A Sharp Computational Phase Transition for the Partition Function of the Transverse-Field Ising Model

Alistair Sinclair
Computer Science Division, UC Berkeley, CA, USA

Thuy-Duong Vuong
Department of Computer Science and Engineering, UC San Diego, La Jolla, CA, USA

Abstract

We study the problem of approximating the partition function of the transverse-field Ising model (TFIM), a widely studied quantum many-body model with important applications in quantum simulation and quantum annealing. Despite its fundamental importance, the algorithmic landscape for computing the TFIM partition function has remained poorly understood beyond restricted parameter regimes.

We provide a precise characterization of the temperature regimes in which efficient approximation is possible, establishing a sharp computational phase transition. Let J denote the symmetric interaction matrix and $\Delta(J) = \lambda_{\max}(J) - \lambda_{\min}(J)$ denote its spectral width. We show that, for all inverse temperatures $\beta \in [0, 1/\Delta(J)]$, there exists an efficient classical randomized algorithm that approximates the partition function $\mathrm{tr}(e^{-\beta H})$ to within an arbitrarily small multiplicative factor. To obtain this result, we apply the standard Trotter decomposition to map the quantum model to a classical spin system, and then leverage new techniques in Markov chain analysis to derive an efficient algorithm that samples from and computes the partition function of the resulting distribution. This temperature threshold is tight: for $\beta > 1/\Delta(J)$, we show that there exists interaction matrix J for which approximating the partition function, even within an exponential factor, is NP-hard and thus is unlikely to admit an efficient classical or quantum algorithm.

2012 ACM Subject Classification Theory of computation → Random walks and Markov chains; Theory of computation → Design and analysis of algorithms

Keywords and phrases Computational phase transition, partition function, transverse-field Ising, sampling, counting

Digital Object Identifier 10.4230/LIPIcs.TQC.2026.1

Funding *Alistair Sinclair*: Supported in part by NSF grant CCF-223109.

1 Introduction

A central problem in quantum statistical mechanics and quantum information theory is to determine when the partition function of a quantum many-body system can be efficiently approximated. The partition function encodes the system's thermodynamic behavior and plays a fundamental role in quantum simulation, quantum phase transitions, and complexity theory [6].

In this work, we consider the *transverse-field Ising model* (TFIM), one of the most extensively studied quantum spin systems. The TFIM serves as a canonical model for quantum phase transitions [26, 12], and plays an important role in quantum annealing [21, 24, 31]. Despite its central role, prior work has made only limited progress in understanding the computational complexity of approximating its partition function, with existing results confined to restricted parameter regimes such as the ferromagnetic case [5] or very high temperature [11].

We consider the general TFIM with Hamiltonian

$$H = \frac{1}{2} \sum_{i,j\in[n]} J_{ij} Z_i Z_j + \sum_{i=1}^{n} h_i^z Z_i + \sum_{i=1}^{n} h_i^x X_i, \tag{1}$$

where $Z_i = I^{\otimes(i-1)} \otimes Z \otimes I^{\otimes(n-i)}$ and $X_i = I^{\otimes(i-1)} \otimes X \otimes I^{\otimes(n-i)}$ denote the Pauli operators acting on site i, with

$$Z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \qquad X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Here $J = (J_{ij})_{i,j\in[n]} \in \mathbb{R}^{n\times n}$ is the *interaction matrix*, $\mathbf{h}^x = (h_i^x)_{i\in[n]} \in \mathbb{R}^n$ is the *transverse field*, and $\mathbf{h}^z = (h_i^z)_{i\in[n]} \in \mathbb{R}^n$ is the *longitudinal field*.

Without loss of generality, we may assume that J is symmetric, as symmetrizing by replacing J with $\frac{J+J^\intercal}{2}$ leaves the Hamiltonian invariant. Since J is symmetric, its eigenvalues are real. We let $\lambda_{\max}(J)$ and $\lambda_{\min}(J)$ denote the maximum and minimum eigenvalues of J, and let $\Delta(J) = \lambda_{\max}(J) - \lambda_{\min}(J)$ be its spectral width.

We provide a precise characterization of the temperature regimes in which the partition function of the TFIM can be efficiently approximated. Our main result establishes a sharp computational phase transition: for all inverse temperatures β in the interval $[0, \frac{1}{\Delta(J)}]$, there exists an efficient *classical* algorithm for approximating the partition function $\mathrm{tr}(\exp(-\beta H))$ (see Theorem 20 and Theorem 1). Our approximation algorithm applies in a much broader range of parameters than previous works (see Section 1.2 for details).

For $\beta > 1/\Delta(J)$, we show that the problem is NP-hard, and thus unlikely to admit any efficient *classical or quantum* algorithm (see Theorem 2). To the best of our knowledge, this is the first example of a quantum many-body problem where approximating the partition function exhibits a provable computational phase transition.

▶ **Theorem 1** (Informal; see Theorem 20 for the formal statement)**.** *For any $\beta \in [0, \frac{1}{\Delta(J)}]$, and $\varepsilon, \delta \in (0,1)$, there exists a classical randomized algorithm running in*

$$\log \delta^{-1} \cdot \mathrm{poly}(n, \beta \left\|\boldsymbol{h}^x\right\|_1, \beta \left\|\boldsymbol{h}^z\right\|_1, \varepsilon^{-1})$$

time that outputs $\hat{Z}$ such that

$$\Pr\left(e^{-\varepsilon}\hat{Z} \le \mathrm{tr}(\exp(-\beta H)) \le e^{\varepsilon}\hat{Z}\right) \ge 1 - \delta.$$

▶ **Theorem 2.** *For any $\varepsilon > 0$, there exists an interaction matrix J such that at $\beta = \frac{1+\varepsilon}{\Delta(J)}$, approximating the partition function $\mathrm{tr}(\exp(-\beta H))$ is NP-hard for any transverse field $\mathbf{h}^x$ satisfying $\left\|\beta \mathbf{h}^x\right\|_1 \le c(\varepsilon)n$, where $c(\varepsilon) > 0$ is a constant dependent only on ε. Moreover, assuming* RP $\neq$ NP *(respectively,* BQP $\not\supseteq$ NP*), no polynomial-time classical (respectively, quantum) algorithm can approximate the partition function within a multiplicative factor* $\exp(c(\varepsilon)n)$.

Theorem 2 is a straightforward consequence of [16], which establishes NP-hardness in the classical limit when $\mathbf{h}^x = 0$. Specifically, [16] shows that, when $\mathbf{h}^x = 0$, for any $\varepsilon > 0$ and $\beta > \frac{1+\varepsilon}{\Delta(J)}$, approximating the partition function within an exponential multiplicative factor $\exp(c(\varepsilon)\, n)$ is NP-hard. In the regime considered in Theorem 2, where the transverse field is dominated by the interaction matrix, the partition function $Z = \mathrm{tr}(\exp(-\beta H))$ is very close to $\mathrm{tr}(\exp(-\beta H_z))$, which is the partition function of a classical Ising model and is therefore NP-hard to approximate in the low-temperature regime $\beta > \frac{1}{\Delta(J)}$ due to [16]; hence, in the same temperature regime, approximating the partition function Z is also NP-hard.

▶ Remark 3. Theorem 2 implies that approximating the partition function is NP-hard when $\beta > \frac{1}{\Delta(J)}$ and the transverse field is dominated by the interaction matrix, i.e., $\|\mathbf{h}^x\|_1 \leq n \cdot O_\varepsilon(1)\Delta(J)$. In the special case of a uniform transverse field, i.e., $h_i^x = h \,\forall i$, this condition simplifies to $|h| \leq O_\varepsilon(1)\Delta(J)$. On the other hand, when the transverse field is very strong and dominates the interaction matrix, the problem should become easy even at low temperatures, since the system effectively behaves like non-interacting spins. Understanding exactly how the strength of the transverse field affects the temperature at which this transition to an easy regime occurs remains an intriguing open problem.

1.1 Techniques

As in previous works [5, 11], we map the partition function of the transverse-field Ising model (TFIM)

$$Z := \operatorname{tr}(\exp(-\beta H))$$

to the partition function of a classical Ising model using the Trotter decomposition.

For intuition, we sketch this mapping here. Write $H_z := \frac{1}{2}\sum_{i,j\in[n]} J_{ij} Z_i Z_j$ and $H_x := \sum_{i=1}^n h_i^x X_i$; then [5] shows that for $M = \operatorname{poly}(\|\beta H_z\|, \|\beta H_x\|, \frac{1}{\varepsilon})$ a sufficiently large natural number, we have

$$\begin{aligned} Z &\approx \operatorname{tr}((e^{\frac{-\beta H_z}{M}} e^{\frac{-\beta H_x}{M}})^M) \\ &= \sum_{\mathbf{s}^{(1)},\cdots,\mathbf{s}^{(M)}} \prod_{k=1}^{M} \langle \mathbf{s}^{(k)}| e^{\frac{-\beta H_z}{M}} |\mathbf{s}^{(k)}\rangle \langle \mathbf{s}^{(k)}| e^{\frac{-\beta H_x}{M}} |\mathbf{s}^{(k+1)}\rangle \\ &= A \cdot \sum_{(\mathbf{s}^{(i)})_{i\in[M]}} \exp\left(\frac{-\beta}{2M}\sum_{k=1}^{M}\sum_{i,j=1}^{n} J_{ij} s_{k,i} s_{k,j} - \frac{\beta}{M}\sum_{k=1}^{M}\sum_{i=1}^{n} h_i^z s_{k,i} + \sum_{i=1}^{n}\sum_{k=1}^{M} K_i s_{k,i} s_{k+1,i}\right) \end{aligned} \tag{2}$$

with

$$A = \prod_{i:h_i^x\neq 0} \prod_{i:h_i^x\neq 0} \left(\frac{1}{2}\sqrt{|e^{2h_i^x/M} - e^{-2h_i^x/M}|}\right)^M$$

and $K_i = g(\frac{-\beta h_i^x}{M})$ where $g(a) = \frac{1}{2}\ln|\frac{e^a+e^{-a}}{e^a-e^{-a}}|$. Here we are using the conventions $\mathbf{s}^{(i)} = (s_{k,i})_{k\in[M]}$ and $\mathbf{s}^{(M+1)} = \mathbf{s}^{(1)}$.

Equation (2) says that, up to the constant factor A, the partition function of the transverse field Ising model is approximately equivalent to the partition function of a classical Ising model defined on nM vertices. This classical model describes a distribution over configurations $(s_{k,i})_{k\in[M],i\in[n]} \in \{\pm 1\}^{nM}$, in which the external field acting on $s_{k,i}$ is $-\beta h_i^z/M$ and the interactions are as follows:

- For any $k \in [M]$ and $i, j \in [n]$, the interaction between $s_{k,i}$ and $s_{k,j}$ is given by $-\beta J_{ij}/M$.
- For any $i \in [n]$ and $k \in [M]$, the interaction between $s_{k,i}$ and $s_{k+1,i}$ is $K_i = g\left(-\frac{\beta h_i^x}{M}\right)$ as defined above.

The following informal statement quantifies the approximation in Equation (2).

▶ **Lemma 4** (Informal version of Lemma 9). *Consider the TFIM model in Equation* (1) *and an inverse temperature* $\beta \geq 0$. *Let* $Z = \mathrm{tr}(\exp(-\beta H))$. *For any* $\varepsilon > 0$, *there exists* $M = \mathrm{poly}(\|\beta H_z\|, \|\beta H_x\|, \frac{1}{\varepsilon})$ *and* $H_{classical} : \{\pm 1\}^{nM} \to \mathbb{R} \cup \{\pm\infty\}$ *corresponding to a classical Ising model so that for* $Z' = \sum_\sigma \exp(H_{classical}(\sigma))$, *we have*

$$\exp(-\varepsilon)Z' \leq Z \leq \exp(\varepsilon)Z'.$$

Thus, we have reduced the problem of computing the partition function of the transverse field Ising model to that of a classical Ising model on nM vertices, where M is polynomial in the model and accuracy parameters. We can now estimate the partition function of the resulting classical Ising model using classical Markov chain Monte Carlo methods together with counting-to-sampling techniques. Using simulated annealing [13, 30], we reduce the problem of approximating the partition function to sampling from the corresponding Gibbs distribution. Since the classical Ising model arising from the mapping contains interaction terms that depend on the transverse field and may be unbounded, we cannot directly apply existing rapid mixing results based on the boundedness of the interaction matrix [18, 32, 23, 14, 1]. Instead, we view the corresponding classical Ising model as a distribution on $(\mathbb{R}^M)^n$, and invoke a recent result of Anari *et al.* [3] to show that the associated Glauber dynamics is rapidly mixing whenever $\beta < 1/\Delta(J)$. Each update of the Glauber dynamics on the $(\mathbb{R}^M)^n$ model reduces to sampling from a one-dimensional Ising model on M vertices, which can be efficiently implemented using dynamic programming. Combining these ingredients gives an algorithm to approximate the partition function when $\beta < 1/\Delta(J)$; we can push the boundary to $\beta = 1/\Delta(J)$ using a standard perturbative argument. First, we observe that for $\beta < 1/\Delta(J)$, the approximation algorithm's runtime depends polynomially on the gap $\frac{1}{1-\beta\Delta(J)}$. To estimate $\mathrm{tr}(\exp(-\beta H))$ at $\beta = 1/\Delta(J)$, we efficiently estimate the partition function $\mathrm{tr}(\exp(-\beta' H))$ where $\beta' = \beta(1 - \epsilon/n)$ by noting that $\frac{1}{1-\beta'\Delta(J)} = O(n/\epsilon)$, then observe that $\mathrm{tr}(\exp(-\beta' H))$ approximates $\mathrm{tr}(\exp(-\beta H))$ within an $\exp(\epsilon)$ multiplicative factor. For details, see the proof of Theorem 20, which is the formal version of Theorem 1.

1.2 Related work

The problem of approximating partition functions of classical and quantum spin systems has a long history. For many classical spin systems, the computational complexity landscape is now well understood; see, e.g., [28, 29, 10, 8, 9, 16, 14, 1, 2]. In contrast, progress on the quantum side has been more limited, and existing algorithmic results typically require strong structural or high-temperature assumptions.

Several recent works study general quantum Hamiltonians of the form $H = \sum_a \lambda_a G_a$, where $|\lambda_a| \leq 1$, each G_a is a tensor product of single-site Pauli operators acting on a subset a, and the maximum degree of the interaction graph is denoted by $\Delta = \max_i |\{a : i \in a\}|$. In this setting, Haah *et al.* [17] and Bakshi *et al.* [4] give a $(n/\varepsilon)^{O(\log \Delta)}$ -time algorithm that approximates $\mathrm{tr}(e^{-\beta H})$ within a multiplicative factor $1 + \varepsilon$, provided that $\beta < \frac{1}{4e^2(\Delta+1)^2}$. Helmuth and Mann [19] consider classical Hamiltonians perturbed by local quantum terms and obtain efficient approximation algorithms when the perturbation strength is sufficiently small and the inverse temperature lies in an appropriate regime. When specialized to the transverse-field Ising model (TFIM), however, these results impose restrictive bounds on the transverse and longitudinal fields, making them substantially weaker than ours.

Next, we discuss prior works that focus on the TFIM. Brayvi [5] gives an FPRAS for $\mathrm{tr}(e^{-\beta H})$ in the ferromagnetic (stoquastic) regime, where

$$-\beta J_{ij} \geq 0, \qquad -\beta h_i^x \geq 0, \qquad \text{and} \qquad -\beta h_i^z \geq 0.$$

In this case, the Trotter mapping gives a ferromagnetic classical Ising model, allowing the use of classical Markov chain techniques for ferromagnetic systems [20]. We emphasize that our result does not require any sign constraints on the interaction terms J_{ij} or the longitudinal and transverse fields h_i^z, h_i^x.

More recently, Crosson and Slezak [11] establish an FPRAS for the TFIM under a Dobrushin-type condition

$$\beta < \frac{1}{2(\Delta+2)B}, \quad \text{where} \quad B = \max_{i \neq j} |J_{ij}| \qquad \text{and} \qquad \Delta = \max_i |\{j \neq i : J_{ij} \neq 0\}|.$$

This condition is strictly stronger than ours, which implies an efficient approximation algorithm whenever $\beta \leq \frac{1}{2\Delta B}$. To see this, assuming without loss of generality that $J_{ii} = 0$ for all $i \in [n]$ [1], the Gershgorin circle theorem implies

$$\Delta(J) \leq 2\|J\|_\infty \leq 2\Delta B, \quad \text{where} \quad \|J\|_\infty = \max_i \sum_{j \neq i} |J_{ij}|.$$

We remark that the Dobrushin-type conditions $\beta = O(1/\|J\|_\infty)$ or $\beta < O(\frac{1}{\Delta B})$ can be substantially more restrictive than the spectral constraint $\beta < \frac{1}{\Delta(J)}$. In particular, when the interaction terms have mixed signs, significant cancellations may occur, leading to $\Delta(J) \ll \|J\|_\infty$. A concrete and important example arises when the couplings are i.i.d. Gaussian with normalized variance, $J_{ij} \sim \mathcal{N}(0, 1/n)$, as in the Sherrington–Kirkpatrick model and related spin-glass systems [27, 25]. In this case, random matrix theory implies that $\Delta(J) = \Theta(1)$, while $\|J\|_\infty = \Theta(\sqrt{n})$; hence, our Theorem 20 implies an efficient approximation algorithm when $\beta = O(1)$ while [11] requires a much higher temperature–$\beta = O(\frac{1}{\sqrt{n}})$. We can also consider diluted spin-glass models [14, 1], where the matrix J is supported on a random d-regular graph G with Rademacher disorder, i.e., $J_{ij} \sim \text{Uniform}(\{\pm 1\})$ if i, j are neighbors in G, and $J_{ij} = 0$ otherwise. For this model, Friedman's theorem implies that $\Delta(J) \leq 4\sqrt{d-1}$ (see [14, 1] for details), yet $\|J\|_\infty = d$; hence, Theorem 20 implies an efficient approximation algorithm when $\beta = O(\frac{1}{\sqrt{d}})$ while the Dobrushin-type conditions require $\beta = O(\frac{1}{d})$.

1.3 Organization

Following some background and notational preliminaries in Section 2, in Section 3 we spell out the details of how how to map the partition function of the transverse-field Ising model to that of a classical Ising model. In Section 4, we show how to sample from the resulting classical Ising model using Markov chains. In Section 5, we combine these ingredients to prove our main results Theorem 20, Theorem 1 and Theorem 2.

2 Preliminaries

2.1 Notation

For a vector $x \in \mathbb{R}^N$, we let $\|x\|$ denote its Euclidean norm. For matrix $A \in \mathbb{R}^{N \times N}$, we let $\|A\|$ denote its operator norm, i.e., $\|A\| = \|A\|_{\text{op}} = \max_{x \neq 0} \frac{\|Ax\|}{\|x\|}$. If A is Hermitian, its eigenvalues are real, and we let $\lambda_{\max}(A)$ and $\lambda_{\min}(A)$ denote its maximum and minimum eigenvalues respectively, and let $\Delta(J) = \lambda_{\max}(A) - \lambda_{\min}(J)$.

[1] Changing diagonal entries rescales the partition function by a constant factor and does not affect approximation complexity.

2.2 Linear algebra

▶ **Proposition 5.** *Let $H, H' \in \mathbb{R}^{N\times N}$ be Hermitian matrices where $\|H - H'\| \le \epsilon$. Then*

$$e^{-\epsilon}\operatorname{tr}(\exp(H)) \le \operatorname{tr}(\exp(H')) \le e^{\epsilon}\operatorname{tr}(\exp(H)).$$

Proof. This is a simple consequence of Weyl's inequality. Since H, H' are Hermitian, their eigenvalues are real. Let $\lambda_1 \ge \lambda_2 \ge \cdots \ge \lambda_N$ and $\lambda'_1 \ge \lambda'_2 \ge \cdots \ge \lambda'_N$ be the eigenvalues of H and H', respectively; then by Weyl's inequality

$$|\lambda_i - \lambda'_i| \le \|H - H'\| \le \epsilon,$$

and thus

$$\begin{aligned} e^{-\epsilon}\operatorname{tr}(\exp(H)) = e^{-\epsilon}\sum_i \exp(\lambda_i) \le \operatorname{tr}(\exp(H')) &= \sum_i \exp(\lambda'_i) \\ &\le e^{\epsilon}\sum_i \exp(\lambda_i) = e^{\epsilon}\operatorname{tr}(\exp(H)). \end{aligned}$$

◀

2.3 Markov chains and mixing time

Our analysis of the mixing time of the Glauber dynamics for the classical distribution arising from the quantum-to-classical mapping is based on establishing a fundamental information-theoretic inequality known as the *Approximate Tensorization of Entropy* (ATE). Let $\Omega = \Omega_1 \times \cdots \Omega_n$ be a product space. For $i \in [n]$, define the Markov kernel $D_{\sim i}$ as one that maps $x = (x_1, \ldots, x_n) \in \Omega$ deterministically to $x_{\sim i}$ by erasing its ith component, i.e.,

$$x_{\sim i} \coloneqq (x_1, \ldots, x_{i-1}, \varnothing, x_{i+1}, \ldots, x_n).$$

Let $D_{n\to n-1}$ be the Markov kernel that picks a coordinate $i \in [n]$ uniformly at random and erases it, i.e., returns the average of $D_{\sim 1}, \ldots, D_{\sim n}$. We call this the "down operator."

The Glauber dynamics associated with a probability distribution μ on Ω is the Markov chain that, at each step, selects a coordinate $i \in [n]$ uniformly at random, erases the current value of x_i, and resamples it according to the conditional distribution $\mu(\cdot \mid x_{\sim i})$. Equivalently, each transition consists of first applying the down operator $D_{n\to n-1}$ and then restoring the erased coordinate by sampling from the appropriate conditional marginal of μ.

▶ **Definition 6** (Approximate Tensorization of Entropy, [22, 7])**.** *Let μ be a probability measure on $\mathbb{R}^n$. We say μ satisfies* Approximate Tensorization of Entropy (ATE) with constant C *if for every other measure ν absolutely continuous w.r.t. μ,*

$$\mathcal{D}_{KL}(\nu D_{n\to n-1} || \mu D_{n\to n-1}) \le \left(1 - \frac{1}{Cn}\right)\mathcal{D}_{KL}(\nu||\mu).$$

ATE is a very strong property, which yields an array of useful consequences "for free" once established. In particular, it yields functional inequalities and implies fast mixing of the Glauber dynamics.

▶ **Proposition 7** (Fact 3.5 of [10])**.** *If ATE holds with constant C for measure μ, then:*

1. *The MLSI (Modified Log-Sobolev Inequality) for the Glauber dynamics holds with constant $1/Cn$.*

2. *As a consequence of the MLSI, on a discrete space, defining*

$$\mu_{\min} = \min\{\,\mu(x) \mid x \in \mathit{supp}(\mu)\,\},$$

the Glauber dynamics on μ has ϵ-mixing time $O(Cn(\log\log(1/\mu_{\min}) + \log(1/\epsilon)))$. Equivalently, if $\hat{\mu}$ denotes the distribution obtained after

$$t = O\big(Cn\big(\log\log(1/\mu_{\min}) + \log(1/\epsilon)\big)\big)$$

steps of the Glauber dynamics starting from an arbitrary initial state, then $d_{\mathrm{TV}}(\hat{\mu}, \mu) \le \epsilon$.

3 Mapping the transverse-field Ising model to a classical Ising model

In this section, we reproduce the standard mapping from the partition function of the transverse-field Ising model to that of a classical Ising model using the Trotter decomposition, as was sketched in Section 1.1. While this mapping is standard (see, e.g., [11, 5]), we include it here to keep our discussion self-contained.

We will use the following fact from [5].

▶ **Proposition 8** ([5, Lemma 3]). *Let $Z = \mathrm{tr}(\exp(H_z + H_x))$, $Z' = \mathrm{tr}((e^{\frac{H_z}{M}} e^{\frac{H_x}{M}})^M)$ where $M \in \mathbb{N}, M \ge 3$, and*

$$H_z = \frac{1}{2}\sum_{i,j\in[n]} J_{ij} Z_i Z_j + \sum_{i\in[n]} h_i^z Z_i\,; \qquad H_x = \sum_{i\in[n]} h_i^x X_i. \tag{3}$$

If $M \ge \max\left\{2(\|H_z\| + \|H_x\|), \sqrt{\frac{12(\|H_z\|+\|H_x\|)^3}{\epsilon}}\right\}$ then

$$e^{-\epsilon} Z \le Z' \le e^{\epsilon} Z.$$

Next, we record the standard argument that Z' corresponds to the partition function of a classical Ising model. Since the proof is standard, we defer it to the appendix.

▶ **Lemma 9** ([11, 5]). *Let $Z' = \mathrm{tr}((e^{\frac{H_z}{M}} e^{\frac{H_x}{M}})^M)$ where H_z and H_x are defined as in (3). Let $D_M = \left\{\frac{1}{\sqrt{M}}\boldsymbol{x} \mid \boldsymbol{x} \in \{\pm 1\}^M\right\} \subseteq S^{M-1} = \{\,\boldsymbol{x} \in \mathbb{R}^M \mid \|\boldsymbol{x}\| = 1\,\}$. For $i \in [n]$,*

- *If $h_i^x = 0$: let $\boldsymbol{v} = \frac{1}{\sqrt{M}}\begin{bmatrix}1 & 1 & \cdots & 1\end{bmatrix}^\intercal$ and $\Omega_i = \{\boldsymbol{v}, -\boldsymbol{v}\} \subseteq D_M$ and let $w^{(i)} : \Omega_i \to \mathbb{R}$ be defined by $w^{(i)}(\boldsymbol{v}) = h_i^z$ and $w^{(i)}(-\boldsymbol{v}) = -h_i^z$.*
- *If $h_i^x \ne 0$: Let $\Omega_i = D_M$ and $w^{(i)} : \Omega_i \to \mathbb{R}$ be defined by, for $\boldsymbol{x} = \begin{bmatrix}x_1 & x_2 & \cdots & x_M\end{bmatrix}^\intercal$,*

$$w^{(i)}(\boldsymbol{x}) = \frac{h_i^z}{\sqrt{M}}\sum_{k=1}^{M} x_k + K_i M(x_1x_2 + x_2x_3 + \cdots + x_Mx_1),$$

where

$$K_i = \ln\sqrt{\frac{e^{h_i^x/M} + e^{-h_i^x/M}}{|e^{h_i^x/M} - e^{-h_i^x/M}|}} = \frac{1}{2}\ln\left(\frac{e^{h_i^x/M} + e^{-h_i^x/M}}{|e^{h_i^x/M} - e^{-h_i^x/M}|}\right).$$

Let $H_{classical} : \Omega_1 \times \Omega_2 \times \cdots \times \Omega_n \to \mathbb{R}_{\ge 0}$ be defined by

$$H_{classical}((\boldsymbol{x}_i)_{i\in[n]}) = \frac{1}{2}\sum_{i,j\in[n]} J_{ij}\langle \boldsymbol{x}_i, \boldsymbol{x}_j\rangle + \sum_{i=1}^{n} w^{(i)}(\boldsymbol{x}_i).$$

Then

$$Z' = A \cdot \sum_{\mathbf{x}_1 \in \Omega_1, \cdots, \mathbf{x}_n \in \Omega_n} \exp(H_{classical}((\mathbf{x}_i)_{i\in[n]})),$$

where

$$A = \prod_{i:h_i^x \neq 0} \left(\frac{1}{2} \sqrt{|e^{2h_i^x/M} - e^{-2h_i^x/M}|} \right)^M .$$

4 Sampling from the classical Ising model

In this section, we will record the ingredients for establishing an efficient algorithm to sample from the classical Ising model that arises from the quantum-to-classical mapping described in Lemma 9 in the regime $\beta\Delta J < 1$. In Section 5, we combine this with the standard counting-to-sampling argument implies that the partition function can be efficiently approximated in the same regime. After appropriate transformations, the classical Ising model can be reformulated as the $O(M)$-spin model in (4). Lemma 10 establishes *Approximate Tensorization of Entropy* (ATE), which implies fast mixing of the Glauber dynamics by Proposition 7, and each step of the Glauber dynamics can be efficiently implemented by Proposition 12.

▶ **Lemma 10.** *Consider the distribution $\nu_\beta : (S^{M-1})^n \to \mathbb{R}_{\geq 0}$ defined by*

$$\nu_\beta((\boldsymbol{x}_i)_{i\in[n]}) \propto \exp\left(\frac{\beta}{2} \sum_{i,j\in[n]} J_{ij} \langle \boldsymbol{x}_i, \boldsymbol{x}_j \rangle \right) \prod_{i=1}^n \mu^{(i)}(\boldsymbol{x}_i), \tag{4}$$

where for each $i \in [n]$, $\mu^{(i)} : S^{M-1} \to \mathbb{R}_{\geq 0}$ is an arbitrary density function on the sphere $S^{M-1} = \{ \boldsymbol{x} \in \mathbb{R}^M : \|\boldsymbol{x}\| = 1 \}$. Suppose $J \succeq 0$ and $\beta \in [0, \frac{1}{\|J\|_{OP}})$. Then the distribution ν_β satisfies ATE with constant $C = \frac{1}{1-\beta\cdot\|J\|_{OP}}$.

Following [15], we say that a measure $\mu : \mathbb{R}^k \to \mathbb{R}_{\geq 0}$ is *β-semi-log-concave* if $\mathrm{Cov}(\mathcal{T}_{\mathbf{w}}\mu) \preceq \beta I$ for every $\mathbf{w} \in \mathbb{R}^k$, where the measure $\mathcal{T}_{\mathbf{w}}\mu : \mathbb{R}^k \to \mathbb{R}_{\geq 0}$ is defined by $\mathcal{T}_{\mathbf{w}}\mu(\mathbf{x}) \propto \exp(\langle \mathbf{w}, \mathbf{x} \rangle)\mu(\mathbf{x})$. We used the following proposition in the proof of Lemma 10.

▶ **Proposition 11.** *Consider a distribution $\mu : \mathbb{R}^k \to \mathbb{R}_{\geq 0}$ with*

$$\mathrm{supp}(\mu) \subseteq \{\boldsymbol{x} \in \mathbb{R}^k : \|\boldsymbol{x}\| \leq 1\}.$$

Then μ is 1-semi-log-concave.

Proof. Let

$$\Sigma := \mathrm{Cov}(\mu) = \mathbb{E}_{X\sim\mu}\left[(X - \mathbb{E}X)(X - \mathbb{E}X)^\top\right].$$

We first show that the operator norm of Σ is at most 1.

For any unit vector $\mathbf{v} \in \mathbb{R}^k$, we have

$$\begin{aligned}
\mathbf{v}^\top \Sigma \mathbf{v} &= \mathbb{E}_{X\sim\mu}\left[(\mathbf{v}^\top X - \mathbb{E}[\mathbf{v}^\top X])^2\right] \\
&= \mathrm{Var}_{X\sim\mu}(\mathbf{v}^\top X) \\
&\leq \mathbb{E}_{X\sim\mu}\left[(\mathbf{v}^\top X)^2\right] \\
&\leq \mathbb{E}_{X\sim\mu}\left[(\|\mathbf{v}\|\,\|X\|)^2\right] \\
&= \mathbb{E}_{X\sim\mu}\left[\|X\|^2\right] \\
&\leq 1,
\end{aligned}$$

where the last inequality follows from the assumption that $\mathrm{supp}(\mu) \subseteq \{\mathbf{x} : \|\mathbf{x}\| \leq 1\}$. Since this holds for all unit vectors v, we conclude that $\|\Sigma\|_{\mathrm{op}} \leq 1$.

Now let $\mathbf{w} \in \mathbb{R}^k$ and consider the tilted measure $\mathcal{T}_{\mathbf{w}}\mu$. Tilting does not change the support of the distribution, so

$$\operatorname{supp}(\mathcal{T}_{\mathbf{w}}\mu) \subseteq \{\mathbf{x} \in \mathbb{R}^k : \|\mathbf{x}\| \leq 1\}.$$

Applying the same argument as above shows that

$$\|\operatorname{Cov}(\mathcal{T}_{\mathbf{w}}\mu)\|_{\mathrm{op}} \leq 1.$$

Therefore, μ is 1-semi-log-concave. ◀

Proof of Lemma 10. This is essentially due to [3, Theorem 9], which assumes that the diagonal of J is constant, i.e., $J_{ii} = \alpha \ \forall i \in [n]$, which is not necessarily true in our case. However, by following that proof, we see that this assumption can be relaxed to $J_{ii} \geq \alpha \ \forall i \in [n]$. To apply their result, we only need to verify that, for each $i \in [n]$ and $s \geq 0$, the distribution $\mathcal{G}_{-s}\mu^{(i)}$ defined by $\mathcal{G}_{-s}\mu^{(i)} \propto \exp(s\,\|\mathbf{x}\|^2)\mu^{(i)}(\mathbf{x})$ is 1-semi-log-concave. Since $\mu^{(i)}$ is supported on S^{M-1}, the distribution $\mathcal{G}_{-s}\mu^{(i)}$ is simply $\mu^{(i)}$ itself, and is 1-semi-log-concave by Proposition 11. We may now apply [3, Theorem 9] with $\rho(s) = 1 \, \forall s \geq 0$ and $\alpha = 0, \eta = 0$, and deduce that the distribution ν_β satisfies Approximate Tensorization of Entropy with constant

$$C = \exp\Big(\int_0^{\|\beta J\|_{\mathrm{OP}}} q(z)dz\Big),$$

where $q \equiv q_{\eta,\rho}$ solves the equation

$$q(z) = 1 + \int_0^z q(s)^2 ds.$$

It is easy to see that the function $q(z) = \frac{1}{1-z}$ solves this equation, and hence ν_β satisfies ATE with constant $C = \frac{1}{1-\|\beta J\|_{\mathrm{OP}}}$. The lemma follows. ◀

▶ **Proposition 12.** *Let $H_{classical} : \Omega_1 \times \Omega_2 \times \cdots \times \Omega_n \to \mathbb{R}$ be as in Lemma 9. For $\beta \in \mathbb{R}$, let $\nu_\beta : \Omega_1 \times \Omega_2 \times \cdots \times \Omega_n \to \mathbb{R}_{\geq 0}$ be the distribution defined by $\nu_\beta(\sigma) \propto \exp(\beta H_{classical}(\sigma))$. Each step of the Glauber dynamics on ν_β samples from a 1-dimensional Ising model, and thus can be done in $O(M^2 + Mn)$ time.*

Proof. Each step of Glauber samples from the conditional distribution $\nu(\mathbf{x}_i|(\mathbf{x}_j)_{j\neq i}) : \Omega_i \to \mathbb{R}_{\geq 0}$. Let $\mathbf{x}_i = \frac{1}{\sqrt{M}}(s_{1,i}, \cdots, s_{M,i}) \in D_M$. Sampling from $\nu(\mathbf{x}_i|(\mathbf{x}_j)_{j\neq i})$ reduces to sampling from the Ising model $\pi : \{\pm 1\}^M \to \mathbb{R}_{\geq 0}$ where

$$\pi((s_{1,i}, \cdots, s_{M,i})) \propto (\sum_{k=1}^{M} K_i s_{k,i} s_{k+1,i} + \sum_{i=1}^{M} \hat{h}_k s_{k,i})$$

and

$$\hat{h}_k = \frac{1}{M}\sum_{j\neq i} J_{ij} s_{k,j} + \frac{1}{M} h_i^z s_{k,i}.$$

Computing $(\hat{h}_k)_{k\in[M]}$ takes $O(Mn)$ time in total, and sampling takes $O(M^2)$ time, as we prove in Proposition 13 below. ◀

▶ **Proposition 13.** *Consider the 1D Ising model* $\nu : \{\pm 1\}^M \to \mathbb{R}_{\geq 0}$ *with either the open boundary condition*

$$\nu(x) \propto \exp\left(\sum_{i=1}^{M-1} J_{i,i+1} x_i x_{i+1} + \sum_{i=1}^{M} h_i x_i\right), \tag{5}$$

or the closed boundary condition

$$\nu(x) \propto \exp\left(\sum_{i=1}^{M} J_{i,i+1} x_i x_{i+1} + \sum_{i=1}^{M} h_i x_i\right), \qquad x_{M+1} = x_1. \tag{6}$$

In both cases, one can compute all relevant conditional marginals exactly and therefore sample from ν *in* $O(M^2)$ *time.*

Proof. We show how to compute the conditional marginals

$$\nu(X_i = \pm 1 \mid X_{[i-1]} = \sigma_{[i-1]})$$

exactly by dynamic programming, for any $i \in [M]$ and any partial assignment $\sigma_{[i-1]}$ with $\nu(x_{[i-1]} = \sigma_{[i-1]}) > 0$. Once these marginals are available, we may sample sequentially using the standard autoregressive procedure: sample $\sigma_1 \sim \nu(X_1)$, then sample $\sigma_2 \sim \nu(X_2 \mid X_1 = \sigma_1)$, and so on, until we have sampled $\sigma_n \sim \nu(X_n \mid X_{[n-1]} = \sigma_{[n-1]})$, at which point we output $\sigma = \sigma_{[n]}$. The output distribution is exactly ν.

Open boundary case (5). It suffices to compute $\nu(x_1)$, since for $i > 1$ the conditional distribution $\nu(x_i \mid x_{[i-1]} = \sigma_{[i-1]})$ is itself the open-boundary Ising model on a smaller interval.

Define

$$p_N(\sigma_M) = \exp(h_M \sigma_M), \qquad \sigma_M \in \{\pm 1\}.$$

For $k = N-1, \ldots, 1$, compute recursively

$$p_k(\sigma_k) = \sum_{\sigma_{k+1} \in \{\pm 1\}} \exp\left(J_{k,k+1} \sigma_k \sigma_{k+1} + h_k \sigma_k\right) p_{k+1}(\sigma_{k+1}).$$

Then

$$\nu(x_1 = \sigma_1) = \frac{p_1(\sigma_1)}{p_1(1) + p_1(-1)}, \qquad \sigma_1 \in \{\pm 1\}.$$

Closed boundary case (6). When $i > 1$, conditioning on $x_{[i-1]} = \sigma_{[i-1]}$ breaks the cycle and reduces the problem to the open-boundary case.

For the case $i = 1$, define

$$p_M^{(\sigma_1)}(\sigma_M) = \exp\left(J_{N,1}\, \sigma_1 \sigma_M + h_M \sigma_M\right), \qquad (\sigma_1, \sigma_M) \in \{\pm 1\}^2.$$

For $k = M-1, \ldots, 1$, compute

$$p_k^{(\sigma_1)}(\sigma_k) = \sum_{\sigma_{k+1} \in \{\pm 1\}} \exp\left(J_{k,k+1} \sigma_k \sigma_{k+1} + h_k \sigma_k\right) p_{k+1}^{(\sigma_1)}(\sigma_{k+1}).$$

Finally, set

$$p_1(\sigma_1) = p_1^{(\sigma_1)}(\sigma_1)$$

and compute the marginal

$$\nu(x_1 = \sigma_1) = \frac{p_1(\sigma_1)}{p_1(1) + p_1(-1)}.$$

Runtime. Each marginal computation requires $O(M)$ operations. Sampling by autoregression requires computing M marginals, so the runtime is $O(M^2)$. ◀

5 On computing the partition function

In this section, we prove our main results, Theorem 1 and Theorem 2 from the Introduction. We begin with Theorem 2, which is a direct consequence of the following result from [16].

▶ **Theorem 14** ([16, Theorem 1]). *There exists a symmetric matrix $J \in \mathbb{R}^{n\times n}$ with $\Delta(J) = 1+\varepsilon > 1$ such that approximating the partition function $\operatorname{tr}(\exp(\frac{1}{2}\sum_{i,j} J_{ij}Z_iZ_j)) = \sum_{\sigma\in\{\pm1\}^n}\exp(\frac{1}{2}\sum_{i,j}J_{ij}\sigma_i\sigma_j)$ within an $\exp(c(\varepsilon)n)$-multiplicative factor is NP-hard, where the constant $c(\varepsilon) > 0$ depends on ε.*

Proof of Theorem 2. Let $Z = \operatorname{tr}(\exp(-\beta H))$. Let $\tilde{J} = -\beta J$, and $\tilde{Z} = \operatorname{tr}(\exp(-\beta H_z)) = \operatorname{tr}(\exp(\frac{1}{2}\sum_{i,j}\tilde{J}_{ij}Z_iZ_j))$. When $\|\beta\mathbf{h}^x\|_1 \le \frac{c(\varepsilon)n}{2}$, we have

$$\|-\beta H - (-\beta H_z)\| = \left\|\beta\sum_{i\in[n]} h_i^x X_i\right\| \le \|\beta\mathbf{h}^x\|_1 \le \frac{c(\varepsilon)n}{2}$$

and thus by Proposition 5,

$$Z\exp(-\frac{c(\varepsilon)n}{2}) \le \tilde{Z} \le Z\exp(\frac{c(\varepsilon)n}{2}).$$

Thus, if we can output a $\hat{Z}$ that approximates Z within a multiplicative factor $\exp(\frac{c(\varepsilon)n}{2})$, then $\hat{Z}$ will approximate $\tilde{Z}$ within a factor of $\exp(c(\varepsilon)n)$. The claim then follows from Theorem 14. ◀

The remainder of this section is devoted to the proof of Theorem 20, which we will restate more formally later as Theorem 1. We first recall some classical results on counting-to-sampling via simulated annealing.

▶ **Lemma 15.** *Consider a discrete set Ω and a Hamiltonian $H : \Omega \to [-L, L]$ for some $L \ge 0$. For $\beta \in \mathbb{R}$, let μ_β be the distribution on Ω defined by*

$$\mu_\beta(\sigma) = \frac{\exp(-\beta H(\sigma))}{Z(\beta)}, \quad \text{where} \quad Z(\beta) = \sum_{\sigma\in\Omega}\exp(-\beta H(\sigma)).$$

Suppose we know $Z(\beta_{\min})$, and we have oracle access to the Hamiltonian H and to samples from μ_β for $\beta \in [\beta_{\min}, \beta_{\max}]$. For any $\epsilon > 0$, there exists a (classical) algorithm that outputs an estimator $\hat{Z}$ s.t.

$$\Pr\Big((1-\epsilon)Z(\beta_{\max}) \le \hat{Z} \le (1+\epsilon)Z(\beta_{\max})\Big) \ge 3/4.$$

The algorithm uses $k = O\left(\left(\frac{(\beta_{\max}-\beta_{\min})L}{\epsilon}\right)^2\right)$ samples from μ_β and $O(k)$ Hamiltonian evaluations.

Using a standard coupling argument, we can also relax the precondition of Lemma 15 to allow for samples from a distribution that is close to μ_β.

▶ **Lemma 16.** *Consider a discrete set Ω and a Hamiltonian $H : \Omega \to [-L, L]$ for some $L \ge 0$. For $\beta \in \mathbb{R}$, let μ_β be the distribution on Ω defined by*

$$\mu_\beta(\sigma) = \frac{\exp(-\beta H(\sigma))}{Z(\beta)} \quad \text{where} \quad Z(\beta) = \sum_{\sigma\in\Omega}\exp(-\beta H(\sigma)).$$

Suppose we know $Z(\beta_{\min})$, and we have oracle access to the Hamiltonian H and to samples from $\hat{\mu}_\beta$ where $d_{TV}(\hat{\mu}_\beta, \mu_\beta) \leq \epsilon_{TV}$ for $\beta \in [\beta_{\min}, \beta_{\max}]$. For any $\epsilon > 0$, there exists a (classical) algorithm that outputs an estimator $\hat{Z}$ s.t.

$$\Pr\left((1-\epsilon)Z(\beta_{\max}) \leq \hat{Z} \leq (1+\epsilon)Z(\beta_{\max})\right) \geq 3/4 - \epsilon_{TV} k.$$

The algorithm uses $k = O\left(\left(\frac{(\beta_{\max}-\beta_{\min})L}{\epsilon}\right)^2\right)$ samples from $\hat{\mu}_\beta$ and $O(k)$ Hamiltonian evaluations.

Lemmas 15 and 16 are standard, and we leave their proofs to the appendix.

▶ Remark 17. A standard application of Chernoff bounds allows us to improve the success probability in Lemmas 15 and 16 to $1-\delta$ by taking the median of $\Theta(\log(1/\delta))$ independent trials.

In order to apply Lemmas 15 and 16, we need the following proposition which bounds the classical Hamiltonian arising from the quantum-to-classical mapping in Lemma 9, whose proof we leave to the appendix.

▶ **Proposition 18.** *$\forall a \in [-1,1] \setminus \{0\}$, we have $0 \leq \ln\left(\frac{e^a+e^{-a}}{|e^a-e^{-a}|}\right) \leq \ln\frac{2}{|a|}$.*

▶ **Proposition 19.** *Let $H_{classical} : \Omega_1 \times \Omega_2 \times \cdots \times \Omega_n \to \mathbb{R}_{\geq 0}$ be as in Lemma 9. Suppose $M \geq |h_i^x|\ \forall i \in [n]$ and $\min\{\ |h_i^x|\ |\ h_i^x \neq 0\} \geq r$. Let*

$$L = n\|J\| + \|\mathbf{h}^z\|_1 + nM\ln(2M/r)).$$

Then

$$|H((\boldsymbol{x}_i)_{i\in[n]})| \leq L \quad \forall (\boldsymbol{x}_i)_{i\in[n]} \in \Omega_1 \times \Omega_2 \times \cdots \times \Omega_n.$$

Proof. Let $H \equiv H_{\text{classical}}$. By the triangle inequality,

$$|H((\mathbf{x}_i)_{i\in[n]})| \leq |\sum_{i,j\in[n]} J_{ij}\langle \mathbf{x}_i, \mathbf{x}_j\rangle| + \sum_{i\in[n]} |w^{(i)}(\mathbf{x}_i)|. \tag{7}$$

Recall that $\mathbf{x}_i = \frac{1}{\sqrt{M}}(s_{1,i}, \cdots, s_{M,i}) \in D_M$; therefore

$$|\sum_{i,j\in[n]} J_{ij}\langle \mathbf{x}_i, \mathbf{x}_j\rangle| = |\frac{1}{M}\sum_{k=1}^{M}\sum_{i,j\in[n]} J_{ij}s_{k,i}s_{k,j}| \leq \frac{1}{M}\sum_{k=1}^{M}|\sum_{i,j\in[n]} J_{ij}s_{k,i}s_{k,j}| \leq n\|J\|.$$

For $i \in [n]$ and $\mathbf{x} \in \Omega_i$, we will show

$$|w^{(i)}(x)| \leq |h_i^z| + M\ln(2M/r).$$

If $h_i^x = 0$ then $|w^{(i)}(\mathbf{x})| = |h_i^z|$ and we are done. So we assume $h_i^x \neq 0$, and then for any $\mathbf{x} = \begin{bmatrix} x_1 & x_2 & \cdots & x_M \end{bmatrix}^\intercal$, by the triangle inequality

$$\begin{aligned} |w^{(i)}(\mathbf{x})| &\leq \frac{|h_i^z|}{\sqrt{M}}\sum_{k=1}^{M}|x_k| + |K_i|M(|x_1x_2| + |x_2x_3| + \cdots + |x_Mx_1|) \\ &= |h_i^z| + |K_i|M, \end{aligned}$$

where we recall

$$K_i = \ln\sqrt{\frac{e^{h_i^x/M}+e^{-h_i^x/M}}{|e^{h_i^x/M}-e^{-h_i^x/M}|}} = \frac{1}{2}\ln\left(\frac{e^{h_i^x/M}+e^{-h_i^x/M}}{|e^{h_i^x/M}-e^{-h_i^x/M}|}\right)$$

satisfies $0 \leq K_i \leq \ln\frac{2M}{|h_i^x|} \leq \ln(2M/r)$ by Proposition 18. Plugging this back into Equation (7) finishes the proof. ◀

Finally, we are ready to prove Theorem 1. In fact, we will prove a more formal version of the theorem, which we now state.

▶ **Theorem 20** (Formal version of Theorem 1). *Consider the TFIM model in Equation* (1). *We assume without loss of generality that* $J = (J_{ij})_{i,j\in[n]} \in \mathbb{R}^{n\times n}$ *is a symmetric matrix. Let* $\lambda_{\max}(J)$ *and* $\lambda_{\min}(J)$ *denote its maximum and minimum eigenvalues, and let* $\Delta(J) = \lambda_{\max}(J) - \lambda_{\min}(J)$. *Let* $Z := \operatorname{tr}(\exp(-\beta H))$.

Suppose $\beta \in [0, \frac{1}{\Delta(J)})$. *Let* $\|\boldsymbol{h}^z\|_1 = \sum_{i=1}^n |h_i^z|$ *and* $\|\boldsymbol{h}^x\|_1 = \sum_{i=1}^n |h_i^x|$. *For any* $\epsilon, \delta \in (0,1)$, *there exists a classical algorithm that runs in*

$$\frac{1}{1-\beta\Delta(J)} \cdot \log \delta^{-1} \cdot \operatorname{poly}(n, \beta \|\boldsymbol{h}^x\|_1, \beta \|\boldsymbol{h}^z\|_1, \epsilon^{-1})$$

time and outputs $\hat{Z}$ *such that,*

$$\Pr\left(e^{-\epsilon}\hat{Z} \le Z \le e^{\epsilon}\hat{Z}\right) \ge 1-\delta.$$

Moreover, for the critical value $\beta = \frac{1}{\Delta(J)}$, *the same approximation guarantee holds with runtime* $\log \delta^{-1} \cdot \operatorname{poly}(n, \beta \|\boldsymbol{h}^x\|_1, \beta \|\boldsymbol{h}^z\|_1, \epsilon^{-1})$.

Proof. We begin with a proof of the first statement; the second statement will follow relatively straightforwardly from the first.

Letting $\tilde{J} = -\beta J + \beta\lambda_{\max}(J)\cdot I$, we have $0 \preceq \tilde{J} \preceq (\beta\Delta(J))\cdot I \preceq I$. Now let $\tilde{\mathbf{h}}^z = -\beta\mathbf{h}^z, \tilde{\mathbf{h}}^x = -\beta\mathbf{h}^x$, and

$$H_z = \frac{1}{2}\sum_{i,j\in[n]} \tilde{J}_{ij} Z_i Z_j + \sum_{i\in[n]} \tilde{h}_i^z Z_i\,, \qquad H_x = \sum_{i\in[n]} \tilde{h}_i^x X_i.$$

This implies that $\operatorname{tr}(\exp(-\beta H))/\operatorname{tr}(\exp(H_z+H_x)) = c$ for some constant $c > 0$ that we can efficiently compute, so we only need to approximate $Z := \operatorname{tr}(\exp(H_z+H_x))$.

Note that, since H_z is a diagonal matrix with diagonal entries given by $\frac{1}{2}\sum_{i,j\in[n]} \tilde{J}_{ij} z_i z_j + \sum_{i\in[n]} \tilde{h}_i^z z_i$ for $z \in \{\pm 1\}^n$, we obtain the following bounds:

$$\|H_z\| \le \max_{z\in\{\pm1\}^n} |\frac{1}{2}\sum_{i,j\in[n]} \tilde{J}_{ij} z_i z_j + \sum_{i\in[n]} \tilde{h}_i^z z_i| \le n\left\|\tilde{J}\right\| + \left\|\tilde{\mathbf{h}}^z\right\|_1 \le (\beta\Delta(J))n + \left\|\tilde{\mathbf{h}}^z\right\|_1;$$
$$\|H_x\| \le \sum_{i\in[n]} |\tilde{h}_i^x|\cdot\|X_i\| = \sum_{i\in[n]} |\tilde{h}_i^x| = \left\|\tilde{\mathbf{h}}^x\right\|_1.$$

Letting

$$C := (\beta\Delta(J))n + \left\|\tilde{\mathbf{h}}^x\right\|_1 + \left\|\tilde{\mathbf{h}}^z\right\|_1$$

we have $C \ge \|H_z\| + \|H_x\|$. Let $r > 0$ and $M \in \mathbb{N}_{\ge 3}$ be parameters to be chosen later. Letting

$$I_{\text{bad}} = \{\, i \in [n] : \tilde{h}_i^x \ne 0 \text{ and } |\tilde{h}_i^x| \le r \,\},$$

and

$$\tilde{H} = H_z + \tilde{H}_x, \quad \tilde{H}_x = \sum_{i\in[n]\setminus I_{\text{bad}}} h_i^x X_i,$$

we get

$$\left\|H - \tilde{H}\right\| \leq \sum_{i \in I_{\text{bad}}} |\tilde{h}_i^x| \cdot \|X_i\| \leq rn.$$

Hence by Proposition 5, setting $Z = \text{tr}(\exp(H_z + H_x))$ and $\tilde{Z} = \text{tr}(\exp(H_z + \tilde{H}_x))$ and choosing $r = \frac{\epsilon}{3n}$, we have

$$e^{-\epsilon/3}\tilde{Z} \leq Z \leq e^{\epsilon/3}\tilde{Z}. \tag{8}$$

We now approximate $\tilde{Z}$. Set

$$M = \left(\frac{3C^3}{\epsilon}\right)^{1/2}. \tag{9}$$

By Proposition 8 and Lemma 9, let the classical Hamiltonian $H_{\text{classical}}$ be defined as in Lemma 9 wrt $\tilde{H}$ and M, so that, for

$$\tilde{A} = \prod_{i:\tilde{h}_i^x \neq 0} \left(\frac{1}{2}\sqrt{|e^{2\tilde{h}_i^x/M} - e^{-2\tilde{h}_i^x/M}|}\right)^M,$$

we have

$$Z' = \tilde{A} \sum_{\mathbf{x}_1 \in \Omega_1, \cdots, \mathbf{x}_n \in \Omega_n} \exp(H_{\text{classical}}((\mathbf{x}_i)_{i\in[n]}))$$

and

$$\exp(-\epsilon/3)\tilde{Z} \leq Z' \leq \tilde{Z}\exp(\epsilon/3).$$

In particular, combining this with Equation (8) we obtain

$$\exp(-2\epsilon/3)Z' \leq Z \leq Z'\exp(2\epsilon/3). \tag{10}$$

Next, we approximate Z'. Setting

$$L = C + nM\log(M/r) = \tilde{O}(nM),$$

yields

$$|H_{\text{classical}}((\mathbf{x}_i)_{i\in[n]})| \leq L \forall (\mathbf{x}_i)_{i\in[n]} \in \Omega_1 \times \cdots \times \Omega_n.$$

For $\beta' \in [0,1]$, consider the distribution defined by $\nu_{\beta'}((\mathbf{x}_i)_{i\in[n]}) \propto \exp(\beta' H_{\text{classical}}((\mathbf{x}_i)_{i\in[n]}))$. Applying Lemma 10 and Proposition 7 to this distribution, and noting that

$$(\nu_{\beta'})_{\min} := \min\{\, \nu_{\beta'}(x) \mid x \in \text{supp}(\nu_{\beta'}) \,\} \geq \exp(-2L)2^{-nM},$$

we obtain that the Glauber dynamics[2] outputs a sample from $\hat{\nu}_{\beta'}$ s.t. $d_{TV}(\hat{\nu}_{\beta'}, \nu_{\beta'}) \leq \epsilon_{TV}$ in

$$T = O\left(\frac{1}{1 - \beta\Delta(J)} \cdot n\log\left(\frac{n}{\epsilon_{TV}}\right)\text{poly}\log(C, \epsilon^{-1})\right)$$

steps, where each step takes $O(M^2)$ time by Proposition 12.

[2] We can initialize the Glauber dynamics at any $(\mathbf{x}_i)_{i\in[n]} \in \Omega_1 \times \cdots \times \Omega_n$, for example, at $(\mathbf{x}_i)_{i\in[n]}$ where $\mathbf{x}_i = \frac{1}{\sqrt{M}}\begin{bmatrix}1 & 1 & \cdots & 1\end{bmatrix}^\intercal \forall i \in [n]$.

Applying Lemma 16 and Proposition 19 for $\epsilon_{TV} = \frac{1}{8k}$, $\beta_{\min} = 0, \beta_{\max} = 1$ (note that $Z(\beta_{\min}) = \prod_{i=1}^{n} |\Omega_i|$ can be easily computed and $M \geq C \geq \max |\tilde{h}_i^x|$), we obtain a classical algorithm that outputs $\hat{Z}$ so that with probability $\geq 5/8$,

$$\exp(-\epsilon/3)Z' \leq \hat{Z} \leq Z' \exp(\epsilon/3),$$

which, by Equation (10), implies

$$\hat{Z} \exp(-\epsilon) \leq Z \leq \hat{Z} \exp(\epsilon)$$

using $k = O\left(\frac{L^2}{\epsilon^2}\right)$ samples from $\hat{\nu}_{\beta'}$ and evaluations of $H_{\text{classical}}$.

The overall running time of the algorithm is

$$O(kTM^2 + kn^2M) = \tilde{O}\Big(\frac{1}{1-\beta\Delta(J)} \cdot \epsilon^{-4} n^3 ((\beta\Delta(J))n + \beta \left\|\mathbf{h}^x\right\|_1 + \beta \left\|\mathbf{h}\right\|_1^z)^6\Big),$$

where $\tilde{O}$ hides polylogarithmic factors.

We can boost the success rate from $5/8$ to $1-\delta$ for any $\delta > 0$ at the cost of an extra $O(\log(1/\delta))$ factor in the runtime (see Remark 17). This finishes the proof of the first statement.

For the proof of the second statement, where $\beta = \frac{1}{\Delta(J)}$, let $\hat{J} = (1 - \frac{\epsilon}{2n\beta\Delta(J)})\tilde{J}$ and

$$\tilde{Z} = \operatorname{tr}(\exp(\hat{H}_z + H_x)) \quad \text{where} \quad \hat{H}_z = \frac{1}{2} \sum_{i,j \in [n]} \hat{J}_{ij} Z_i Z_j + \sum_{i \in [n]} \tilde{h}_i^z Z_i.$$

Note that

$$\left\|\hat{H}_z - H_z\right\| \leq \frac{\epsilon}{2n\beta\Delta(J)} \cdot \max_{z \in \{\pm 1\}^n} \left|\frac{1}{2} \sum_{i,j \in [n]} \tilde{J}_{ij} z_i z_j\right| \leq \frac{\epsilon}{2n\beta\Delta(J)} \cdot n \cdot \left\|\tilde{J}\right\| \leq \epsilon/2.$$

By Proposition 5,

$$\tilde{Z} \exp(-\epsilon/2) \leq Z \leq \tilde{Z} \exp(\epsilon/2).$$

Noting that $\left\|\hat{J}\right\| \leq 1 - \frac{\epsilon}{2n}$, we can now use the first statement to output $\hat{Z}$ that approximates $\tilde{Z}$ within a factor of $\exp(\epsilon/2)$, and thus approximates Z within a factor of $\exp(\epsilon)$. The runtime guarantee follows from the first statement. ◀

References

1. Nima Anari, Vishesh Jain, Frederic Koehler, Huy Tuan Pham, and Thuy-Duong Vuong. Entropic independence i: Modified log-Sobolev inequalities for fractionally log-concave distributions and high-temperature Ising models. *arXiv preprint*, 2021. arXiv:2106.04105.
2. Nima Anari, Vishesh Jain, Frederic Koehler, Huy Tuan Pham, and Thuy-Duong Vuong. Entropic independence: optimal mixing of down-up random walks. In *Proceedings of the 54th Annual ACM SIGACT Symposium on Theory of Computing*, STOC 2022, pages 1418–1430, New York, NY, USA, 2022. Association for Computing Machinery. doi:10.1145/3519935.3520048.
3. Nima Anari, Frederic Koehler, and Thuy Duong Vuong. Trickle-down in localization schemes and applications. *Proceedings of the 56th Annual ACM Symposium on Theory of Computing*, 2024. URL: https://api.semanticscholar.org/CorpusID:270392522.
4. Ainesh Bakshi, Allen Liu, Ankur Moitra, and Ewin Tang. High-temperature gibbs states are unentangled and efficiently preparable. *2024 IEEE 65th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 1027–1036, 2024. URL: https://api.semanticscholar.org/CorpusID:268680477.

5 Sergey Bravyi. Monte carlo simulation of stoquastic hamiltonians. *Quantum Inf. Comput.*, 15:1122–1140, 2014. doi:10.26421/QIC15.13-14-3.

6 Sergey Bravyi, Anirban Narayan Chowdhury, David Gosset, and Pawel Wocjan. Quantum hamiltonian complexity in thermal equilibrium. *Nature Physics*, 18:1367–1370, 2021. URL: https://api.semanticscholar.org/CorpusID:240288401.

7 Pietro Caputo, Georg Menz, and Prasad Tetali. Approximate tensorization of entropy at high temperature, 2014. doi:10.48550/arXiv.1405.0608.

8 Xiaoyu Chen, Weiming Feng, Yitong Yin, and Xinyuan Zhang. Rapid mixing of glauber dynamics via spectral independence for all degrees. *2021 IEEE 62nd Annual Symposium on Foundations of Computer Science (FOCS)*, pages 137–148, 2021. URL: https://api.semanticscholar.org/CorpusID:235254724.

9 Yuansi Chen and Ronen Eldan. Localization schemes: A framework for proving mixing bounds for markov chains (extended abstract). *2022 IEEE 63rd Annual Symposium on Foundations of Computer Science (FOCS)*, pages 110–122, 2022. doi:10.1109/FOCS54457.2022.00018.

10 Zongchen Chen, Kuikui Liu, and Eric Vigoda. Optimal mixing of glauber dynamics: Entropy factorization via high-dimensional expansion. In *Proceedings of the 53rd Annual ACM SIGACT Symposium on Theory of Computing*, pages 1537–1550, 2021. doi:10.1145/3406325.3451035.

11 Elizabeth Crosson and Samuel Slezak. Classical simulation of high temperature quantum ising models. *Quantum*, 9:1237, 2025. doi:10.22331/q-2025-02-18-1237.

12 Amit Dutta, Gabriel Aeppli, Bikas K. Chakrabarti, Uma Divakaran, Thomas F. Rosenbaum, and Diptiman Sen. *Quantum Phase Transitions in Transverse Field Spin Models: From Statistical Physics to Quantum Information*. Cambridge University Press, Cambridge, 2015. doi:10.1017/CBO9781107706057.

13 Martin Dyer, Leeds K, and Alan Frieze. Computing the volume of convex bodies: A case where randomness provably helps. *Proc Symp Appl Math*, 44, February 1998. doi:10.1090/psapm/044/1141926.

14 Ronen Eldan, Frederic Koehler, and Ofer Zeitouni. A spectral condition for spectral gap: fast mixing in high-temperature ising models. *Probability Theory and Related Fields*, 182:1035–1051, 2020. URL: https://api.semanticscholar.org/CorpusID:220546472.

15 Ronen Eldan and Omer Shamir. Log concavity and concentration of lipschitz functions on the boolean hypercube. *Journal of functional analysis*, 282(8):109392, 2022.

16 Andreas Galanis, Alkis Kalavasis, and Anthimos Vardis Kandiros. On Sampling from Ising Models with Spectral Constraints. In Amit Kumar and Noga Ron-Zewi, editors, *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques (APPROX/RANDOM 2024)*, volume 317 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 70:1–70:14, Dagstuhl, Germany, 2024. Schloss Dagstuhl – Leibniz-Zentrum für Informatik. doi:10.4230/LIPIcs.APPROX/RANDOM.2024.70.

17 Jeongwan Haah, Robin Kothari, and Ewin Tang. Learning quantum hamiltonians from high-temperature gibbs states and real-time evolutions. *Nature Physics*, 20:1027–1031, 2021. URL: https://api.semanticscholar.org/CorpusID:282058447.

18 Thomas P Hayes. A simple condition implying rapid mixing of single-site dynamics on spin systems. In *2006 47th Annual IEEE Symposium on Foundations of Computer Science (FOCS'06)*, pages 39–46. IEEE, 2006. doi:10.1109/FOCS.2006.6.

19 Tyler Helmuth and Ryan L. Mann. Efficient algorithms for approximating quantum partition functions at low temperature. *Quantum*, 7:1155, October 2023. doi:10.22331/q-2023-10-25-1155.

20 Mark Jerrum and Alistair Sinclair. Polynomial-time approximation algorithms for the ising model. *SIAM Journal on Computing*, 22(5):1087–1116, 1993. doi:10.1137/0222066.

21 Tadashi Kadowaki and Hidetoshi Nishimori. Quantum annealing in the transverse ising model. *Physical Review E*, 58(5):5355–5363, 1998. doi:10.1103/PhysRevE.58.5355.

22 Katalin Marton. An inequality for relative entropy and logarithmic sobolev inequalities in euclidean spaces. *Journal of Functional Analysis*, 264(1):34–61, 2013.

23 Katalin Marton. Logarithmic sobolev inequalities in discrete product spaces. *Combinatorics, Probability and Computing*, 28(6):919–935, 2019. doi:10.1017/S0963548319000099.

24 R. Martoňák, G. E. Santoro, and E. Tosatti. Quantum annealing by the path-integral monte carlo method: The two-dimensional random ising model. *Physical Review B*, 66(9):094203, 2002. doi:10.1103/PhysRevB.66.094203.

25 Marc Mézard, Giorgio Parisi, and Miguel Angel Virasoro. *Spin glass theory and beyond: An Introduction to the Replica Method and Its Applications*, volume 9. World Scientific Publishing Company, 1987.

26 Subir Sachdev. *Quantum Phase Transitions*. Cambridge University Press, Cambridge, 2 edition, 2011.

27 David Sherrington and Scott Kirkpatrick. Solvable model of a spin-glass. *Physical review letters*, 35(26):1792, 1975.

28 Allan Sly. Computational transition at the uniqueness threshold. In *2010 IEEE 51st Annual Symposium on Foundations of Computer Science*, pages 287–296. IEEE, 2010. doi:10.1109/FOCS.2010.34.

29 Allan Sly and Nike Sun. The computational hardness of counting in two-spin models on d-regular graphs. *2012 IEEE 53rd Annual Symposium on Foundations of Computer Science*, pages 361–369, 2012. URL: https://api.semanticscholar.org/CorpusID:1610854.

30 Daniel Stefankovic, Santosh S. Vempala, and Eric Vigoda. Adaptive simulated annealing: A near-optimal connection between sampling and counting. *48th Annual IEEE Symposium on Foundations of Computer Science (FOCS'07)*, pages 183–193, 2006. URL: https://api.semanticscholar.org/CorpusID:2713057.

31 Yuki Susa, Yusuke Yamashiro, Masayuki Yamamoto, Itay Hen, Daniel A. Lidar, and Hidetoshi Nishimori. Quantum annealing of the p-spin model under inhomogeneous transverse field driving. *Physical Review A*, 98(4):042326, 2018. doi:10.1103/PhysRevA.98.042326.

32 Liming Wu. Poincaré and transportation inequalities for gibbs measures under the dobrushin uniqueness condition. *The Annals of Probability*, 34(5):1960–1989, 2006.

A Deferred proofs

Proof of Proposition 8. Inserting resolutions of the identity in terms of the computational basis, i.e., $I = \sum_{\mathbf{s}^{(k)}} |\mathbf{s}^{(k)}\rangle \langle \mathbf{s}^{(k)}|$, we have

$$\begin{aligned} Z' &= \operatorname{tr}((\sum_{\mathbf{s}^{(1)}} |\mathbf{s}^{(1)}\rangle \langle \mathbf{s}^{(1)}|) e^{\frac{H_z}{M}} e^{\frac{H_x}{M}} (\sum_{\mathbf{s}^{(2)}} |\mathbf{s}^{(2)}\rangle \langle \mathbf{s}^{(2)}|) e^{\frac{H_z}{M}} e^{\frac{H_x}{M}} \cdots (\sum_{\mathbf{s}^{(M)}} |\mathbf{s}^{(M)}\rangle \langle \mathbf{s}^{(M)}|) e^{\frac{H_z}{M}} e^{\frac{H_x}{M}}) \\ &= \sum_{\mathbf{s}^{(1)},\cdots,\mathbf{s}^{(M)}} \prod_{k=1}^{M} \langle \mathbf{s}^{(k)}| e^{\frac{H_z}{M}} e^{\frac{H_x}{M}} |\mathbf{s}^{(k+1)}\rangle \end{aligned} \tag{11}$$

with periodic boundary condition $\mathbf{s}^{(M+1)} = \mathbf{s}^{(1)}$. Since H_z is diagonal in the computational basis, the above can be rewritten as

$$Z' = \sum_{\mathbf{s}^{(1)},\cdots,\mathbf{s}^{(M)}} \prod_{k=1}^{M} \langle \mathbf{s}^{(k)}| e^{\frac{H_z}{M}} |\mathbf{s}^{(k)}\rangle \langle \mathbf{s}^{(k)}| e^{\frac{H_x}{M}} |\mathbf{s}^{(k+1)}\rangle . \tag{12}$$

We identify the computational basis vector $\mathbf{s}^{(k)}$ with $(s_{k,1}, \cdots, s_{k,N}) \in \{\pm 1\}^n$. Then

$$\begin{aligned} \langle \mathbf{s}^{(k)}| e^{\frac{H_z}{M}} |\mathbf{s}^{(k)}\rangle &= \exp\Big(\frac{1}{M}(\frac{1}{2}\sum_{i,j} J_{ij} s_{k,i} s_{k,j} + \sum_{i=1}^{n} h_i^z s_{k,i})\Big) \\ &= \exp\Big(\frac{1}{2}\sum_{i,j\in[n]} J_{ij}\langle \mathbf{x}_i, \mathbf{x}_j\rangle + \sum_{i\in[n]} \frac{h_i^z}{\sqrt{M}} \sum_{k=1}^{M} (\mathbf{x}_i)_k\Big), \end{aligned} \tag{13}$$

where we let $\mathbf{x}_i = \frac{1}{\sqrt{M}}(s_{1,i}, s_{2,i}, \cdots, s_{M,i}) \in D_M$.

Because $H_x = \sum_i h_i^x X_i$, we have

$$e^{H_x} = \prod_i e^{h_i^x X_i}. \tag{14}$$

If $h_i^x = 0$ then $e^{h_i^x X_i} = I$ is diagonal in the computational basis, so

$$\prod_{k=1}^{M} \langle \mathbf{s}^{(k)} |\, e^{\frac{H_x}{M}} \,| \mathbf{s}^{(k+1)} \rangle \neq 0 \Rightarrow s_{k,i} = s_{k+1,i}\ \forall k \in [M] \Rightarrow \mathbf{x}_i \in \{\pm \mathbf{v}\} = \Omega_i. \tag{15}$$

Next, consider i where $h_i^x \neq 0$. We consider the power e^{aX} of the Pauli matrix $X = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$, where $a \in \mathbb{R}_{\neq 0}$. Since X has eigenvalues $1, -1$ with corresponding eigenvectors $v = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}$ and $w = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{bmatrix}$, e^{aX} has eigenvalues e^a, e^{-a} with the same eigenvectors, and thus

$$e^{aX} = e^a v v^\intercal + e^{-a} w w^\intercal = \frac{1}{2} \begin{bmatrix} e^a + e^{-a} & e^a - e^{-a} \\ e^a - e^{-a} & e^a + e^{-a} \end{bmatrix} = C_a \cdot \begin{bmatrix} e^{K_a} & \operatorname{sign}(a) e^{-K_a} \\ \operatorname{sign}(a) e^{-K_a} & e^{K_a} \end{bmatrix}, \tag{16}$$

where $C_a = \frac{1}{2}\sqrt{(e^a + e^{-a}) \cdot |e^a - e^{-a}|} = \frac{1}{2}\sqrt{|e^{2a} - e^{-2a}|}$ and

$$K_a = g(a) = \ln \sqrt{\frac{e^a + e^{-a}}{|e^a - e^{-a}|}} = \frac{1}{2} \ln \frac{e^a + e^{-a}}{|e^a - e^{-a}|}.$$

Let $R = \{\, i \in [n] \mid h_i^x \neq 0 \,\}$, $R_+ = \{\, i \in [n] \mid h_i^x > 0 \,\}$ and $R_- = \{\, i \in [n] \mid h_i^x < 0 \,\}$. For $i \in R$, let

$$K_i = g(h_i^x / M) \quad \text{and} \quad C_i = \frac{1}{2}\sqrt{|e^{2h_i^x/M} - e^{-2h_i^x/M}|}. \tag{17}$$

We have

$$\begin{aligned}
&\prod_{i=1}^{M} \langle s^{(k)} |\, e^{\frac{1}{M} H_x} \,| s^{(k+1)} \rangle \\
=&\mathbf{1}[s_{k,i} = s_{k+1,i} \forall k \in [M], i \notin R] \left(\prod_{i \in R} \prod_{k=1}^{M} C_i \exp(K_i\, s_{k,i}\, s_{k+1,i}) \right) \prod_{i \in R_-} \prod_{k=1}^{M} (s_{k,i} s_{k+1,i}) \\
=&\mathbf{1}[s_{k,i} = s_{k+1,i} \forall k \in [M], i \notin R] \prod_{i \in R} \prod_{k=1}^{M} C_i \exp(K_i\, s_{k,i}\, s_{k+1,i}),
\end{aligned} \tag{18}$$

where the second equality follows because, for any $i \in [n]$ and assignment $(s_{k,i})_{k \in [M]} \in \{\pm 1\}^M$, $\prod_{k=1}^{M} (s_{k,i} s_{k+1,i}) = \prod_{k=1}^{M} s_{k,i}^2 = 1$.

Hence, substituting Equations (13) and (18) into Equation (12), we obtain

$$Z' = \prod_{i \in R} C_i^M \times Z'' \quad \text{with}$$

$$Z'' = \sum_{\substack{(\mathbf{s}^{(i)})_i \in \{\pm 1\}^n \\ s_{1,i} = \cdots = s_{M,i} \forall i \notin R}} \exp\left(\frac{1}{2M} \sum_{k=1}^{M} \left(\sum_{i,j=1}^{n} J_{ij} s_{k,i} s_{k,j} + \frac{1}{M} \sum_{i=1}^{n} h_i^z s_{k,i} + \sum_{i \in R} K_i s_{k,i} s_{k+1,i} \right) \right)$$

$$= \sum_{\mathbf{x}_i \in \Omega_i \forall i \in [n]} \exp\left(\frac{1}{2} \sum_{i,j \in [n]} J_{ij} \langle \mathbf{x}_i, \mathbf{x}_j \rangle + \sum_{i \in [n]} w^{(i)}(\mathbf{x}_i) \right)$$

thus

$$Z' = \prod_{i : h_i^x \neq 0} \left(\frac{1}{2} \sqrt{|e^{2h_i^x/M} - e^{-2h_i^x/M}|} \right)^M \sum_{\mathbf{x}_1 \in \Omega_1, \cdots, \mathbf{x}_n \in \Omega_n} \exp\left(H_{\text{classical}}((\mathbf{x}_i)_{i \in [n]}) \right),$$

which concludes the proof of the lemma. ◀

Proof of Lemma 15. For the proof, we will need the following result due to [13] (restated in [30, Theorem 2.2]).

▶ **Proposition 21** ([30, Theorem 2.2]). *Let $W_1, \ldots, W_\ell$ be independent random variables with $\frac{\mathbb{E}[W_i^2]}{\mathbb{E}[W_i]^2} \leq B$ for $i \in [\ell]$. Let $\widehat{W} = W_1 \cdots W_\ell$. Let S_i be the average of $16B\ell/\varepsilon^2$ independent random samples from W_i for $i \in [\ell]$. Let $\widehat{S} = S_1 S_2 \cdots S_\ell$. Then*

$$\Pr\left((1-\varepsilon)\mathbb{E}[\widehat{W}] \leq \widehat{S} \leq (1+\varepsilon)\mathbb{E}[\widehat{W}] \right) \geq \tfrac{3}{4}.$$

Let $\ell = L(\beta_{\max} - \beta_{\min})$, and define the sequence $\beta_{\min} = \beta_0 < \beta_1 < \cdots < \beta_\ell = \beta_{\max}$ where $\beta_i = \beta_{\min} + \frac{i}{L}$ $\forall i \in \{0, \cdots, \ell\}$. For $i \in \{1, \cdots, \ell\}$, let W_i be the random variable obtained by sampling σ from $\mu_{\beta_{i-1}}$ and then outputting the value $\exp(-(\beta_i - \beta_{i-1})H(\sigma))$. Note that

$$\begin{aligned}
\mathbb{E}[W_i] &= \sum_{\sigma \in \Omega} \mu_{\beta_{i-1}}(\sigma) \exp(-(\beta_i - \beta_{i-1})H(\sigma)) \\
&= \frac{1}{Z(\beta_{i-1})} \sum_{\sigma \in \Omega} \exp(-(\beta_i - \beta_{i-1})H(\sigma)) \exp(-\beta_{i-1}H(\sigma)) \\
&= \frac{1}{Z(\beta_{i-1})} \sum_{\sigma \in \Omega} \exp(-\beta_i H(\sigma)) \\
&= \frac{Z(\beta_i)}{Z(\beta_{i-1})},
\end{aligned}$$

and thus for $\widehat{W} = W_1 W_2 \cdots W_\ell$,

$$\mathbb{E}[\widehat{W}] = \prod_{i=1}^{\ell} \mathbb{E}[W_i] = \frac{Z(\beta_{\max})}{Z(\beta_{\min})}.$$

Since $(\beta_i - \beta_{i-1}) = \frac{1}{L}$ and $|H(\sigma)| \leq L$, $\frac{\mathbb{E}[W_i^2]}{\mathbb{E}[W_i]^2} \leq B$ for $B = e^2$.

Let S_i be the average of $16B\ell/\varepsilon^2$ independent random samples from W_i for $i \in [\ell]$, and let $\widehat{S} = S_1 S_2 \cdots S_\ell$. The algorithm will output $\hat{Z} = Z(\beta_{\min})\widehat{S}$. By Proposition 21,

$$\Pr\left((1-\epsilon)Z(\beta_{\max}) \leq \hat{Z} \leq (1+\epsilon)Z(\beta_{\max}) \right) = \Pr\left((1-\varepsilon)\mathbb{E}[\widehat{W}] \leq \widehat{S} \leq (1+\varepsilon)\mathbb{E}[\widehat{W}] \right) \geq \frac{3}{4}. \blacktriangleleft$$

Proof of Lemma 16. We use the same algorithm as in Lemma 15. We couple the distributions μ_β and $\hat{\mu}_\beta$ via a coupling Π_β such that

$$\mathbb{P}_{(\sigma,\sigma')\sim\Pi_\beta}[\sigma \neq \sigma'] \leq \epsilon_{TV}.$$

By a union bound, with probability at least $1 - k\,\epsilon_{TV}$, all k samples drawn from $\hat{\mu}_\beta$ are identical to the corresponding samples drawn from μ_β. Consequently, with probability at least $1 - k\,\epsilon_{TV}$, the execution of the algorithm using samples from $\hat{\mu}_\beta$ is identical to that of the algorithm using samples from μ_β. Hence the guarantee on the output follows from Lemma 15. ◀

Proof of Proposition 18. Clearly, $\frac{e^a+e^{-a}}{|e^a-e^{-a}|} \geq 1$, so $0 \leq \ln\left(\frac{e^a+e^{-a}}{|e^a-e^{-a}|}\right)$.

Wlog, we can assume $a \in (0,1]$. The inequality $\ln\left(\frac{e^a+e^{-a}}{|e^a-e^{-a}|}\right) \leq \ln\frac{2}{|a|}$ becomes equivalent to

$$f(a) := e^a - e^{-a} - \frac{a}{2}\cdot(e^a + e^{-a}) \geq 0.$$

We calculate

$$f'(a) = e^a + e^{-a} - \frac{1}{2}(e^a + e^{-a}) - \frac{a}{2}\cdot(e^a - e^{-a}) = \frac{1}{2}((e^a + e^{-a}) - a\cdot(e^a - e^{-a}))$$

and

$$2f''(a) = e^a - e^{-a} - (e^a - e^{-a}) - a(e^a + e^{-a}) = -a(e^a + e^{-a}) \leq 0 \quad \forall a \in [0, +\infty).$$

Thus,

$$\forall a \in [0,1] : f'(a) \geq f'(1) = \frac{1}{e} > 0,$$

and hence

$$\forall a \in [0,1] : f(a) \geq f(0) = 0.$$ ◀