

Improved Product-State Approximation Algorithms for Quantum Local Hamiltonians

Thiago Bergamaschi ✉

Department of Computer Science, University of California, Berkeley, CA, USA

Abstract

The ground state energy and the free energy of Quantum Local Hamiltonians are fundamental quantities in quantum many-body physics, however, it is QMA-Hard to estimate them in general. In this paper, we develop new techniques to find classical, additive error product-state approximations for these quantities on certain families of Quantum k -Local Hamiltonians. Namely, those which are either dense, have low threshold rank, or are defined on a sparse graph that excludes a fixed minor, building on the methods and the systems studied by Brandão and Harrow, Gharibian and Kempe, and Bansal, Bravyi and Terhal.

We present two main technical contributions. First, we discuss a connection between product-state approximations of local Hamiltonians and combinatorial graph property testing. We develop a series of *weak Szemerédi regularity* lemmas for k -local Hamiltonians, built on those of Frieze and Kannan and others. We use them to develop *constant time* sampling algorithms, and to characterize the “vertex sample complexity” of the Local Hamiltonian problem, in an analog to a classical result by Alon, de la Vega, Kannan and Karpinski. Second, we build on the information-theoretic product-state approximation techniques by Brandão and Harrow, extending their results to the free energy and to an asymmetric graph setting. We leverage this structure to define families of algorithms for the free energy at low temperatures, and new algorithms for certain sparse graph families.

2012 ACM Subject Classification Theory of computation → Quantum information theory; Theory of computation → Approximation algorithms analysis

Keywords and phrases Approximation Algorithms, Quantum Information

Digital Object Identifier 10.4230/LIPIcs.ICALP.2023.20

Category Track A: Algorithms, Complexity and Games

Related Version *Full Version*: <https://arxiv.org/abs/2210.08680>

Funding *Thiago Bergamaschi*: NSF GRFP No. DGE 2146752.

1 Introduction

The mean-field approximation is a popular heuristic in quantum many-body physics, in which product-states are used as an ansatz for generic quantum states. The low-energy states of quantum systems may be highly entangled objects, and possibly exponentially more complex than simple (unentangled) product states. This often makes computing properties of these low-energy states classically intractable. From a complexity-theoretic point of view, the mean-field approach casts these quantum problems that are in the complexity class QMA [34], into problems in NP, since product-states have a polynomial-size description and can act as classical, efficiently verifiable certificates. However, in the absence of a hardness-of-approximation result for QMA [1, 6, 2, 29] and assuming $\text{QMA} \neq \text{NP}$, it is generally unknown if the ground states of quantum systems can even have “good” approximations with succinct classical descriptions, let alone if we can compute or approximate them efficiently.

In this work, we develop a series of classical algorithms to efficiently find mean-field approximations for quantum systems described by *local Hamiltonians*, and we develop new techniques to show that good mean-field approximations exist for fairly general classes of



© Thiago Bergamaschi;
licensed under Creative Commons License CC-BY 4.0

50th International Colloquium on Automata, Languages, and Programming (ICALP 2023).

Editors: Kousha Etessami, Uriel Feige, and Gabriele Puppis; Article No. 20; pp. 20:1–20:18

Leibniz International Proceedings in Informatics



LIPICs Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany



these systems. A local Hamiltonian corresponds to a sparse matrix $H \in \mathbb{C}^{d^n \times d^n}$ which is exponentially large in the number n of quantum particles (or qudits), and can be described as a sum over “local” terms $H = \sum_{e \in E} h_e$ defined by some hypergraph $G = ([n], E)$. H is said to be k -local if each hyperedge $e \in E$ is a k -tuple of vertices in $[n]$, and is said to have “bounded” interaction strengths if the operator norm $\|h_e\|_\infty$ is at most a constant independent of n for each hyperedge in E ¹.

It is well known that the existence of product-state approximations to H is very sensitive to the structure of the underlying interaction graph G . In a seminal result, Brandão and Harrow [15] proved that so long as H has bounded interaction strengths, and is defined on a graph G of high degree or small expansion, then there exists a product state which approximates the ground state energy of H up to an additive error $\epsilon \cdot m$ (scaling with the number of edges or “interactions” m of H). Their results can be interpreted as rigorous proofs of accuracy of the mean-field approximation to the ground state energy of certain systems, and they opened the door to classical approximation schemes to find these “good” mean-field solutions efficiently. One of the main focuses of this work is to relax certain assumptions on the structure of the interaction graphs G , to extend the scope of their algorithms and existence statements.

The second main focus of this work is to study the structure and classical computation of properties of quantum systems in thermal equilibrium. The Helmholtz Free Energy $F(\beta)$ of a Quantum Local Hamiltonian H at a given temperature β^{-1} arises as an approximate counting analog to the ground state energy, as it reveals the degeneracy of the ground state (the number of QMA witnesses), the density of states of the Hamiltonian, as well as the existence of phase transitions. Quantitatively, $F(\beta)$ can be described as the optimum of a maximum entropy program:

$$F(\beta) \equiv \min_{\rho \geq 0, \text{Tr}\rho=1} f(\rho) = \min_{\rho \geq 0, \text{Tr}\rho=1} \text{Tr}[H\rho] - S(\rho)/\beta \quad (1)$$

where the optimizer $\rho \propto e^{-\beta H}$ of the program above is called the Gibbs state of H . The computational complexity, and in particular the hardness of approximation of $F(\beta)$ is similarly not comprehensively understood. While QMA-Hard to estimate in general due to a reduction to the “low temperature” limit, and exactly computable in polynomial time using a #P oracle [19], it would seem there is much to uncover regarding the computational tradeoffs between error and temperature [16].

1.1 Our Main Contributions

In this section we overview our main contributions, which we present formally and in more detail in section 2.2.

Rigorous Mean-Field Approximations and Guarantees in NP

Our first contributions concern improvements and extensions to the existence statements by Brandão and Harrow [15]. Their methods had roots in the information-theoretic techniques by [39] and [10], developed in the context of approximating CSPs using the Lasserre Hierarchy. Informally, we show how to use their self-decoupling arguments to construct mixed states which are tensor products of single-particle mixed states, which approximate the Free Energy up to an additive error. We view these results as rigorous proofs of accuracy for the mean-field

¹ Please refer to section 2.1 for more background on local Hamiltonians and Schatten norms.

approximation to the Free Energy of Quantum Local Hamiltonians, and they imply that approximating the Free Energy of dense Hamiltonians up to an extensive error (scaling with the number of edges) is in NP.

► **Theorem 1.** Fix $k, d = O(1)$, and an inverse temperature β . Let $H = \sum_{e \in E} h_e$ be a k -Local Hamiltonian on n qudits of local dimension d , and m interactions each of strength $\|h_e\|_\infty \leq 1$. Then, there exists a product state σ_β such that

$$F \leq f(\sigma_\beta) = \text{Tr}[H\sigma_\beta] - S(\sigma_\beta)/\beta \leq F + O(n^{\frac{k-1}{3}} m^{2/3}) \quad (2)$$

That is, σ_β is an $O(n^{\frac{k-1}{3}} m^{2/3})$ additive error approximation to the Free Energy of H .

Note that when $k = 2$ (Hamiltonians on Graphs), the error becomes $O(n^{1/3} m^{2/3}) = O(m/D^{1/3})$, which recovers Brandão and Harrow's [15] result in terms of the *average* degree $D = m/n$ of the graph.

We emphasize two important points about the result above. First and foremost, the existence of approximations to $F(\beta)$ in NP implies that we can now use classical approximation schemes to search for optimal mean-field approximations to the free energy, and they will also be good additive approximations to the “entangled value” of $F(\beta)$. As we later discuss, this enables us to import practically all the previous machinery of approximation schemes for the ground state energy, to the Free Energy, developing novel algorithms for many quantum systems and improving on recent results by Bravyi et al. [16].

The second point of emphasis is that the result above holds at all temperatures β^{-1} . In this fashion, we are able to bypass the “low temperature bottleneck” of many approximation schemes for the Free Energy which constrain approaches in previous work, such as the polynomial interpolation method [11, 27] or Markov Chain Monte Carlo methods. We present a comprehensive comparison with previous work and the scope of our techniques for thermal systems in section 2.3.

Hamiltonian Regularity Lemmas and Approximation Algorithms

From an algorithmic point of view, our main contribution is a connection between product state approximations and graph property testing. We discuss quantum analogs of the *weak Szemerédi regularity lemmas* for dense graphs, hyper-graphs and low-threshold rank graphs [22, 3, 23], developed in the context of additive approximation schemes for Max-Cut and Max-kCSPs. At their heart lies a powerful combinatorial characterization of these systems, Szemerédi's celebrated regularity lemma [42], which states that dense graphs can be approximately decomposed into unions of complete bipartite graphs. We develop natural, constructive generalizations of these results for Quantum Local Hamiltonians, by combining our new product state approximations with multi-coloured versions of known weak regularity results, leading to improved approximation algorithms and novel structural characterizations of local Hamiltonians. Our central result in this vein is an additive error approximation scheme for dense k -Local Hamiltonians, which runs in *constant time*:

► **Theorem 2.** Fix $d, k = O(1)$, $\epsilon > 0$, and let $H = \sum_e h_e$ be a k -Local Hamiltonian on n qudits of local dimension d and bounded strength interactions $\|h_e\|_\infty \leq 1$. Then, there exists a randomized algorithm which runs in time $2^{\text{poly}(1/\epsilon)}$, and with probability .99 returns an estimate for the ground state energy of H accurate up to an additive error of $\epsilon \cdot n^k$.

We report our sampling algorithms, including that in theorem 2, in the *probe model of computation* introduced by Goldreich et al. [26]. In a nutshell, the time complexity measured above corresponds to the number of queries to a description of H , see section 2.1 for more details.

Our simplest algorithm is remarkably clean to describe, and is based on the “Vertex Sample Complexity” result for Max-kCSPs by Alon et al. [3]: Given a Hamiltonian H , sample a uniformly random subset of qudits $Q \subset [n]$ of certain constant size $|Q| = q = \text{poly}(1/\epsilon)$, and let H_Q be the $d^q \times d^q$ matrix corresponding to the restriction to the interactions of H contained entirely in Q . Then, exactly diagonalize the (constant-sized) $d^q \times d^q$ matrix H_Q , and output its lowest eigenvalue multiplied by $(n/q)^k$. For small constant $d, k, \epsilon > 0$, it is clear that this approach requires just a constant number of queries to H . The challenge, of course, lies in proving that this estimate corresponds to a $\epsilon \cdot n^k$ additive error estimate to the true ground state energy of the original Hamiltonian H .

In the body, we show how these ideas can be used to develop improvements in runtime from $n^{\text{poly}(1/\epsilon)}$ to $\text{poly}(n, 1/\epsilon) + 2^{\text{poly}(1/\epsilon)}$ or just $2^{\text{poly}(1/\epsilon)}$ for a wide range of problems on Quantum Local Hamiltonians, such as approximation schemes for the ground state energy, the Free Energy, and for Hamiltonians defined on low threshold rank graphs.

2 Technical Overview

2.1 Background and Notation

Linear Algebra and Matrix Norms. Given an $w \times w$ matrix A we refer to $\|A\|_p$ as the Schatten p -norm of A (the L_p norm of the singular values of A), and we refer to $|A|_p$ as the L_p norm of the w^2 -dimensional vectorization of A . The graph decompositions are phrased in terms of the *cut norm* $\|A\|_C$ introduced by [22], defined by

$$A^+ = \max_{S_1, S_2 \subseteq [w]} \sum_{i \in S_1, j \in S_2} A_{ij} \text{ and } \|A\|_C = \max(A^+, (-A)^+) \quad (3)$$

where we have $\|A\|_C \leq \|A\|_{\infty \rightarrow 1} = \sup_{x \neq 0} \frac{|Ax|_1}{|x|_\infty} \leq 4 \cdot \|A\|_C$.

Asymptotic Notation. For any function $f(n)$ we refer to the asymptotic notation $\tilde{O}(f(n)) = O(f(n) \text{polylog}(f(n))) \leq c_1 \cdot f(n) \log^{c_2} f(n)$ for a choice of real positive constants c_1, c_2 .

Local Hamiltonians. We denote a k -Local Hamiltonian on n qudits of local dimension d via a $d^n \times d^n$ Hermitian matrix, which can be expressed as a sum of local interactions $H = \sum_{e \in E} h_e$. By “local”, we simply mean that each summand $h_e = H_e \otimes \mathbb{I}_{V \setminus e}$ acts non-trivially only on k particles at a time, as indicated by each k -tuple $e = (u_1 \cdots u_k)$ in a set of hyper-edges E . In this manner, we can specify any Local Hamiltonian “instance” simply by specifying the $d^k \times d^k$ submatrices of each local term. If $d, k = O(1)$, then the input has a polynomial-sized description in n . For notational convenience, we often omit the trivial support $\mathbb{I}_{V \setminus e}$. The ground state energy and the ground state of H are its minimum eigenvalue and corresponding eigenvector, and the *variational* minimum energy of H is the minimum energy of H among all product states $\min_{\rho = \otimes \rho_u} \text{Tr}[H \otimes_u \rho_u]$ with $\rho_u \in \mathbb{C}^{d \times d}$ and $\rho_u \geq 0, \text{Tr}_u[\rho_u] = 1$.

Interaction Graphs. We refer to the “Interaction Graph” $G = ([n], E)$ of a 2-Local Hamiltonian H as the graph with undirected edges $e = (u, v) \in E$ whenever the particles u, v interact non-trivially in H . That is, whenever the spectral norm is non-zero $\|H_e\|_\infty \neq 0$. By expressing each $d^2 \times d^2$ Hermitian matrix $H_{u,v} = \sum_{i,j \in [d^2]} H_{u,v}^{i,j} \cdot \sigma_u^i \otimes \sigma_v^j$ in an orthogonal basis decomposition, and grouping all the interactions with the same basis i, j , we refer to the i, j “Pauli Graph” as the subgraph of G induced on all the *directed* edges $e = (u, v)$ with

non-zero $H_{u,v}^{i,j} = d^{-2} \text{Tr}[H_{u,v} \sigma_u^i \otimes \sigma_v^j]$, with weighted adjacency matrix $J^{ij} = \{H_{u,v}^{i,j}\}_{u,v \in [n]}$. We note that the matrices J^{ij} are degenerate, since $J^{ij} = (J^{ji})^T$, but we often brush over this issue via a handshaking argument. If we are given a density matrix $\rho = \otimes \rho_u$ which is a product of single qudit density matrices with a basis decomposition $\rho_u = d^{-1} \sum_i \alpha_u^i \cdot \sigma^i$, then the energy of ρ , $\text{Tr}[H\rho]$ is a polynomial over the real variables α :

$$\sum_{(u,v) \in E} \text{Tr}[H_{u,v} \rho_u \otimes \rho_v] = d^{-2} \sum_{(u,v) \in E} \sum_{i,j \in [d^2]} H_{u,v}^{i,j} \alpha_u^i \cdot \alpha_v^j = (2d^2)^{-1} \sum_{i,j \in [d^2]} \sum_{u \neq v \in [n]} J_{uv}^{ij} \alpha_u^i \cdot \alpha_v^j \quad (4)$$

Model of Computation. We report our sampling algorithms in the *probe model of computation* introduced by [26] in the context of graph property testing. That is, we assume we can sample a uniformly random vertex or hyper-edge in $O(1)$ time (or “probes”). Formally, fixed a k -Local Hamiltonian “instance” $H = \sum_{e \in E} H_e \otimes \mathbb{I}_{V \setminus e}$, for any k -tuple of vertices/hyper-edge $e = (u_1 \cdots u_k)$, $u_i \in [n]$, we assume we can query the (constant-sized) $d^k \times d^k$ sub-matrix H_e in $O(1)$ time. We emphasize that since our goal is often a sublinear time algorithm, we always enforce that our algorithms output estimates for the energy (or free energy), and *implicit* descriptions of product states. If requested, these implicit descriptions can always be expanded into n -qudit product states in an additional $\text{poly}(n, 1/\epsilon)$ time.

2.2 Our Results

Approximation Guarantees in NP

The first of our results are rigorous proofs of accuracy of the mean-field approximation on Quantum k -Local Hamiltonians. We argue the existence of product states, or products of single-particle mixed states, which provide additive error approximations to the ground state energy and the free energy of these systems. We build on the information-theoretic techniques by Brandão and Harrow [15], presenting an extension to the free energy and modestly refining their techniques on generic (hyper-) graphs.

► **Theorem 3.** *Fix $k, d = O(1)$. Let $H = \sum_{e \in E} h_e$ be a k -Local Hamiltonian on n qudits of local dimension d , and m interactions each of strength $\|h_e\|_\infty \leq 1$. Then, there exists a product state $|\psi\rangle = \otimes_{u \in [n]} |\psi_u\rangle$, $|\psi_u\rangle \in \mathbb{C}^d$ such that*

$$\langle \psi | H | \psi \rangle \leq \min_{\phi} \langle \phi | H | \phi \rangle + O(n^{\frac{k-1}{3}} m^{2/3}) \quad (5)$$

In the body, we prove more general versions of the theorem above sensitive to the matrix of interaction strengths of H . Theorem 3 matches the previous results in [15] whenever the Hamiltonian is defined on D -regular or dense graphs $m = \Omega(n^k)$, and generalizes their statements to just depend on the number of edges m . In the setting of Theorem 3, whenever $m = \Omega(n^{k-1}/\epsilon^3)$, approximating the ground state energy of H up to additive error $\epsilon \cdot m$ is in the complexity class NP, as the product state has a polynomial size description and acts as a classical witness. While these optimal product states may still be NP-Hard to find in the worst case, there are many examples where one can approximate these solutions efficiently.

To extend both these information-theoretic ideas and algorithmic applications to the free energy, we need further insights on the structure of these product state approximations. We discuss² how the “entanglement-breaking” procedure of [15], not only approximately

² In section B of the full version.

preserves the energy, but in fact also increases the entropy as well. When applied to the Gibbs state, we show one can carefully extract a tensor product of single particle mixed-states which is a good approximation to the free energy. We formalize this statement in Theorem 4,

► **Theorem 4.** *Fix $k, d = O(1)$, and an inverse temperature β . Let $H = \sum_{e \in E} h_e$ be a k -Local Hamiltonian on n qudits of local dimension d , and m interactions each of strength $\|h_e\|_\infty \leq 1$. Then, there exists a product state $\sigma_\beta = \otimes_{u \in [n]} \sigma_u, \sigma_u \in \mathbb{C}^{d \times d}$ such that*

$$f(\sigma_\beta) = \text{Tr}[H\sigma_\beta] - S(\sigma_\beta)/\beta \leq F + O(n^{\frac{k-1}{3}} m^{2/3}) \quad (6)$$

We emphasize that the statement above implies a product state approximation exists at all temperatures β^{-1} (and recovers the ground state approximation at $T = 0$), and moreover uses very little of the underlying graph structure apart from the *average* dense condition.

Hamiltonian Weak Regularity Lemmas

We develop an approach to designing approximations algorithms for Local Hamiltonians based on *weak Szemerédi regularity lemmas*, which are approximate decompositions to graphs, matrices, and tensors [42, 22, 3, 23].

The idea behind this construction lies in a powerful tool in extremal combinatorics. In his celebrated regularity lemma, Szemerédi [42] proved that any dense graph can be approximated by a union of a constant number of complete bipartite graphs. However, the number of partitions grew very fast with the intended quality of approximation. Frieze and Kannan [22] developed a constructive decomposition under a weaker notion of approximation, what they referred to as a “weak” regularity lemma. Concretely, they prove that any real $n \times n$ matrix with bounded entries can be decomposed into a sum of $O(1/\epsilon^2)$ cut matrices (complete bipartite graphs), up to an error $\epsilon \cdot n^2$ in the cut norm. Moreover, [22] proved that one can in fact construct such a “cut decomposition” implicitly in time polynomial in $1/\epsilon$, which enabled them to devise constant time sampling-based approximation schemes for many problems on dense graphs.

We define a natural adaptation of their results to a quantum setting, by constructing an approximate decomposition H_D of a Local Hamiltonian H which is a sum over complete, bipartite, sub-Hamiltonians. The structure of H_D can be understood as a “multi-colored” matrix cut decomposition, as essentially we apply the cut decomposition by [22] to each term in a basis decomposition of H . For concreteness, let $H = \sum_{u,v} h_{u,v}$ be a 2-Local Hamiltonian on qubits, and let us consider re-writing its Pauli basis decomposition below. We suppress the identity terms $\otimes_{V \setminus \{u,v\}}$ on the qubits that each interaction acts trivially on.

$$H = \sum_{(u,v) \in E} h_{u,v} = \sum_{(u,v) \in E} \sum_{i,j \in \{I,X,Y,Z\}} h_{u,v}^{i,j} \sigma_u^i \otimes \sigma_v^j = \sum_{i,j \in \{I,X,Y,Z\}} \sum_{u < v} h_{u,v}^{i,j} \sigma_u^i \otimes \sigma_v^j \quad (7)$$

We associate each pair of indices $i, j \in \{I, X, Y, Z\}$ to a color, and consider the $n \times n$ real valued weighted adjacency matrix $J^{ij} = \{h_{u,v}^{i,j}\}_{u,v \in [n]}$ of the i, j “Pauli Graph”. By applying the cut decomposition by [22] to each of these 16 matrices J^{ij} , we construct an approximate decomposition of H into roughly $16 \cdot O(1/\epsilon^2)$ complete bipartite sub-Hamiltonians. In this context, a “complete bipartite sub-Hamiltonian” is defined by two Pauli matrices, (say, X, Y), two subsets $S, T \subset [n]$ (which, for now, we assume to be disjoint), and an interaction strength $\alpha \in \mathbb{R}$, and can be expressed as $\alpha \sum_{u \in S, v \in T} X_u \otimes Y_v$.

In the body we argue that the approximation guarantees in the cut norm are precisely what we need to ensure that for any product state $\sigma = \otimes_u \sigma_u$, the energy of σ under H or H_D are close: $\text{Tr}[H\sigma] \approx \text{Tr}[H_D\sigma]$. By further combining this product state regularity with our asymmetric product state approximations, we prove a stronger property on the spectra of H_D :

► **Lemma 5 (Informal).** *Fix $d, k = O(1)$ and a constant $\epsilon > 0$, and let $H = \sum_e h_e$ be a k -Local Hamiltonian on n qudits of local dimension d and m interactions of strength bounded by $\|h_e\|_\infty \leq 1$. Then, there exists a decomposition $H_D = \sum_i^s D^{(i)}$ of H into $s = O(1/\epsilon^2)$ complete bipartite sub-Hamiltonians such that*

$$\|H - H_D\|_\infty \leq \epsilon \cdot n^{k/2} m^{1/2} \quad (8)$$

Additive Error Approximation Schemes

Leveraging the structure of the Hamiltonian regularity (lemma 5) in combination with the product state approximation toolkit enables us to devise a series of approximation schemes for Quantum Local Hamiltonians. We follow the ideas of [22, 3, 23] in establishing LP relaxations to Max Cut and other Max CSPs, and we develop an SDP relaxation scheme for finding the minimal energy product state of a Local Hamiltonian. These ideas enable us to devise an efficient additive error approximation scheme for dense Hamiltonians,

► **Theorem 6 (Theorem 2, restatement).** *Fix $d, k = O(1)$ and $\epsilon > 0$. Let $H = \sum_e h_e$ be a k -Local Hamiltonian on n qudits of local dimension d , and m interactions of bounded strength $\|h_e\|_\infty \leq 1$. There exists a randomized algorithm which runs in time $2^{\tilde{O}(1/\epsilon^{2k-2})}$ in the probe model of computation, and with probability .99 computes an estimate for the ground state energy of H accurate up to an additive error of $\epsilon \cdot n^{k/2} \sqrt{m}$.*

We note that $n^{k/2} \sqrt{m} \geq m$, and thus in polynomial or sublinear time this approximation scheme only provides a non-trivial guarantee when the hyper-graph is dense, $m = \Omega(n^k / \log^c n)$ for some small positive constant c . However, it provides an improvement over the $n^{O(1/\epsilon^2)}$ time algorithms by [24] and [15] in this additive error regime. On the other hand, a simple explicit variant of this result provides a sub-exponential time approximation algorithm whenever $m = \omega(n^{k-1} \log n)$:

► **Theorem 7.** *In the context of Theorem 2, there exists a randomized algorithm which runs in time $\tilde{O}(n^k) \cdot 2^{\tilde{O}(n^k/\epsilon^2 m)}$ and with high probability computes an estimate for the ground state energy of H accurate up to an additive error of $\epsilon \cdot m$.*

Concretely, the key idea behind these relaxations is that for any product state $\sigma = \otimes_u \sigma_u$, the energy of σ on the cut decomposition H_D is a simple function of the average magnetization of a small number of subsets of the n qudits. To illustrate how this enables a relaxation scheme, consider a single complete bipartite sub-Hamiltonian, such as $H_{S,T} = \sum_{u \in S, v \in T} X_u \otimes Y_v$. The energy of σ on $H_{S,T}$ is

$$\text{Tr}[H_{S,T}\sigma] = \sum_{u \in S, v \in T} \text{Tr}_{u,v}[X_u \sigma_u \otimes Y_v \sigma_v] = \left(\sum_{u \in S} \text{Tr}[X_u \sigma_u] \right) \cdot \left(\sum_{v \in T} \text{Tr}[Y_v \sigma_v] \right), \quad (9)$$

simply the product of the average X direction magnetization of $S \subset [n]$ with the average Y magnetization of T . If we fix a “guess” $r, c \in [-n, n]$, one can introduce affine constraints on the single particle density matrices σ_u , constraining their average magnetizations to lie within a $\pm \gamma \cdot n$ range of the guess r, c :

$$r - \gamma \cdot n \leq \sum_{u \in S} \text{Tr}[X_u \sigma_u] \leq r + \gamma \cdot n, \quad (10)$$

$$c - \gamma \cdot n \leq \sum_{v \in T} \text{Tr}[Y_v \sigma_v] \leq c + \gamma \cdot n. \quad (11)$$

Then, we are guaranteed that any product state σ which is feasible for the constraints above must have energy in a range around the guess: $|\text{Tr}[H_{S,T}\sigma] - r \cdot c| \leq (2 \cdot \gamma + \gamma^2) \cdot n^2$. In this manner, one can discretize over the space of “guesses” (r, c) and define an overlapping set of convex constraints on the description of the product states σ , such that every product state is feasible for at least one set of constraints. Approximating the ground state energy among product states ultimately reduces to checking the feasibility of a constant number of SDPs, one for each guess of r, c , and outputting whichever gives us the smallest energy estimate.

Using the techniques by [23], we can extend these insights to the setting of symmetric 2-Local Hamiltonians defined on graphs of low threshold rank. They proved that the weak regularity results of [22] could be extended to low-threshold rank graphs, by constructing a cut decomposition of a low rank approximation to the normalized adjacency matrix of these graphs. While in the appendix we formalize approximation algorithms for generic symmetric Hamiltonians (on low threshold rank graphs), perhaps the most faithful extension of this result to the quantum setting would be its application to approximating the Quantum Max Cut [25, 37, 36, 38]. Given an undirected graph $G = (V, E)$, the “Quantum Max-Cut” corresponds to the maximum eigenvalue of the Hamiltonian

$$H = \frac{1}{2} \sum_{e \in E} \left(\mathbb{I}_u \otimes \mathbb{I}_v - X_u \otimes X_v - Y_u \otimes Y_v - Z_u \otimes Z_v \right) \otimes \mathbb{I}_{V \setminus \{u,v\}} \quad (12)$$

If A is the adjacency matrix of G and D the diagonal matrix of degrees, the δ -SOS threshold rank $t_\delta(A)$ of A is the number of eigenvalues of the normalized adjacency matrix $D^{-1/2}AD^{-1/2}$ which are outside of the range $[-\delta, \delta]$. We prove

► **Theorem 8.** *Fix $\epsilon, \delta > 0$. Let $G = (V, E)$ be a graph on n vertices and m edges with adjacency matrix A and threshold rank $t \equiv t_{\epsilon/2}(A)$. Then, there exists an algorithm which finds an $\epsilon \cdot m + O(n^{1/3}m^{2/3})$ additive error approximation to the Quantum Max Cut of G in time $\text{poly}(n, 1/\epsilon, t) + 2^{\tilde{O}(t/\epsilon^2)}$.*

For instance, sparse D -regular random graphs have $\Theta(D^{-1/2})$ -SOS threshold rank 1. In this manner, for any constant ϵ and if $D = \Omega(1/\epsilon^3)$, then one can compute an $\epsilon \cdot m$ approximation to the Quantum Max Cut of a D -regular random graph in polynomial time.

A series of works [30, 31, 32] showed that the matrix weak regularity lemma [22] could be used to approximate the free energy of Ising Models, and to give interesting structural results on the quality of the mean-field approximation and the “vertex sample complexity” of these systems. They observed that the maximum entropy program subject to the linear relaxation constraints described above, reveals properties of the Gibbs distribution and enables an additive error approximation to the free energy at all temperatures. By combining these ideas with the Hamiltonian regularity Lemma 5 and Theorem 4 on product state approximations to the free energy, we develop a series of additive error approximation schemes for the free energy of Quantum Local Hamiltonians. The first of which is a constant time approximation scheme, which provides an additive error guarantee in a low temperature regime.

► **Theorem 9.** Fix $k, d = O(1)$, and $\epsilon, \delta > \omega(n^{-1/(2k-2)})$ and an inverse temperature $\beta > 0$, and let H be a k -Local Hamiltonian on n qudits of local dimension d and m bounded strength interactions. Then, there exists an algorithm that runs in time $2^{\tilde{O}(\epsilon^{2-2k})} \cdot O(\delta^{-2})$ in the probe model of computation, that returns an estimate to the free energy accurate up to an additive error of $\epsilon n^{k/2} m^{1/2} + \delta n/\beta$ and is correct with probability .99.

We emphasize that the free energy is a convex program regularized by temperature, and thereby our approximation schemes often incur a tradeoff between combinatorial errors and thermal (temperature dependent) errors. In the low temperature regime, whenever $\beta = \Omega(n^{1-k/2} m^{-1/2})$, the algorithm above recovers the behavior of the ground state energy approximation scheme, and is largely temperature independent. However, as the temperature increases and surpasses the threshold, the leading source of error becomes the thermal error $\delta n/\beta$. In our second algorithm, we show that an explicit approach significantly improves this thermal error dependence, at the cost of a polynomial runtime.

► **Theorem 10.** Fix $k, d = O(1)$, and $\epsilon, \delta > 0$ and an inverse temperature $\beta > 0$, and let H be a k -Local Hamiltonian on n qudits of local dimension d and m bounded strength interactions. Then, there exists an algorithm that runs in time $2^{\tilde{O}(\epsilon^{-2})} \cdot \tilde{O}(n^k \log 1/\delta)$, that returns an estimate to the free energy accurate up to an additive error of $\epsilon n^{k/2} m^{1/2} + \delta n/\beta$ and is correct with high probability.

The Vertex Sample Complexity

The Regularity Lemma Lemma 20 enables us to derive an insightful structural statement for Local Hamiltonians. Namely, the definition of a “vertex sample complexity” for Local Hamiltonians of bounded interaction strengths, in an analogy to the vertex sample complexity of Max-kCSPs of [3] and [4]. They showed that the restriction of any Max-kCSP to a uniformly random sample of $\text{poly}(1/\epsilon)$ variables, sufficed to estimate the maximum number of satisfiable clauses up to an additive error of $\epsilon \cdot n^k$. We develop a generalization of this result to Quantum Local Hamiltonians, by combining the Hamiltonian regularity lemma with some extensions to the proof techniques by [3] to SDPs.

► **Theorem 11.** Fix $d, k = O(1)$ and $\epsilon > 0$, and let H be a k -local Hamiltonian on n qudits of local dimension d and m bounded interaction strengths. Let $Q \subset [n]$ be a uniformly random sample of $q = \Omega(\epsilon^{-6} \log 1/\epsilon)$ of those qudits, and let H_Q be the sum of interactions with support contained entirely in Q . Then, with probability 0.99,

$$\left| \min_{\rho} \text{Tr}[H\rho] - \frac{n^k}{q^k} \min_{\rho_Q} \text{Tr}[H_Q\rho_Q] \right| \leq \epsilon \cdot n^k \quad (13)$$

We rely crucially on the guarantee of product state approximations to Quantum Local Hamiltonians in this regime of additive error. Indeed, one of the directions of the statement above is quite intuitive for both classical and quantum systems: If the ground state energy of H is low, then the ground state energy of the restriction H_Q can't be much higher than the estimate. This is since the reduced density matrix $\rho_Q = \text{Tr}_{V \setminus Q}[\psi]$ of the ground state ψ of H , probably also has low energy $\text{Tr}[H_Q\rho_Q] \approx \frac{q^2}{n^2} \cdot \text{Tr}[H\psi]$, and the true ground state energy of H_Q can only be lower than that.

In the converse, however, lies an interesting “semi-classical” characterization of this additive error regime. Note that if the ground state energy of H is “high”, then in particular there doesn't exist any product states with low energy on H . Using the proof techniques in [3], we show this implies the existence of a certain succinctly describable classical certificate

to this product-state “infeasibility”, which we sample from to prove the absence of product states with low energy on H_Q . Here is where we require the product state approximations of Theorem 3: for sufficiently large Q , the absence of low energy product states for H_Q must imply a high ground state energy for H_Q . In this sense, the ground state energy of H_Q can’t be much lower than its estimate either.

As a straightforward corollary to this structural result, now we can easily devise an algorithm which provides an additive error guarantee by exactly diagonalizing the Hamiltonian H_Q on $q = \tilde{O}(\epsilon^{-6})$ vertices in time $2^{\tilde{O}(1/\epsilon^6)}$. However, we can in fact do slightly better, simply by applying the additive error, product state approximation algorithm by [24] to the subsample:

► **Corollary 12.** *Fix $d, k = O(1)$ and $\epsilon > 0$, and let H be a k -Local Hamiltonian on n qudits of local dimension d and m bounded interaction strengths. There exists a randomized algorithm which runs in time $2^{O(\epsilon^{-2})}$, and with probability .99 outputs an estimate to the ground state energy accurate up to an additive error of $\epsilon \cdot n^k$.*

Aside from the improved dependence on k in the exponent, this result may seem to only subtly differ from that in Theorem 2. However, we emphasize that Theorem 2 requires an exponential number in $1/\epsilon$ of samples of vertices, whereas Theorem 11 guarantees a polynomial number suffices.

Approximation Schemes on Graphs that exclude a Fixed Minor

Finally, we develop novel singly-exponential time algorithms for sparse, 2-Local Hamiltonians defined on graphs that exclude a fixed minor. Formally, the family of h -minor free graphs are all the graphs G that can not produce another (smaller) graph h , by deleting edges and vertices and by contracting edges [41]. Planar graphs, and bounded genus graphs (such as torioids) are among the interesting special cases of these classes. Our approach builds on previous work by [9] and [15] on planar graphs, using more general combinatorial decompositions [21] and improving on their “quantum-to-classical” mappings. We show how such 2-Local Hamiltonians can be approximately understood as classical Max k -CSPs defined on the high degree vertices in the graph, and develop a dynamic programming algorithm to solve it using a simple hyper-dimensional version of a tree decomposition. Our first result for these systems is a classical algorithm to approximate the ground state energy in time singly exponential in $\text{poly}(1/\epsilon)$,

► **Theorem 13.** *Fix $\epsilon > 0$. Let H be a 2-Local Hamiltonian defined on n qubits and $m = \Theta(n)$ bounded strength interactions of norm < 1 , configured on an h -minor free graph $G = (V, E)$ where the minor is constant size $|h| = O(1)$. Then, we can approximate the ground state energy of H up to additive error $\epsilon \cdot n$, in time $\text{poly}(n) + n \cdot 2^{\text{poly}(1/\epsilon)}$.*

We build on these ideas by combining them with our information-theoretic techniques for the free energy of quantum systems, to construct novel algorithms for the free energy of these classes of sparse graphs at low temperatures as well.

► **Theorem 14.** *Fix $\epsilon > 0$ and an inverse temperature β . Let H be a 2-Local Hamiltonian on n qubits and $m = \Theta(n)$ bounded strength interactions of norm < 1 , configured on an h -minor free graph $G = (V, E)$ where the minor is constant size $|h| = O(1)$. Then, we can approximate the the free energy $F(\beta)$ of H up to additive error $\epsilon \cdot n$, in time $\text{poly}(n) + n \cdot \max(2, \beta^{-1})^{\text{poly}(1/\epsilon)}$, respectively.*

2.3 Related Work

Classical Approximation Schemes for QMA Complete Problems

While the systematic study of approximation algorithms to QMA-Complete problems is still emerging, there are a number of works we would like to highlight on the topic. [9] developed classical approximation schemes for ground state energies of classical and Quantum 2-Local Hamiltonians configured on planar graphs (of bounded degree, in the quantum case). They leveraged Baker’s technique [8] and structural properties of planar graphs to approximately decompose the Hamiltonian into non-interacting partitions, which then could be analyzed by exact diagonalization, or dynamic programming. [24] were among the first to construct an approximation algorithm for the k -Local Hamiltonian Problem. They argued that product states can provide a d^{-k+1} -relative factor approximations to the ground state energy of k -Local Hamiltonians defined on qudits, similarly to how Max Cut admits a $1/2$ multiplicative approximation. They then developed an approximation algorithm for the variational problem of finding the minimal energy product state of a given Local Hamiltonian H . It constructs a product state that provides an (extensive) $\epsilon \cdot n^k$ additive approximation to the ground state energy, in runtime $n^{O(\epsilon^{-2} \log 1/\epsilon)}$. Their approach was based on an adaptation of a classical technique, the “exhaustive sampling method” by [7] to the quantum setting, developed in the context of approximating Max Cut on dense graphs.

Later, [15] developed information-theoretic techniques to argue the existence of product state approximations to the the ground state energy. More precisely, they show that so long as H is *everywhere dense* ($\Omega(n^{k-1})$ minimum degree), has bounded expansion, or is clustered into regions of sub-volume law entanglement entropy, there exist product states that provide *additive error* approximations to the minimum energy. Leveraging their information-theoretic statements, they turned the algorithm of [24] into a PTAS for the ground state energy, albeit only meaningful when the number of interactions $m = \Omega(n^k)$. Additionally, they devise approximation schemes for Quantum Hamiltonians defined on generic planar graphs (not just those of bounded degree), solving an open problem posed by [9]. Their key insight was what we refer to as a “high-low degree” technique, in which one could consider a product state over all vertices of degree larger than some tunable cutoff Δ , and a generic (entangled) quantum state over the hilbert space of the low-degree particles, while incurring only a small error to the ground state energy. It is worthwhile to raise however, that the runtime of the resulting algorithm is triply-exponential in $1/\epsilon$, where the algorithm returns an $\epsilon \cdot n$ additive approximation.

More recently, in the context of *relative error* approximation schemes, [28] showed that one can find a product state within a relative error of l of the ground state of a traceless k -Local Hamiltonian of bounded norm, where l is the maximum degree of the underlying hyper-graph. [18] devised a $O(\log n)$ multiplicative approximation scheme to the ground state energy of 2-Local traceless Hamiltonians by rounding the solutions of SDPs to product states.

Classical Approximation Schemes for the Free Energy of Quantum Systems

Our results also contribute to a rich literature of classical techniques for thermal quantum systems. Perhaps the most well known of these techniques are the Quantum Monte Carlo methods, which approximate the quantum partition function of a quantum system to that of a classical spin system, which in turn is approximated via Markov chain Monte Carlo methods. Despite the enormous practical success of these techniques, rigorous proofs of convergence have only been presented in certain restricted systems [17, 13, 20], and they generically

are efficient only in the high temperature limit. Another high-temperature technique is the polynomial interpolation method [11, 27], based on a Taylor expansion of the partition function in the high temperature limit. Although both of these approaches are only provably efficient either on restricted classes of systems (such as substochastic Hamiltonians) and/or in the high temperature limit (typically β is a constant, or at most $O(\log n)$), they provide quite strong notions of approximation. In fact, they generally provide $(1 + \epsilon)$ multiplicative approximations to the partition function (which translates to an ϵ additive approximation to the free energy), while in this paper we only attempt *extensive*, additive, $\epsilon \cdot m$ error approximations to the free energy.

By approaching the problem via this weaker notion of error, it is possible to devise approximation schemes in a much wider range of temperatures. A recent result by [16] presented an algorithm that estimates the free energy of dense Local Hamiltonians, also building on the information-theoretic techniques by [15]. Their approach is based on a quantum generalization to a classical correlation rounding approach by [40], and their algorithm finds a $\epsilon \cdot n^2$ additive approximation to the free energy of 2-Local Hamiltonians, in runtime $n^{O(\epsilon^{-2})}$.

Comparison to Previous Work

To conclude our introduction we summarize our algorithmic improvements in contrast to previous known constructions for the quantum systems studied. In table 1 below we label the Hamiltonians, and runtime and accuracy guarantees of the additive error approximation schemes in previous work for the systems we consider. In table 2, we present our results for these same systems.

For simplicity, unless otherwise stated we concern ourselves with Quantum Local Hamiltonians of bounded interaction strengths $\|H_e\|_\infty \leq 1$ on n qubits and m interactions. In both tables, we refer to a “low threshold rank” Hamiltonian as having constant ϵ -SOS threshold rank of its interaction graph. With the exception of the recent work by [16], all the results in table 1 concern ground state energy approximation schemes.

■ **Table 1** A summary of previous algorithms for related Quantum Systems.

Result	System/Context	Accuracy	Runtime
[24]	k -local Hamiltonians	$\epsilon \cdot n^k$	$n^{\tilde{O}(\epsilon^{-2})}$
[15]	Low Threshold Rank Hamiltonians	$\epsilon \cdot \sum_{e \in E} \ H_e\ _\infty$	$n^{O(\epsilon^{-1})}$
[16]	Free Energy of 2-local Hamiltonians	$\epsilon \cdot n^2 + \delta \cdot n/\beta$	$n^{\tilde{O}(\epsilon^{-2})} \cdot O(\log 1/\delta)$
[9]	Planar Graphs of bounded degree Δ	$\epsilon \cdot \sum_{e \in E} \ H_e\ _\infty$	$n^{O(1)} \cdot 2^{2^{\text{poly}(\Delta, \epsilon^{-1})}}$
[15]	Planar Graphs	$\epsilon \cdot \sum_{e \in E} \ H_e\ _\infty$	$n^{O(1)} \cdot 2^{2^{2^{\text{poly}(\epsilon^{-1})}}}$

We remark* that the runtime results for k -local Hamiltonians are reported in the probe model [26], and thus may seem a priori incomparable to more standard model runtimes. However, we emphasize that we can easily convert between models by suitably pre-processing the input Hamiltonian and underlying Graph. For instance, if we are allowed query access to the input Hamiltonian in time $O(1)$, and arithmetic operations on entries of H take time $O(1)$, but sampling a random element of $[n]$ takes time $O(\log n)$, then the algorithm of theorem 11 outputs an estimate to the ground state energy in total time $O(\text{poly}(1/\epsilon) \cdot \log n + 2^{\text{poly}(1/\epsilon)})$ -

■ **Table 2** The main algorithms in this work.

System	Context	Accuracy	Runtime*
k -local Hamiltonians	G.S. Energy	$\epsilon \cdot n^k$	$2^{\text{poly}(\epsilon^{-1})}$
	Free Energy	$\epsilon \cdot n^k + \delta \cdot n/\beta$	$2^{\text{poly}(\epsilon^{-1})} \cdot O(\delta^{-2})$
Low Threshold Rank Quantum Max Cut	Maximum Eigenvalue	$\epsilon \cdot m + O(n^{1/3}m^{2/3})$	$(n/\epsilon)^{O(1)} + 2^{\tilde{O}(1/\epsilon^2)}$
h -Minor Free Graphs of bounded degree Δ	G.S. Energy	$\epsilon \cdot n$	$n^{O(1)} + n \cdot 2^{\text{poly}(\Delta, \epsilon^{-1})}$
	Free Energy		
h -Minor Free Graphs	G.S. Energy	$\epsilon \cdot n$	$n^{O(1)} + n \cdot 2^{\text{poly}(\epsilon^{-1})}$
	Free Energy		$n^{O(1)} + n \cdot \max(2, \beta^{-1})^{\text{poly}(\epsilon^{-1})}$

as we only require a $\text{poly}(1/\epsilon)$ number of sampled vertices. On the other hand, if don't have query access to the description of H , simply spending initial $O(d^{2k}(n+m)) = O(n+m)$ preprocessing time to read out the description of H is sufficient to reduce the setting to the previous one, assuming $d, k = O(1)$.

3 Discussion

We conclude this work by raising some open problems. The first of which is a curious gap between the quality of the mean field approximation to classical and Quantum Local Hamiltonians. To contrast our results to those in the classical setting, [14, 12, 30, 32] studied the quality of the mean-field approximation to classical spin glass models with generic interaction matrices. The work of [32] culminated in the result that the mean-field approximation is within an additive error of $O(n^{2/3}m^{1/3})$ of the free energy, a strictly better dependence on the number of interactions than our upper bound, $O(n^{1/3}m^{2/3})$. As both these results have roots in the information-theoretic techniques by [39], it seems intriguing to ask whether there is some deeper structure. A possible direction would be to combine the regularity insights with the correlation rounding techniques, as in [32]. However, there remain certain technical obstacles to approaching the free energy of quantum systems with the regularity lemma, namely analyzing the matrix exponential of the cut decomposition H_D .

Another interesting problem is to improve the weak regularity results for “low threshold rank” Hamiltonians (Such as theorem 8 and section G of the full version). While we are able to devise approximation schemes based on graph regularity for a range of Hamiltonians whose interaction graphs have low threshold rank, we are unable to provide an actual construction of an approximate Hamiltonian H' . It would also be interesting to see whether the coarsest partition technique could be lifted to be applied to more general low threshold rank Hamiltonians, as opposed to relying on the high degree of symmetry of the Quantum Max Cut.

Finally, while the focus of this paper is on product-state approximations, the author considers it to be an outstanding open problem whether one can devise entangled ansatz's for classical approximation schemes to quantum problems. For examples, see [33, 5, 35], who devised low-depth quantum circuits which perform slightly better than the best product state on certain Hamiltonians.

4 Organization

In the appendix, we present a proof of the Hamiltonian regularity lemma 20, and, for readability, defer to the full version (<https://arxiv.org/abs/2210.08680>) our information-theoretic statements and algorithms.

References

- 1 Scott Aaronson. The quantum pcp manifesto, October 2006. URL: <http://www.scottaaronson.com/blog/?p=139>.
- 2 Dorit Aharonov, Itai Arad, Zeph Landau, and Umesh V. Vazirani. The detectability lemma and quantum gap amplification. In *STOC '09*, 2009.
- 3 Noga Alon, Wenceslas Fernandez de la Vega, Ravi Kannan, and Marek Karpinski. Random sampling and approximation of max-csp problems. *Electron. Colloquium Comput. Complex.*, 2002.
- 4 Gunnar Andersson and Lars Engebretsen. Property testers for dense constraint satisfaction programs on finite domains. *Random Struct. Algorithms*, 21:14–32, 2002.
- 5 Anurag Anshu, David Gosset, and Karen J. Morenz. Beyond product state approximations for a quantum analogue of max cut. In *TQC*, 2020.
- 6 Itai Arad. A note about a partial no-go theorem for quantum pcp. *Quantum Inf. Comput.*, 11:1019–1027, 2011.
- 7 Sanjeev Arora, David R. Karger, and Marek Karpinski. Polynomial time approximation schemes for dense instances of np-hard problems. *Journal of Computer and System Sciences*, 58:193–210, 1999.
- 8 S. Baker. Approximation algorithms for np-complete problems on planar graphs. *Journal of the ACM*, 1994.
- 9 Nikhil Bansal, Sergey Bravyi, and Barbara M. Terhal. Classical approximation schemes for the ground-state energy of quantum and classical ising spin hamiltonians on planar graphs. *Quantum Inf. Comput.*, 9:701–720, 2009.
- 10 Boaz Barak, Prasad Raghavendra, and David Steurer. Rounding semidefinite programming hierarchies via global correlation. *2011 IEEE 52nd Annual Symposium on Foundations of Computer Science*, pages 472–481, 2011.
- 11 Alexander I. Barvinok. Combinatorics and complexity of partition functions. In *Algorithms and combinatorics*, 2016.
- 12 Anirban Basak and Sumit Mukherjee. Universality of the mean-field for the potts model. *Probability Theory and Related Fields*, 168:557–600, 2015.
- 13 Thiago Bergamaschi. Simulated quantum annealing is efficient on the spike hamiltonian. *arXiv:Quantum Physics*, 2020. [arXiv:2011.15094](https://arxiv.org/abs/2011.15094).
- 14 Christian Borgs, Jennifer T. Chayes, László Miklós Lovász, Vera T. Sós, and Katalin Vesztegombi. Convergent sequences of dense graphs ii. multiway cuts and statistical physics. *Annals of Mathematics*, 176:151–219, 2012.
- 15 Fernando G. S. L. Brandão and Aram Wettroth Harrow. Product-state approximations to quantum ground states. In *STOC '13*, 2013.
- 16 Sergey Bravyi, Anirban Narayan Chowdhury, David Gosset, and Pawel Wocjan. On the complexity of quantum partition functions. *arXiv*, abs/2110.15466, 2021. [arXiv:2110.15466](https://arxiv.org/abs/2110.15466).
- 17 Sergey Bravyi and David Gosset. Polynomial-time classical simulation of quantum ferromagnets. *Physical review letters*, 119 10:100503, 2017.
- 18 Sergey Bravyi, David Gosset, Robert Koenig, and Kristan Temme. Approximation algorithms for quantum many-body problems. *Journal of Mathematical Physics*, 2019.
- 19 Brielin Brown, Steven T. Flammia, and Norbert Schuch. Computational difficulty of computing the density of states. *Physical review letters*, 107 4:040501, 2011.

- 20 Elizabeth Crosson and Aram Wettroth Harrow. Rapid mixing of path integral monte carlo for 1d stoquastic hamiltonians. *Quantum*, 5:395, 2021.
- 21 Erik D. Demaine, Mohammad Taghi Hajiaghayi, and Ken-ichi Kawarabayashi. Algorithmic graph minor theory: Decomposition, approximation, and coloring. *46th Annual IEEE Symposium on Foundations of Computer Science (FOCS'05)*, pages 637–646, 2005.
- 22 Alan M. Frieze and Ravi Kannan. Quick approximation to matrices and applications. *Combinatorica*, 19:175–220, 1999.
- 23 Shayan Oveis Gharan and Luca Trevisan. A new regularity lemma and faster approximation algorithms for low threshold rank graphs. In *APPROX-RANDOM*, 2013.
- 24 Sevag Gharibian and Julia Kempe. Approximation algorithms for qma-complete problems. *2011 IEEE 26th Annual Conference on Computational Complexity*, pages 178–188, 2011.
- 25 Sevag Gharibian and Ojas Parekh. Almost optimal classical approximation algorithms for a quantum generalization of max-cut. *arXiv*, abs/1909.08846, 2019. [arXiv:1909.08846](https://arxiv.org/abs/1909.08846).
- 26 Oded Goldreich, Shafi Goldwasser, and Dana Ron. Property testing and its connection to learning and approximation. *Electron. Colloquium Comput. Complex.*, 3, 1998.
- 27 Aram Wettroth Harrow, Saeed Adel Mehraban, and Mehdi Soleimanifar. Classical algorithms, correlation decay, and complex zeros of partition functions of quantum many-body systems. *Proceedings of the 52nd Annual ACM SIGACT Symposium on Theory of Computing*, 2020.
- 28 Aram Wettroth Harrow and Ashley Montanaro. Extremal eigenvalues of local hamiltonians. *arXiv*, abs/1507.00739, 2015. [arXiv:1507.00739](https://arxiv.org/abs/1507.00739).
- 29 Matthew B. Hastings. Trivial low energy states for commuting hamiltonians, and the quantum pcg conjecture. *Quantum Inf. Comput.*, 13:393–429, 2013.
- 30 Vishesh Jain, Frederic Koehler, and Elchanan Mossel. The mean-field approximation: Information inequalities, algorithms, and complexity. In *COLT*, 2018.
- 31 Vishesh Jain, Frederic Koehler, and Elchanan Mossel. The vertex sample complexity of free energy is polynomial. In *COLT*, 2018.
- 32 Vishesh Jain, Frederic Koehler, and Andrej Risteski. Mean-field approximation, convex hierarchies, and the optimality of correlation rounding: a unified perspective. *Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing*, 2019.
- 33 Robbie King. An improved approximation algorithm for quantum max-cut. *arXiv*, abs/2209.02589, 2022. [arXiv:2209.02589](https://arxiv.org/abs/2209.02589).
- 34 Alexei Y. Kitaev, Alexander Shen, and Mikhail N. Vyalyi. Classical and quantum computation. In *Graduate studies in mathematics*, 2002.
- 35 Eun Jee Lee. Optimizing quantum circuit parameters via sdp. In *International Symposium on Algorithms and Computation*, 2022.
- 36 Ojas Parekh and Kevin Thompson. Application of the level-2 quantum lasserre hierarchy in quantum approximation algorithms. In *ICALP*, 2021.
- 37 Ojas Parekh and Kevin Thompson. Beating random assignment for approximating quantum 2-local hamiltonian problems. *arXiv*, abs/2012.12347, 2021. [arXiv:2012.12347](https://arxiv.org/abs/2012.12347).
- 38 Ojas Parekh and Kevin Thompson. An optimal product-state approximation for 2-local quantum hamiltonians with positive terms. *arXiv*, abs/2206.08342, 2022. [arXiv:2206.08342](https://arxiv.org/abs/2206.08342).
- 39 Prasad Raghavendra and Ning Tan. Approximating csps with global cardinality constraints using sdp hierarchies. In *SODA*, 2012.
- 40 Andrej Risteski. How to calculate partition functions using convex programming hierarchies: provable bounds for variational methods. *arXiv*, abs/1607.03183, 2016. [arXiv:1607.03183](https://arxiv.org/abs/1607.03183).
- 41 Neil Robertson and Paul D. Seymour. Graph minors. ii. algorithmic aspects of tree-width. *J. Algorithms*, 7:309–322, 1986.
- 42 Endre Szemerédi. Regular partitions of graphs. *Colloq. Internat. CNRS*, 260, 1975.

A The Hamiltonian Regularity Lemma

Let us begin by reviewing the cut decomposition of [22]. The key intuition behind their result is the notion that dense graphs can be roughly viewed as a sum of complete bipartite sub-graphs between subsets of vertices in the graph. Each of these bipartite sub-graphs is essentially a “cut” in the graph, hence the name.

► **Definition 15.** *Given two sets $S, T \subset [n]$ and a number $d \in \mathbb{R}$, the $n \times n$ cut matrix $D = \text{CUT}(S, T, d)$ is defined by $D_{u,v} = d \cdot \delta_{u \in S} \delta_{v \in T}$.*

► **Definition 16.** *A “cut decomposition” expresses a real matrix J as the sum*

$$J = \sum_{k=0}^s D^{(k)} + W \quad (14)$$

where each $D^{(k)}$ is a cut matrix defined on sets $R_k, L_k \subset [n]$, and of weight d_k . Such a decomposition is said to have width s , coefficient length $(\sum d_k^2)^{1/2}$, and error $\|W\|_{\infty \rightarrow 1}$.

The main result of [22] is precisely an algorithm to efficiently find such a decomposition:

► **Theorem 17** ([22]). *Let J be an arbitrary real matrix and fix a constant $\epsilon > 0$. Then there exists a cut decomposition of width $O(\epsilon^{-2})$, coefficient length $O(\|J\|_F/n)$, error at most $\epsilon n \|J\|_F$, and such that $\|W\|_F \leq \|J\|_F$. Moreover, with probability $1 - \delta$ said decomposition can be found implicitly in time $2^{\tilde{O}(\epsilon^{-2})}/\delta^2$, and explicitly in time $\tilde{O}(n^2/\epsilon^4) + 2^{\tilde{O}(\epsilon^{-2})}/\delta^2$.*

► **Remark 18.** The key point of the cut decomposition is that the number of cuts only depends on the quality of the approximation, not the size of the graph.

Perhaps the main tool we introduce in this work is a generalization of this result to the quantum setting. We exploit the fact that quantum density matrices and quantum Hamiltonians can be expressed in a Pauli basis, to reduce the problem of decomposing Hamiltonians into that of a “multi-colored” cut decomposition. For simplicity, here we discuss the case of 2-Local Hamiltonians, on qudits of local dimension $d = 2^{d'}$ which is a power of 2, and defer further generalizations to the appendix.

Let $H = \sum H_e$ be 2-local Hamiltonian defined on n qudits, and define $\mathbb{P}_{\log d} = \{\mathbb{I}, X, Y, Z\}^{\otimes \log d}$ be the set of Pauli operators acting on a single qudit. Any operator $h_{u,v}$ acting on the Hilbert space of 2 qudits can be decomposed into basis of $\mathbb{P}_{\log d} \otimes \mathbb{P}_{\log d}$:

$$H_{uv} = \sum_{i,j \in [d^2]} h_{uv}^{ij} \sigma_i^u \otimes \sigma_j^v \quad (15)$$

Where the h_{uv}^{ij} are all real coefficients. Group the coefficients of the interactions defined on the same Pauli matrices i, j into an interaction matrix $J^{ij} = \{h_{uv}^{ij}\}_{u,v}$, i.e., a matrix for each of d^4 “colors”. We note that this essentially defines $O(d^4)$ different weighted adjacency matrices. Now, let us apply the regularity lemma of [22] on each of the colored interaction/adjacency matrices J^{ij} above. By construction, for each pair (i, j) one can express

$$J^{ij} = \sum_{k=1}^s D^{ijk} + W^{ij} \equiv D^{ij} + W^{ij} \quad (16)$$

Where $D^{ijk} = \text{CUT}(R^{ijk}, L^{ijk}, d^{ijk})$ are the s cut matrices of the interaction $i, j \in [d^2]$, defined on partitions $\{R^{ijk}, L^{ijk}\}$ of the vertex set of the graph, and real constants d^{ijk} , for $k \in [s]$. We can thereby define the cut decomposition H_D of the Hamiltonian H to be the edges of the D^{ijk} crossing any such cut:

The Hamiltonian Cut Decomposition:
$$H_D = \frac{1}{2} \sum_{\substack{i,j \in [d^2] \\ k \in [s]}} \sum_{\substack{u \in R^{ijk} \\ v \neq u, v \in L^{ijk}}} D_{uv}^{ijk} \sigma_i^u \otimes \sigma_j^v \otimes \mathbb{I}_{V \setminus \{u,v\}} \quad (17)$$

where we appropriately order the tensor product such that $u < v$ and add a factor of $1/2$ via a handshaking argument. More importantly, we filter out the diagonal entries D_{uu}^{ijk} , since the cuts S, T returned by the cut decomposition in Theorem 17 need not be disjoint, and Local Hamiltonians can't have "self-edges" in a basis decomposition. While unfortunately we no longer can interpret the interaction graph of H_D as an exact sum of complete bipartite sub-Hamiltonians, fortunately, we will later recover this interpretation in an approximate sense.

We dedicate the rest of this section to proving two interesting properties of H_D . First, we argue that the energy of any product state $\rho = \otimes_{u \in V} \rho_u$ is close, whether in H or H_D , arising from the combinatorial structure of the decomposition. Then, we leverage our product state approximation toolkit, to argue that H_D is in fact close to H in the spectral norm $\|H - H_D\|_\infty$.

► **Theorem 19.** *Let $H = \sum_{u,v} H_{u,v}$ be a 2-Local Hamiltonian defined on qudits of local dimension $d = 2^{d'} = O(1)$, let $J_{uv} = \|H_{uv}\|_\infty$ be the matrix of interaction strengths, and let H_D be the Hamiltonian cut decomposition of H of width $s = O(\epsilon^{-2})$. Then, for all product states $\rho = \otimes_{u \in V} \rho_u$,*

$$|\text{Tr}[(H - H_D)\rho]| \leq \epsilon n \|J\|_F \quad (18)$$

Moreover, with probability $1 - \delta$ said decomposition can be found implicitly in time $2^{\tilde{O}(\epsilon^{-2})}/\delta^2$, and explicitly in time $\tilde{O}(n^2/\epsilon^4) + 2^{\tilde{O}(\epsilon^{-2})}/\delta^2$

Proof. By restricting our attention to product states, we are able to essentially decouple the "colors" (different Pauli terms) in the Cut Decomposition.

$$|\text{Tr}[(H - H_D)\rho]| = \left| \sum_{u < v} \sum_{i,j} (h_{uv}^{ij} - D_{uv}^{ij}) \text{Tr}[\sigma_u^i \otimes \sigma_v^j \rho] \right| = \quad (19)$$

$$= \left| \frac{1}{2} \sum_{i,j} \sum_{u \neq v} W_{uv}^{ij} \text{Tr}[\sigma_u^i \rho_u] \text{Tr}[\sigma_v^j \rho_v] \right| \leq \sum_{i,j} \left| \sum_{u \neq v} W_{uv}^{ij} \text{Tr}[\sigma_u^i \rho_u] \text{Tr}[\sigma_v^j \rho_v] \right| = \quad (20)$$

$$= \sum_{i,j} \left| \sum_v \left(\sum_{u: u \neq v} W_{uv}^{ij} \text{Tr}[\sigma_u^i \rho_u] \right) \text{Tr}[\sigma_v^j \rho_v] \right| \leq \sum_{i,j} \sum_v \left| \sum_u W_{uv}^{ij} \text{Tr}[\sigma_u^i \rho_u] \right| + \sum_{i,j} \sum_v |W_{vv}^{ij}| \leq \quad (21)$$

$$\leq \sum_{ij} \left(\|W^{ij}\|_{\infty \rightarrow 1} + n \cdot \max_v |W_{vv}^{ij}| \right), \quad (22)$$

where we re-introduced the diagonal terms to obtain the $\infty \rightarrow 1$ norm. From Theorem 17 we can pick a width $s = O(d^8 \epsilon^{-2}) = O(\epsilon^{-2})$ s.t. $\|W^{ij}\|_{\infty \rightarrow 1} \leq \epsilon n \|J^{ij}\|_F / d^4$. Finally, the original interaction graph has no diagonal elements ($J_{vv}^{ij} = 0$), and thus the Cauchy-Schwartz inequality tells us the diagonal entries of D^{ij} are bounded: $|W_{vv}^{ij}| = |J_{vv}^{ij} - D_{vv}^{ij}| \leq \sum_k |d^{ijk}| \leq s^{1/2} \cdot (\sum (d^{ijk})^2)^{1/2} \leq s^{1/2} \cdot \|J^{ij}\|_F / n$. The observation $\|J^{ij}\|_F \leq \|J\|_F$ and assuming $\epsilon^{-2} = o(n)$ concludes the proof. ◀

By combining the product state cut decomposition above with our results on product state approximations in theorem 3 and in section B of the full version, we can extend our results to entangled states as well.

20:18 Improved Product-State Approximation Algorithms for Quantum Local Hamiltonians

► **Lemma 20** (The Hamiltonian Weak Regularity Lemma). *In the context of Theorem 19, $\|H - H_D\| \leq \epsilon \cdot n \|J\|_F$.*

Proof. By Schatten norm duality, there exists a normalized state ψ^* s.t.

$$\|H - H_D\|_\infty = \max_{\psi} |\text{Tr}[(H - H_D)\psi]| = |\text{Tr}[(H - H_D)\psi^*]| \quad (23)$$

We now apply the product state approximation Theorem 3 on the state ψ^* and Hamiltonian $H' = H - H_D$, to argue there exists a separable state σ s.t.

$$|\text{Tr}[(H - H_D)(\psi^* - \sigma)]| \leq \epsilon n \|J\|_F / 2 \quad (24)$$

where we observe that if J' is the matrix of interaction strengths of $H' = H - H_D$, then $\|J'\|_1 \leq n \|J'\|_F$ (Cauchy-Schwartz) and $\|J'\|_F \leq \sum_{i,j \in [d^2]} \|W^{ij}\|_F \leq O(d^4 \|J\|_F)$ by means of a triangle inequality and the guarantees on W in Theorem 17. Since σ is separable, we can appropriately pick the width $s = O(\epsilon^{-2})$ in Theorem 19 to guarantee

$$|\text{Tr}[(H - H_D)\sigma]| \leq \epsilon n \|J\|_F / 2 \quad (25)$$

and thereby via the triangle inequality:

$$\|H - H_D\|_\infty \leq |\text{Tr}[(H - H_D)(\psi^* - \sigma)]| + |\text{Tr}[(H - H_D)\sigma]| \leq \epsilon n \|J\|_F \quad (26)$$

◀

Using the existing technology of matrix regularity lemmas, in the full version we present extensions to the result above for Local Hamiltonians defined on hyper-graphs and for graphs of low threshold rank.