O(1)$. A closely related problem of verifying consistency of local reduced density matrices with non-negative matrix elements was studied by Liu [27]. Finally, Stoq-LH is contained in the class

[c] Intuitively, the lack of dependence on J reflects the fact that zero-temperature properties of the ferromagnetic Ising model are trivial — all spins are oriented in the same direction. Since we have included the inverse temperature into the coefficients $J_{u,v}$, the limit $J \to \infty$ corresponds to the zero temperature. Similarly, if $J_{u,v} \to \infty$ for some pair of spins u, v, the problem with n spins can be reduced to the problem with $n - 1$ spins by merging u and v into one effective spin.

QMA (a quantum analogue of NP) since estimating the ground state energy of a general local Hamiltonian is known to be QMA-complete problem [12]. Interestingly, the problem of estimating the *largest* eigenvalue of a local stoquastic Hamiltonian was shown to be QMA-complete by Jordan et al. [28].

TIM occupies a special place in the family of stoquastic Hamiltonians due to its simplicity and a vast body of work devoted to it. In particular, TIM defines one of the four classes in the complexity classification of 2-local quantum Hamiltonians developed recently by Cubitt and Montanaro [29] (with the other three classes being P, NP, and QMA). Experimental implementation of quantum annealing algorithms based on TIM has been demonstrated [30], see also [31, 32] for possible interpretation of these experiments. It was shown recently that any k-local stoquastic Hamiltonian can be represented as an effective low-energy theory emerging from TIM on a constant-degree graph [33]. As a consequence, Stoq-LH for TIM Hamiltonians is as hard as Stoq-LH for general stoquastic Hamiltonians. Theorem 2 implies that a presence of antiferromagnetic spin couplings in a TIM Hamiltonian is essential for this hardness result.

To put Theorem 2 in a broader context, let us briefly discuss the previous work on complexity of the *classical* ferromagnetic Ising model with local magnetic fields,

$$H = -\sum_{1\le u<v\le n} J_{u,v} Z_u Z_v - \sum_{1\le u\le n} h_u Z_u, \qquad J_{u,v} \ge 0.$$

In the case when the local fields are uniform, i.e. $h_u \ge 0$ or $h_u \le 0$ for all u, the minimum energy problem is trivial since the ground state is given by $Z_u = +1$ or $Z_u = -1$ for all u respectively. In the general case when the local fields may take both positive and negative values, the minimum energy of H can be computed in time $O(n^3)$ by a reduction to the Maximum Flow problem [34]. Suppose now that ones goal is to compute the partition function $\mathcal{Z} = \mathrm{Tr}\, e^{-H}$. In the case of uniform local fields $\mathcal{Z}$ is known to admit FPRAS while an exact computation of $\mathcal{Z}$ is $\#P$-hard [15]. In the general case when the local fields may have both positive and negative signs, approximating $\mathcal{Z}$ with a small multiplicative error is known to be as hard as the approximating the number of independent sets in a bipartite graph which is unlikely to have a polynomial time algorithm [35]. Finally, we note that if one does not insist on rigorous bounds on the running time, there exist alternative more practical algorithms for approximating the partition function $\mathcal{Z}$, such as the Swendsen-Wang algorithm [36]. The latter is known to have the running time growing exponentially with n in the case of the ferromagnetic 3-state Potts model [37].

4 Guided Stoquastic Hamiltonians

In this section we prove Theorem 1. We start from the containment in MA which is by far the most difficult part. We shall describe a classical probabilistic algorithm that takes as input a problem instance $\mathcal{I} = (n, H = \sum_\alpha H_\alpha, \lambda_{yes}, \lambda_{no})$ and a witness string $\mathcal{W}$ of length $poly(n)$. The algorithm runs in time $poly(n)$ and outputs 'accept' or 'reject'. Let $P_{acc} = P_{acc}(\mathcal{I}, \mathcal{W})$ be the acceptance probability. The statement that Guided Stoq-LH is contained in promise-MA is equivalent to the following conditions.

Completeness: If $\mathcal{I}$ is a yes-instance then $P_{acc}(\mathcal{I}, \mathcal{W}) \ge 2/3$ for some witness $\mathcal{W}$.

Soundness: If $\mathcal{I}$ is a no-instance then $P_{acc}(\mathcal{I}, \mathcal{W}) \le 1/3$ for any witness $\mathcal{W}$.

Following the standard terminology, we shall refer to the party running the algorithm as Arthur and the party providing the witness as Merlin.

Let $(n, H = \sum_\alpha H_\alpha, \lambda_{yes}, \lambda_{no})$ be an instance of Guided Stoq-LH. Let λ be the ground state energy of H. One can easily check that $-J \le \lambda \le J$, where $J \equiv \sum_\alpha \|H_\alpha\|$. We can assume that $-J \le \lambda_{yes} < \lambda_{no} \le J$ since otherwise the problem becomes trivial. To convince Arthur that $\lambda \le \lambda_{yes}$ Merlin will present a witness that consists of three parts:

- A real number $-J \le \lambda_M \le \lambda_{yes}$
- A classical circuit that computes some function $\phi_M : \{0,1\}^n \to \mathbb{R}$.
- A binary string $x_M \in \{0,1\}^n$

Let us agree that for a yes-instance $\lambda_M = \lambda$ is the ground state energy of H while the function ϕ_M computes amplitudes of some guiding state for H. The string x_M must satisfy several technical conditions stated below. For a no-instance Merlin may try to cheat, that is, λ_M, ϕ_M, x_M could be arbitrary. We shall assume that Arthur rejects the witness right away if it does not fit the specified format, for example, if $\lambda_M > \lambda_{yes}$.

4.1 Verification algorithm

To define Arthur's verification algorithm we first convert H into a sparse non-negative matrix G such that $\|G\| = 1$ for yes-instances and $\|G\| < 1$ for no-instances. To this end choose $\beta = 1/(2J)$ and define

$$G = I - \beta(H - \lambda_M I). \tag{9}$$

Note that $\|\beta(H - \lambda_M I)\| \le \beta(\|H\| + |\lambda_M|) \le 2\beta J = 1$. This shows that G is a positive semidefinite operator with real non-negative matrix elements in the standard basis for both yes- and no-instances. The operator norm of G is equal to its largest eigenvalue, that is,

$$\|G\| = 1 - \beta(\lambda - \lambda_M). \tag{10}$$

Furthermore, ψ is a ground state of H iff ψ is an eigenvector of G with the eigenvalue $\|G\|$. For a yes-instance $\lambda_M = \lambda$, that is $\|G\| = 1$. For a no-instance $\lambda \ge \lambda_{no}$ whereas $\lambda_M \le \lambda_{yes}$. Therefore

$$\text{yes-instance} \quad \Rightarrow \quad \|G\| = 1 \tag{11}$$

$$\text{no-instance} \quad \Rightarrow \quad \|G\| \le 1 - \Delta, \tag{12}$$

where Δ is a "decision gap" defined as

$$\Delta = \beta(\lambda_{no} - \lambda_{yes}) \ge poly(1/n). \tag{13}$$

The next step is to convert G into a random walk. Here we adopt the Projection Monte Carlo method with a variable number of walkers which has been previously proposed in the context of Quantum Monte Carlo simulations [13, 14]. (As was pointed out in Ref. [13], this method has a more favorable scaling of statistical fluctuations in comparison to more widely used Green's function Monte Carlo [23, 5, 6].) A state of the walk is defined as a function

$$\gamma : \{0,1\}^n \to \{0,1,2,\ldots\} \tag{14}$$

that assigns a non-negative integer $\gamma(x)$ to each binary string $x \in \{0,1\}^n$. Loosely speaking, the function γ describes a population of walkers distributed over points of the Boolean cube

$\{0,1\}^n$. The meaning of $\gamma(x)$ is the occupation number of a point x. We shall say that a point x is *empty* or *occupied* if $\gamma(x)=0$ or $\gamma(x)\geq 1$ respectively. Our definition of the walk will depend on three parameters: the number of steps L, a cutoff population size Γ_{max}, and a cutoff guiding state amplitude ϕ_{min}. We shall choose $L=poly(n)$, $\Gamma_{max}=poly(n)$, and $\phi_{min}=2^{-n-1}$. A walk with L steps is a random sequence of states $\gamma_0,\gamma_1,\gamma_2,\ldots,\gamma_L$. We choose the initial state γ_0 as

$$\gamma_0(x)=\begin{cases} 1 & \text{if} \quad x=x_M, \\ 0 & \text{otherwise.} \end{cases} \tag{15}$$

Here x_M is the string received from Merlin. Transition probabilities of the walk will be related to matrix elements of G. Define a non-negative matrix P of size 2^n such that

$$\langle x|P|y\rangle = \frac{\phi(y)}{\phi(x)}\,\langle x|G|y\rangle, \tag{16}$$

where ϕ is a "regularized version" of ϕ_M defined as

$$\phi(x)=\begin{cases} \phi_M(x) & \text{if} \quad \phi_{min}\leq\phi_M(x)\leq 1, \\ 1 & \text{if} \quad \phi_M(x)>1, \\ \phi_{min} & \text{if} \quad \phi_M(x)<\phi_{min}. \end{cases} \tag{17}$$

Recall that $\phi_{min}=2^{-n-1}$. Note that the circuit computing $\phi_M(x)$ can be easily converted to the one computing $\phi(x)$ without substantial increase in size. The matrix element $\langle x|P|y\rangle$ will determine the rate at which walkers are created at a point y at step $t+1$ per each walker located at a point x at step t. More formally, given a real number $p\geq 0$ let $\mathrm{Pois}[p]$ be a non-negative integer random variable drawn from the Poisson distribution with the mean p. In other words, $k=\mathrm{Pois}[p]$ iff $\Pr[\,k\,]=e^{-p}p^k/k!$ for $k\geq 0$. Let us agree that $\mathrm{Pois}[0]=0$ with probability one. For each $y\in\{0,1\}^n$ define

$$\gamma_{t+1}(y)=\sum_{x\in\{0,1\}^n}\mathrm{Pois}[\gamma_t(x)\langle x|P|y\rangle], \tag{18}$$

where all terms represent independent Poisson variables. We shall need these well-known properties of the Poisson distribution:

$$\mathbb{E}[\mathrm{Pois}[p]]=p, \quad \mathbb{E}\Big[\mathrm{Pois}[p]^2\Big]=p^2+p, \quad \mathrm{Pois}[p]+\mathrm{Pois}[q]=\mathrm{Pois}[p+q]. \tag{19}$$

The last equality involves a sum of two independent Poisson-distributed variables. The total population size at a step t is

$$\Gamma_t=\sum_{x\in\{0,1\}^n}\gamma_t(x). \tag{20}$$

Arthur's verification algorithm is defined as follows.

1. Receive λ_M,ϕ_M,x_M from Merlin.

2. Make L steps of the random walk defined above. Abort and reject unless $\Gamma_t\leq\Gamma_{max}$ for all steps t.

3. Accept if $\Gamma_L \geq 1$. Reject if $\Gamma_L = 0$.

It is worth pointing out that if the population becomes empty at some step, that is, $\Gamma_t = 0$, then $\Gamma_{t'} = 0$ for all $t' > t$ with probability one. Thus if Arthur observes $\Gamma_t = 0$ at some step, he can safely abort the protocol and reject the witness right away.

First let us check that Arthur can implement the above algorithm in polynomial time for any Merlin's witness. Suppose Arthur has already implemented the first t steps of the algorithm for some $t \geq 0$ and needs to implement the next step. Obviously, the number of occupied points at step t is at most Γ_t. Since Arthur has not aborted the algorithm yet, one has $\Gamma_t \leq \Gamma_{max}$ and thus there are at most $poly(n)$ occupied points. Arthur can store the function $\gamma_t(x)$ efficiently as a list of pairs $(x, \gamma_t(x))$ which includes only occupied points x. Next Arthur needs to generate the function γ_{t+1} according to Eq. (18). Note that if some point x is empty at step t, that is, $\gamma_t(x) = 0$, such point does not contribute to the sum in Eq. (18) since $\mathrm{Pois}[0] = 0$ with probability one. By the same reason, a point y can be occupied at step $t+1$ with a non-zero probability only if $\langle x|P|y\rangle > 0$ for some point x which is occupied at step t. The number of such points y is at most $poly(n)$ since G (and thus P) has at most $poly(n)$ non-zero matrix elements in each row. Hence for a given function γ_t there are at most $poly(n)$ Poisson variables that Arthur has to generate in order to determine the function γ_{t+1}. To generate each of those variables Arthur has to compute $\langle x|P|y\rangle$. This requires computing $\langle x|G|y\rangle$ and the ratio $\phi(y)/\phi(y)$. Both computations can be done in time $poly(n)$ since G is a sum of local operators while ϕ is described by a polynomial-size circuit. Finally, generating a Poisson random variable with a specified mean can be done in constant time. To avoid complications related to approximating the Poisson distribution we shall assume that Arthur has an access to a device that takes as input a mean p and outputs a random non-negative integer drawn from $\mathrm{Pois}[p]$.

4.2 Miscellaneous

Here we state some basic facts needed for the proof of completeness and soundness conditions. Define a non-negative state

$$|\phi\rangle = \sum_{x\in\{0,1\}^n} \phi(x)\,|x\rangle, \tag{21}$$

where $\phi(x)$ is the regularized version of ϕ_M defined in Eq. (17). Note that ϕ may or may not be normalized. We shall need the first and the second moments of Γ_t. Here and below the probability distribution of Γ_t is obtained by iterating Eq. (18) *without* imposing a constraint $\Gamma_t \leq \Gamma_{max}$.

Lemma 1 *For any $t = 1, \ldots, L$ one has*

$$\mathbb{E}[\Gamma_t] = \frac{1}{\phi(x_M)}\langle x_M|G^t|\phi\rangle, \tag{22}$$

and

$$\mathbb{E}\big[\Gamma_L^2\big] = \frac{1}{\phi(x_M)}\sum_{s=0}^{L}\sum_{y\in\{0,1\}^n}\frac{1}{\phi(y)}\cdot\langle x_M|G^s|y\rangle\cdot\langle y|G^{L-s}|\phi\rangle^2. \tag{23}$$

The proof based on Eq. (19) is a straightforward calculation, so we postpone it until the end of this section.

Lemma 2 *Suppose H admits a guiding state. Then there exists at least one guiding state ϕ such that $\langle x|\phi\rangle \geq 2^{-n-1}$ for all $x \in \{0,1\}^n$.*

Proof: Indeed, let $|\omega\rangle = \sum_x \omega(x)\,|x\rangle$ be some guiding state. By definition, it means that $\omega(x) \geq 0$ for all x, $\sum_x \omega^2(x) = 1$, the function $x \to \omega(x)$ can be computed by a polynomial-size circuit, and H has a non-negative normalized ground state ψ such that $\langle x|\psi\rangle \leq poly(n)\cdot\omega(x)$ for all x. Define a function

$$\phi(x) = C_n(\omega(x) + 2^{-n})$$

where $C_n > 0$ is a constant chosen such that $\sum_x \phi^2(x) = 1$. Simple algebra shows that $1 - 2^{-\Omega(n)} \leq C_n \leq 1$, that is, $C_n \approx 1$ for large n. Thus $\phi(x) \geq 2^{-n-1}$ and $\langle x|\psi\rangle \leq C_n^{-1} poly(n)\phi(x) = poly(n)\phi(x)$ for all x. The function $x \to \phi(x)$ has a polynomial-size circuit since one extra addition and a multiplication by a constant can only increase the circuit size by $poly(\log(n))$. Therefore $|\phi\rangle = \sum_x \phi(x)\,|x\rangle$ is the desired guiding state. □

4.3 Proof of soundness

Consider a no-instance. We have to prove that the acceptance probability P_{acc} is small for any witness λ_M, ϕ_M, x_M. One can get an upper bound on P_{acc} by omitting the tests $\Gamma_t \leq \Gamma_{max}$ since Arthur rejects whenever one of these tests fails. Thus

$$P_{acc} \leq \Pr[\,\Gamma_L \geq 1\,] \leq \mathbb{E}[\Gamma_L] = \frac{1}{\phi(x_M)}\langle x_M|G^L|\phi\rangle \leq \frac{\|\phi\|}{\phi(x_M)}\cdot\|G\|^L. \tag{24}$$

Here we used Eq. (22). By definition of ϕ one has $2^{-n-1} \leq \phi(x) \leq 1$ for all $x \in \{0,1\}^n$, see Eq. (17). Thus $\|\phi\| \leq 2^{n/2}$ and $\phi(x_M) \geq 2^{-n-1}$. Furthermore, $\|G^L\| = \|G\|^L \leq (1-\Delta)^L$, where $\Delta = \beta(\lambda_{no} - \lambda_{yes})$ is the decision gap, see Eq. (12,13). Hence

$$P_{acc} \leq 2^{O(n)}(1-\Delta)^L \leq 2^{O(n)}e^{-\Delta L}. \tag{25}$$

Here we used a bound $1 - s \leq e^{-s}$ which holds for all $s \geq 0$. Recall that $\Delta \geq poly(1/n)$. Hence choosing $L = \Omega(n\Delta^{-1}) = poly(n)$ we can make $P_{acc} \leq 2^{-n}$ for any no-instance and for any Merlin's witness.

4.4 Proof of completeness

Consider a yes-instance and Merlin's witness (λ_M, ϕ_M, x_M), where $\lambda_M = \lambda$ is the ground state energy and ϕ_M computes amplitudes of some guiding state satisfying conditions of Lemma 2. Then the regularized version of ϕ_M defined in Eq. (17) coincides with ϕ_M. Hence $|\phi\rangle = \sum_x \phi(x)\,|x\rangle$ is a normalized guiding state for H. We will prove that $P_{acc} \geq poly(1/n)$ for some choice of the string x_M (although the proof is not constructive). Recall that x_M determines the initial state of the walk, see Eq. (15). Indeed, by definition of the protocol, Arthur accepts iff $\Gamma_t \leq \Gamma_{max}$ for all $t = 1, \ldots, L$ and $\Gamma_L \geq 1$. By the union bound,

$$1 - P_{acc} \leq \Pr[\,\Gamma_L = 0\,] + \sum_{t=1}^{L} \Pr[\,\Gamma_t > \Gamma_{max}\,]. \tag{26}$$

Since Γ_L takes non-negative integer values, one can use the second moment bound:

$$\Pr[\,\Gamma_L \geq 1\,] \geq \frac{\mathbb{E}[\Gamma_L]^2}{\mathbb{E}[\Gamma_L^2]}. \tag{27}$$

By Markov's inequality, $\Pr[\,\Gamma_t > \Gamma_{max}\,] \leq \mathbb{E}[\Gamma_t]/\Gamma_{max}$. Hence

$$P_{acc} \geq \frac{\mathbb{E}[\Gamma_L]^2}{\mathbb{E}[\Gamma_L^2]} - \frac{1}{\Gamma_{max}} \sum_{t=1}^{L} \mathbb{E}[\Gamma_t]. \tag{28}$$

It suffices to show that there exists a string x_M and $poly(n)$ functions $p(n), q(n)$ independent of Γ_{max} such that

$$\frac{\mathbb{E}\big[\Gamma_L^2\big]}{\mathbb{E}[\Gamma_L]^2} \leq p(n) \quad \text{and} \quad \mathbb{E}[\Gamma_t] \leq q(n) \quad \text{for all } t = 1, \ldots, L. \tag{29}$$

(Note that all expectation values above depend on x_M.) Indeed, in this case Eq. (28) implies

$$P_{acc} \geq \frac{1}{p(n)} - \frac{Lq(n)}{\Gamma_{max}} \geq \frac{1}{2p(n)} \tag{30}$$

if we choose $\Gamma_{max} = 2p(n)q(n)L = poly(n)$. Let us now prove existence of a string x_M satisfying Eq. (29). Since ϕ is a guiding state for H, there exists a non-negative normalized ground state $|\psi\rangle = \sum_x \psi(x)\,|x\rangle$ such that

$$\psi(x) \leq r(n) \cdot \phi(x) \quad \text{for all } x, \tag{31}$$

where $r(n) \leq poly(n)$, see Definition 1. Define a set

$$\mathcal{S} = \{x \in \{0,1\}^n \,:\, \frac{\psi(x)}{\phi(x)} \geq \frac{\langle\psi|\phi\rangle}{2}\} \tag{32}$$

and a probability distribution

$$\pi(x) = \frac{\psi(x)\phi(x)}{\langle\psi|\phi\rangle}. \tag{33}$$

(One may think of π as a "steady state" of P since $\pi P = \pi$.) We claim that

$$\pi(\mathcal{S}) \equiv \sum_{x\in\mathcal{S}} \pi(x) \geq \frac{1}{2}. \tag{34}$$

Indeed, one has

$$1 = \sum_x \pi(x) = \pi(\mathcal{S}) + \langle\psi|\phi\rangle^{-1} \sum_{x\notin\mathcal{S}} \phi^2(x) \cdot \frac{\psi(x)}{\phi(x)} \leq \pi(\mathcal{S}) + \frac{1}{2}\sum_{x\notin\mathcal{S}} \phi^2(x) \leq \pi(\mathcal{S}) + \frac{1}{2}. \tag{35}$$

Here the last bound uses normalization of ϕ. This proves Eq. (34). For each $t = 1, \ldots, L$ define a function

$$V_t(x) = \frac{\langle x|G^t|\phi\rangle}{\phi(x)}.$$

From Eq. (22) one infers that $\mathbb{E}[\Gamma_t] = V_t(x_M)$. Furthermore,

$$\sum_x \pi(x) V_t(x) = \langle\psi|\phi\rangle^{-1} \sum_x \psi(x)\langle x|G^t|\phi\rangle = 1 \tag{36}$$

since $G^t\psi=\psi$. Next define a function

$$W(x)=\sum_{s=0}^{L}\sum_{y}\frac{\langle x|G^s|y\rangle\cdot\langle y|G^{L-s}|\phi\rangle^2}{\phi(x)\phi(y)}.$$

Here the second sum is over all $y\in\{0,1\}^n$. From Eq. (23) one infers that $\mathbb{E}\big[\Gamma_L^2\big]=W(x_M)$. Taking into account that $G^s\psi=\psi$ for any s one gets

$$\sum_x \pi(x)W(x)=\langle\psi|\phi\rangle^{-1}\sum_{s=0}^{L}\sum_{y}\frac{\psi(y)}{\phi(y)}\cdot\langle y|G^{L-s}|\phi\rangle^2. \tag{37}$$

Note that $\psi(y)/\phi(y)\le r(n)$ and $\langle\psi|\phi\rangle=\sum_x\psi^2(x)\phi(x)/\psi(x)\ge 1/r(n)$ due to Eq. (31). Since G has non-negative matrix elements, $\|G\|=1$, and $\|\phi\|=1$ we arrive at

$$\sum_x \pi(x)W(x)\le r^2(n)\sum_{s=0}^{L}\sum_y\langle y|G^{L-s}|\phi\rangle^2=r^2(n)\sum_{s=0}^{L}\langle\phi|G^{2(L-s)}|\phi\rangle\le r^2(n)(1+L). \tag{38}$$

Combining Eqs. (34,36,38) results in

$$\frac{1}{\pi(\mathcal{S})}\sum_{x\in\mathcal{S}}\pi(x)\left[W(x)+\sum_{t=1}^{L}V_t(x)\right]\le 2(L+1)(1+r^2(n))\equiv q(n). \tag{39}$$

Therefore there must exist $x_M\in\mathcal{S}$ such that

$$W(x_M)\le q(n)\quad\text{and}\quad V_t(x_M)\le q(n)\quad\text{for all } t=1,\ldots,L. \tag{40}$$

Furthermore, for any $x\in\mathcal{S}$ one has

$$V_L(x)=\frac{\langle x|G^L|\phi\rangle}{\phi(x)}\ge\frac{\langle x|G^L|\psi\rangle}{r(n)\phi(x)}=\frac{\psi(x)}{r(n)\phi(x)}\ge\frac{\langle\psi|\phi\rangle}{2r(n)}\ge\frac{1}{2r^2(n)}. \tag{41}$$

This shows that there exists $x_M\in\mathcal{S}$ such that

$$\frac{\mathbb{E}\big[\Gamma_L^2\big]}{\mathbb{E}[\Gamma_L]^2}=\frac{W(x_M)}{V_L(x_M)^2}\le 4r^4(n)q(n)\equiv p(n)\quad\text{and}\quad\mathbb{E}[\Gamma_t]\le q(n)\quad\text{for all } t=1,\ldots,L. \tag{42}$$

This proves the desired bounds in Eq. (29) and shows that Merlin can make Arthur to accept with probability at least $P_{acc}\ge 1/2p(n)\ge poly(1/n)$ provided that $\Gamma_{max}=\Omega(L^3r^8(n))$.

Arthur can achieve the acceptance probability at least $2/3$ as required for the completeness condition by implementing about $2p(n)$ independent rounds of the above algorithm with the same witness. Arthur accepts the witness iff at least one of the rounds outputs 'accept'. For any no-instance each round accepts with probability at most 2^{-n} and thus the full algorithm accepts with probability less than $1/3$ for large enough n. This completes the proof that Guided Stoq-LH is contained in promise-MA.

Finally, the statement that Guided Stoq-LH is complete for promise-MA if $k\ge 6$ follows trivially from Ref. [9]. Indeed, let $\mathcal{L}=\mathcal{L}_{yes}\cup\mathcal{L}_{no}$ be any language in promise-MA, where $\mathcal{L}_{yes}$ and $\mathcal{L}_{no}$ are the sets of yes- and no-instances. Without loss of generality we can assume

that Arthur's verification protocol has perfect completeness [38], that is, for any yes-instance $\mathcal{I} \in \mathcal{L}_{yes}$ there exists a witness $\mathcal{W}$ such that $P_{acc}(\mathcal{I}, \mathcal{W}) = 1$. Using Lemma 3 of Ref. [9] one can efficiently transform any instance $\mathcal{I} \in \mathcal{L}$ into a 6-local stoquastic Hamiltonian H such that its ground state energy satisfies $\lambda = 0$ if $\mathcal{I} \in \mathcal{L}_{yes}$ and $\lambda \geq poly(1/n)$ if $\mathcal{I} \in \mathcal{L}_{no}$. Furthermore, for any yes-instance $\mathcal{I}$ one can choose a ground state of H as a coherent superposition of all computational branches of Arthur's verification algorithm (for some fixed witness $\mathcal{W}$ such that $P_{acc}(\mathcal{I}, \mathcal{W}) = 1$). Hence computing the amplitudes of ψ is equivalent to checking whether a sequence of binary strings represent a valid computational path of the algorithm. This can be checked in time $poly(n)$. The amplitude of any valid computational path depends in a simple way on the number of random bits and the number of computational steps in the verification algorithm, see Ref. [9] for details. Therefore for a yes-instance the ground state ψ itself can be chosen as a guiding state. This shows that any problem in promise-MA can be reduced to Guided Stoq-LH with $k = 6$.

4.5 Open problems

A natural question is whether the above algorithm can be used to estimate the ground state energy without Merlin's assistance assuming that one has a good guess of the guiding state. One possible strategy would be to sweep λ_M over a region that is likely to contain the ground state energy and estimate Arthur's acceptance probability P_{acc} for each value of λ_M by Monte Carlo simulation. One should expect that P_{acc} is negligible unless $\lambda_M \approx \lambda$. Indeed, recall that $\|G\| = 1 - \beta(\lambda - \lambda_M)$. If $\lambda_M > \lambda$ then $\|G\| > 1$ which leads to an exponential growth of the population, see Eq. (22). Accordingly, the test $\Gamma_t \leq \Gamma_{max}$ is likely to fail. On the other hand, if $\lambda_M < \lambda$ one has $\|G\| < 1$ and the final population is likely to be empty. In both cases the outcome of the protocol is 'reject'. Unfortunately, proving that P_{acc} is non-negligible for $\lambda \approx \lambda_M$ requires a careful choice of the initial string x_M. A preliminary analysis shows that x_M can be chosen efficiently if the guiding state obeys a stronger condition $poly(1/n)\langle x|\psi\rangle \leq \langle x|\phi\rangle \leq poly(n)\langle x|\psi\rangle$ for all x.

One may also ask whether Theorem 1 holds for some weaker notion of a guiding state. For example, the pointwise correlation condition in Eq. (4) appears to be unreasonably strong. It would be very desirable to replace Eq. (4) by a bound on some global correlation measure that has a clear physical meaning. Ideally, a guiding state ϕ just needs to have an overlap $\geq poly(1/n)$ with some exact ground state ψ and have efficiently computable amplitudes. A preliminary analysis shows that for Hamiltonians with a polynomial spectral gap a guiding state only needs to satisfy a condition $\sum_x \langle x|\psi\rangle^3/\langle x|\phi\rangle \leq poly(n)$. Alternatively, one can study guiding states ϕ that admit an efficient classical algorithm for sampling a basis vector x from the distribution $\langle x|\phi\rangle^2$.

Finally, it is important to identify non-trivial classes of stoquastic Hamiltonians that actually admit a guiding state. For example, results of Ref. [11] imply that for any frustration-free stoquastic Hamiltonian one can always choose a non-negative ground state with efficiently computable amplitudes. In this case the ground state itself can serve as a guiding state. We anticipate that non-trivial examples of guiding states could be found among tensor network states such as Matrix Product States or PEPS [39]. We conjecture that the ferromagnetic TIM Hamiltonians studied in the next section admit a guiding state.

4.6 Proof of Lemma 1

For any integers $0 \le t \le s$ define a random variable

$$\Gamma_{t,s} = \begin{cases} \Gamma_t & \text{if} \quad s = 0, \\ \sum_{y,z} \gamma_t(y) \langle y|P^s|z\rangle & \text{if} \quad s \ge 1. \end{cases} \tag{43}$$

Applying repeatedly Eq. (19) one can easily show that

$$\mathbb{E}[\Gamma_{t,s}] = \mathbb{E}[\Gamma_{t-1,s+1}]. \tag{44}$$

Note that $\Gamma_{0,t}$ is a deterministic variable,

$$\Gamma_{0,t} = \sum_z \langle x_M|P^t|z\rangle. \tag{45}$$

Hence

$$\mathbb{E}[\Gamma_t] = \mathbb{E}[\Gamma_{t,0}] = \mathbb{E}[\Gamma_{0,t}] = \sum_z \langle x_M|P^t|z\rangle = \frac{1}{\phi(x_M)} \langle x_M|G^t|\phi\rangle. \tag{46}$$

This proves Eq. (22). Let us now compute $\mathbb{E}\left[\Gamma_{t,s}^2\right]$ with respect to the conditional distribution

$$\Pr[\,\gamma_t|\gamma_1,\ldots,\gamma_{t-1}\,] = \Pr[\,\gamma_t|\gamma_{t-1}\,].$$

For a fixed γ_{t-1} the variables $\gamma_t(y)$ are independent Poisson-distributed variables so that

$$\mathbb{E}[\gamma_t(y)] = \sum_x \gamma_{t-1}(x) \langle x|P|y\rangle \tag{47}$$

and

$$\mathbb{E}[\gamma_t(y)\gamma_t(y')] = \mathbb{E}[\gamma_t(y)] \cdot \mathbb{E}[\gamma_t(y')] + \delta_{y,y'} \mathbb{E}[\gamma_t(y)]. \tag{48}$$

Using Eq. (47,48) one easily gets

$$\mathbb{E}\left[\Gamma_{t,s}^2\right] = \Gamma_{t-1,s+1}^2 + \sum_{x,y,z,z'} \gamma_{t-1}(x) \langle x|P|y\rangle \cdot \langle y|P^s|z\rangle \cdot \langle y|P^s|z'\rangle. \tag{49}$$

Taking the expectation value over the distribution of $\gamma_1,\ldots,\gamma_{t-1}$ leads to

$$\mathbb{E}\left[\Gamma_{t,s}^2\right] = \mathbb{E}\left[\Gamma_{t-1,s+1}^2\right] + \sum_{y,z,z'} \langle x_M|P^t|y\rangle \cdot \langle y|P^s|z\rangle \cdot \langle y|P^s|z'\rangle. \tag{50}$$

Iterating Eq. (50) starting from $t = L$, $s = 0$ and using Eq. (45) one gets

$$\mathbb{E}\left[\Gamma_L^2\right] = \sum_{\substack{s+t=L \\ s,t\ge 0}} \sum_y \langle x_M|P^t|y\rangle \left[\sum_z \langle y|P^s|z\rangle\right]^2. \tag{51}$$

Substituting the definition of P, see Eq. (16), leads to Eq. (23). This proves Lemma 1.

5 Approximating the partition function of TIM

In this section we prove Theorem 2. We shall use a notation $\rho(A)$ for the difference between the largest and the smallest eigenvalue of a hermitian matrix A. Note that $\rho(A) \le 2\|A\|$ for any matrix A. The following lemma will be needed to control the error in the Suzuki-Trotter approximation.

Lemma 3 *Consider any pair of hermitian operators A, B and let $\rho = \rho(A) + \rho(B)$. For any $0 \le t \le (2\rho)^{-1}$ there exists a hermitian operator D such that $\|D\| \le 12\rho^3$ and*

$$e^{At/2} e^{Bt} e^{At/2} = e^{(A+B)t + Dt^3}. \tag{52}$$

Since the proof involves a straightforward calculation, we postpone it until the end of this section. Consider now a TIM Hamiltonian

$$H = -A - B, \quad A = \sum_{1 \le u < v \le n} J_{u,v} Z_u Z_v, \quad B = \sum_{1 \le u \le n} h_u X_u, \tag{53}$$

where $J_{u,v} \ge 0$ for all u, v (the ferromagnetic case). Let $\mathcal{Z} = \mathrm{Tr}\big(e^{A+B}\big)$ be the partition function. For any integer $r \ge 1$ define a Suzuki-Trotter approximation to $\mathcal{Z}$ as

$$\mathcal{Z}' = \mathrm{Tr}\left(e^{At} e^{Bt}\right)^r, \quad t \equiv r^{-1}. \tag{54}$$

Note that $\rho \equiv \rho(A) + \rho(B) \le 2(\|A\| + \|B\|) \le poly(n, J)$. Suppose $r \ge 2\rho$ so that $0 \le t \le (2\rho)^{-1}$. Then Lemma 3 implies

$$\mathcal{Z}' = \mathrm{Tr}\left(e^{At/2} e^{Bt} e^{At/2}\right)^r = \mathrm{Tr}\, e^{r\left[(A+B)t + Dt^3\right]} = \mathrm{Tr}\, e^{A+B+C}, \tag{55}$$

where $C = Dr^{-2}$ is a hermitian operator such that $\|C\| \le 12\rho^3 r^{-2}$.

Let λ_i and λ'_i be the i-th largest eigenvalue of the Hamiltonian $A + B$ and $A + B + C$ respectively, where $i = 1, \ldots, 2^n$. By Weyl's inequality and Lemma 3,

$$|\lambda_i - \lambda'_i| \le \|C\| \le \frac{12\rho^3}{r^2}. \tag{56}$$

Choosing

$$r = \max\left[2\rho, \sqrt{12\rho^3 \delta^{-1}}\right] \tag{57}$$

guarantees that conditions of Lemma 3 are satisfied and $\|C\| \le \delta$. Note that $r \le poly(n, J, \delta^{-1})$ since $\rho \le poly(n, J)$. Then $|\lambda'_i - \lambda_i| \le \delta$ and

$$\mathcal{Z}' = \sum_{i=1}^{2^n} e^{\lambda'_i} \le \sum_{i=1}^{2^n} e^{\lambda_i + \delta} \approx (1 + \delta)\mathcal{Z}. \tag{58}$$

Here we assumed for simplicity that $\delta \ll 1$. The same arguments show that $\mathcal{Z}' \ge (1 - \delta)\mathcal{Z}$. Hence $\mathcal{Z}'$ approximates $\mathcal{Z}$ up to a multiplicative error δ.

The remaining step is the standard quantum-to-classical mapping that relates $\mathcal{Z}'$ to the partition function of a classical Ising model [1]. Let $\sigma = (\sigma_1, \ldots, \sigma_n) \in \{\pm 1\}^n$ be a configuration of n classical Ising spins and $|\sigma\rangle$ be the corresponding basis state of n qubits. Then

$$e^{At} = \sum_{\sigma \in \{\pm 1\}^n} e^{E_A(\sigma)} |\sigma\rangle\langle\sigma| \quad \text{where} \quad E_A(\sigma) = \sum_{1 \le u < v \le n} t J_{u,v}\, \sigma_u \sigma_v \tag{59}$$

and

$$e^{Bt} = \Gamma \sum_{\sigma,\sigma' \in \{\pm 1\}^n} e^{E_B(\sigma,\sigma')} |\sigma\rangle\langle\sigma'| \quad \text{where} \quad E_B(\sigma,\sigma') = \sum_{1 \le u \le n} \tilde{h}_u \sigma_u \sigma'_u. \tag{60}$$

Here

$$\tilde{h}_u = -\frac{1}{2} \log\left[\tanh\left(t h_u\right)\right] \quad \text{and} \quad \Gamma = 2^{-n/2} \prod_{u=1}^{n} \sqrt{\sinh\left(2t h_u\right)}. \tag{61}$$

Note that $\tilde{h}_u \ge 0$ since we assumed $h_u \ge 0$. From Eqs. (59,60) one gets

$$\mathcal{Z}' = \mathrm{Tr}\left(e^{At} e^{Bt}\right)^r = \Gamma^r \sum_{\sigma^1,\ldots,\sigma^r \in \{\pm 1\}^n} \exp\left[\sum_{i=1}^{r} E_A(\sigma^i) + E_B(\sigma^i, \sigma^{i+1})\right] \equiv \sum_{\theta \in \{\pm 1\}^{nr}} e^{E(\theta)}. \tag{62}$$

The energy function $E(\theta)$ describes r copies of the n-spin ferromagnetic Ising model such that the i-th copy has the energy function $E_A(\sigma^i)$ and each consecutive pair of copies is coupled by the energy function $E_B(\sigma^i, \sigma^{i+1})$. The multiplicative factor Γ^r can be absorbed into a constant energy shift in $E(\theta)$. Since this factor is easy to compute, for simplicity we shall ignore it. Hence $\mathcal{Z}'$ is the partition function of a classical ferromagnetic Ising model. We can now invoke Theorem 5 of Ref. [15]. It asserts that $\mathcal{Z}'$ admits FPRAS with a running time $O(\delta^{-2} M^3 N^{11} \log N)$, where $N = nr$ is the total number of spins and M is the number of non-zero spin-spin couplings. Note that in our case $M \le nr + \binom{n}{2} r \le n^2 r$. Using Eq. (57) with a conservative estimate $\rho \le n^2 J$ one gets $r \le n^3 J^{3/2} \delta^{-1/2}$. Thus $N \le n^4 J^{3/2} \delta^{-1/2}$, $M \le n^5 J^{3/2} \delta^{-1/2}$ and the FPRAS for $\mathcal{Z}'$ has running time $O(n^{59} J^{21} \delta^{-9})$, ignoring logarithmic factors. Since $\mathcal{Z}'$ approximates $\mathcal{Z}$ with a multiplicative error δ, this proves Theorem 2.

Remark: We note that $\tilde{h}_u$ becomes infinite if $h_u = 0$. One can always assume that $h_u \ge \delta/n$ since changing the Hamiltonian by a perturbation of norm at most δ leads to a multiplicative error of order δ in the partition function.

5.1 Proof of Lemma 3

Since Eq. (52) is invariant under a shift $A \to A + cI$, we can assume that the largest eigenvalue of A is zero. Then $A \le 0$ and $\|A\| = \rho(A)$. Likewise, we can assume that $B \le 0$ and $\|B\| = \rho(B)$. Consider the Taylor series

$$Z(t) \equiv e^{At/2} e^{Bt} e^{At/2} = \sum_{p=0}^{\infty} Z_p \, t^p$$

that converges absolutely for any $t \in \mathbb{C}$. One can easily check that

$$Z_0 = I, \quad Z_1 = A + B, \quad Z_2 = Z_1^2/2.$$

We shall upper bound the norm of higher order coefficients using the Cauchy's formula,

$$Z_p = \frac{1}{2\pi i} \oint_{|t|=R} \frac{Z(t) dt}{t^{p+1}}.$$

Since $Z(t)$ is analytic in the full complex plane, the radius R can be chosen arbitrarily. Let us choose $R = \rho^{-1}$ and let C be the maximum of $\|Z(t)\|$ over the circle $|t| = R$. Then

$\|Z_p\| \le CR^{-p} = C\rho^p$ and

$$Z(t) = \sum_{p=0}^{2} Z_p t^p + \Delta \quad \text{where} \quad \|\Delta\| \le C \sum_{p\ge 3} (\rho t)^p \le 2C\rho^3 t^3.$$

Here we used the assumption that $0 \le t \le (2\rho)^{-1}$. Note that

$$\|Z(t)\| \le \|e^{At/2}\|^2 \cdot \|e^{Bt}\| \le e^{(\|A\|+\|B\|)|t|} = e^{\rho|t|} = e.$$

for any complex t with $|t| = R = \rho^{-1}$. Hence $C \le e$.

Choose any $0 \le t \le (2\rho)^{-1}$ and define

$$U(t) \equiv Z(t) - I = Z_1 t + Z_2 t^2 + \Delta.$$

Taking into account that $e^{At/2} \le I$ and $e^{Bt} \le I$ one easily gets

$$\|U(t)\| \le 2\|e^{At/2} - I\| + \|e^{Bt} - I\| \le t(\|A\| + \|B\|) = \rho t \le \frac{1}{2}. \tag{63}$$

Here we used the inequality $e^x \ge 1 + x$ which holds for $x \le 0$. One can easily check that

$$\|\log(I + U) - U + U^2/2\| \le C'\|U\|^3, \tag{64}$$

for any hermitian operator U such that $\|U\| \le 1/2$, where

$$C' \equiv \max_{x\,:\,|x|\le 1/2} \frac{1}{|x|^3} \left|\log(1+x) - x + x^2/2\right| \approx 0.55.$$

Choose the desired operator D as

$$D = t^{-3}\left[-(A+B)t + \log\left(e^{At/2}e^{Bt}e^{At/2}\right)\right] = t^{-3}\left[-Z_1 t + \log(I + U(t))\right]. \tag{65}$$

By definition, D is hermitian. Combining Eq. (63,64) and using the union bound one gets

$$\|D\| \le t^{-3}\| - Z_1 t + U(t) - U(t)^2/2\| + C'\rho^3.$$

A simple algebra shows that

$$\Gamma \equiv -Z_1 t + U(t) - U(t)^2/2 = \Delta - \frac{Z_1^4 t^4}{8} - \frac{\Delta^2}{2} - \frac{Z_1^3 t^3}{2} - Z_1 \Delta t - \frac{Z_1^2 \Delta t^2}{2}.$$

Taking into account that $\|Z_1\| \le \rho$ and $\rho t \le 1/2$ one arrives at

$$\|\Gamma\| \le (\rho t)^3 \left(2C + \frac{1}{16} + \frac{C^2}{4} + \frac{1}{2} + C + \frac{C}{4}\right) \le (\rho t)^3 \left(\frac{9}{16} + \frac{13C}{4} + \frac{C^2}{4}\right).$$

Therefore

$$\|D\| \le \rho^3 \left(\frac{9}{16} + \frac{13C}{4} + \frac{C^2}{4} + C'\right).$$

Substituting $C \le e$ and $C' \le 0.56$ leads to $\|D\| \le 12\rho^3$.

Acknowledgments

The author would like to thank Toby Cubitt and Graeme Smith for helpful discussions. This work was supported in part by the DARPA QuEST program under contract number HR0011-09-C-0047.

References

1. Subir Sachdev. *Quantum Phase Transitions*. Cambridge University Press, 2007.
2. M. Suzuki, S. Miyashita, and A. Kuroda. Monte Carlo simulation of quantum spin systems. I. *Prog. Theor. Phys.*, 58(5):1377–1387, 1977.
3. N. Prokof'ev, B. Svistunov, and I. Tupitsyn. Exact quantum Monte Carlo process for the statistics of discrete systems. *Pis'ma v Zh.Eks.Teor.Fiz.*, 64(12):911–916, 1996.
4. A. W. Sandvik and J. Kurkijärvi. Quantum Monte Carlo simulation method for spin systems. *Phys. Rev. B*, 43(7):5950, 1991.
5. N. Trivedi and D.M. Ceperley. Green-function Monte Carlo study of quantum antiferromagnets. *Phys. Rev. B*, 40(4):2737, 1989.
6. M. C. Buonaura and S. Sorella. Numerical study of the two-dimensional Heisenberg model using a Green function Monte Carlo technique with a fixed number of walkers. *Phys. Rev. B*, 57(18):11446, 1998.
7. S. Wessel, F. Alet, M. Troyer, and G. G. Batrouni. Quantum Monte Carlo simulations of confined bosonic atoms in optical lattices. *Phys. Rev. A*, 70(5):053615, 2004.
8. M. Troyer and U.-J. Wiese. Computational complexity and fundamental limitations to fermionic quantum Monte Carlo simulations. *Phys. Rev. Lett*, 94(17):170201, 2005.
9. S. Bravyi, D. DiVincenzo, R. Oliveira, and B. Terhal. The complexity of stoquastic local Hamiltonian problems. *Quant. Inf. Comp.*, 8(5):0361–0385, 2008.
10. L. Babai. Trading group theory for randomness. In *Proc. of 17th ACM STOC*, pages 421–429. ACM, 1985.
11. S. Bravyi and B. Terhal. Complexity of stoquastic frustration-free hamiltonians. *SIAM J. on Comp.*, 39(4):1462–1485, 2009.
12. A. Kitaev, A. Shen, and M. Vyalyi. *Classsical and quantum computation*. AMS, 2002.
13. N. Cerf and O. Martin. Projection Monte Carlo methods: an algorithmic analysis. *Int. J. Mod. Phys.* **C**, 06:693, 1995.
14. Roberto I. Oliveira. private communication, 2006.
15. M. Jerrum and A. Sinclair. Polynomial-time approximation algorithms for the Ising model. *SIAM J. on Comp.*, 22(5):1087–1116, 1993.
16. 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(5516):472–475, 2001.
17. R. Somma, D. Nagaj, and M. Kieferová. Quantum speedup by quantum annealing. *Phys. Rev. Lett.*, 109:050501, 2012.
18. A. Yu. Kitaev. Fault-tolerant quantum computation by anyons. *Ann. of Phys.*, 303(1):2–30, 2003.
19. D. Aharonov and A. Ta-Shma. Adiabatic quantum state generation and statistical zero knowledge. In *Proc. 35th ACM STOC*, pages 20–29. ACM, 2003.
20. C. L. Henley. From classical to quantum dynamics at Rokhsar–Kivelson points. *J. of Phys.: Cond. Mat.*, 16(11):S891, 2004.
21. F. Verstraete, M. M. Wolf, D. Perez-Garcia, and J. I. Cirac. Criticality, the area law, and the computational power of projected entangled pair states. *Phys. Rev. Lett.*, 96(22):220601, 2006.
22. M.H. Devoret, A. Wallraff, and J.M. Martinis. Superconducting qubits: a short review. *arXiv preprint cond-mat/0411174*, 2004.
23. J.H. Hetherington. Observations on the statistical iteration of matrices. *Phys. Rev. A*, 30(5):2713, 1984.
24. M. B. Hastings and M.H. Freedman. Obstructions to classically simulating the quantum adiabatic

algorithm. *Quant. Inf. & Comp.*, 13(11-12):1038–1076, 2013.
25. E. Farhi, J. Goldstone, D. Gosset, S. Gutmann, H. B. Meyer, and P. Shor. Quantum adiabatic algorithms, small gaps, and different paths. *Quant. Inf. Comp.*, 11(3):181–214, 2011.
26. E. Farhi, D. Gosset, I. Hen, A.W. Sandvik, P. Shor, A.P. Young, and F. Zamponi. Performance of the quantum adiabatic algorithm on random instances of two optimization problems on regular hypergraphs. *Phys. Rev. A*, 86(5):052334, 2012.
27. Yi-Kai Liu. The local consistency problem for stoquastic and 1D quantum systems. *arXiv preprint 0712.1388*, 2007.
28. S. P. Jordan, D. Gosset, and P. J. Love. QMA-complete problems for stoquastic Hamiltonians and Markov matrices. *Phys. Rev. A*, 81(3):032331, 2010.
29. T. Cubitt and A. Montanaro. Complexity classification of local hamiltonian problems. *arXiv preprint 1311.3161*, 2013.
30. S. Boixo, T. F. Rønnow, S. Isakov, Z. Wang, D. Wecker, D. Lidar, J. M. Martinis, and M. Troyer. Quantum annealing with more than one hundred qubits. *arXiv preprint 1304.4595*, 2013.
31. T. F. Rønnow, Z. Wang, J. Job, S. Boixo, S. Isakov, D. Wecker, J. Martinis, D. Lidar, and M. Troyer. Defining and detecting quantum speedup. *arXiv preprint 1401.2910*, 2014.
32. S. W. Shin, G. Smith, J. Smolin, and U. Vazirani. How "quantum" is the D-wave machine? *arXiv preprint 1401.7087*, 2014.
33. S. Bravyi and M. B. Hastings. On complexity of the quantum Ising model. *preprint arXiv:1410.0703*, 2014.
34. H. Rieger. Frustrated systems: Ground state properties via combinatorial optimization. *preprint arXiv:cond-mat/9705010*, 1997.
35. L. Goldberg and M. Jerrum. The complexity of ferromagnetic Ising with local fields. *Combinatorics, Probability and Computing*, 16(01):43–61, 2007.
36. R. Swendsen and J.-S. Wang. Nonuniversal critical dynamics in Monte Carlo simulations. *Phys. Rev. Lett.*, 58(2):86, 1987.
37. V. K. Gore and M. R. Jerrum. The Swendsen–Wang process does not always mix rapidly. *J. of Stat. Phys.*, 97(1-2):67–86, 1999.
38. M. Furer, O. Goldreich, and Y. Mansour. On completeness and soundness in interactive proof systems. *Advances in Computing Research*, 5:429–442, 1989.
39. F. Verstraete, V. Murg, and J.I. Cirac. Matrix product states, projected entangled pair states, and variational renormalization group methods for quantum spin systems. *Advances in Physics*, 57(2):143–224, 2008.