Performance and limitations of the QAOA at constant levels on large sparse hypergraphs and spin glass models

Joao Basso  
Google Quantum AI  
Venice, CA, USA  
joao.basso@berkeley.edu

David Gamarnik  
MIT  
Cambridge, MA, USA  
gamarnik@mit.edu

Song Mei  
UC Berkeley  
Berkeley, CA, USA  
songmei@berkeley.edu

Leo Zhou  
Caltech  
Pasadena, CA, USA  
leozhou92@gmail.com

Abstract—The Quantum Approximate Optimization Algorithm (QAOA) is a general purpose quantum algorithm designed for combinatorial optimization. We analyze its expected performance and prove concentration properties at any constant level (number of layers) on ensembles of random combinatorial optimization problems in the infinite size limit. These ensembles include mixed spin models and Max-$q$-XORSAT on sparse random hypergraphs. Our analysis can be understood via a saddle-point approximation of a sum-over-paths integral. This is made rigorous by proving a generalization of the multilinear theorem, which is a technical result of independent interest. We then show that the performance of the QAOA at constant levels for the pure $q$-spin model matches asymptotically the ones for Max-$q$-XORSAT on random sparse Erdős-Rényi hypergraphs and every large-girth regular hypergraph. Through this correspondence, we establish that the average-case value produced by the QAOA at constant levels is bounded away from optimality for pure $q$-spin models when $q \geq 4$ and is even. This limitation gives a hardness of approximation result for quantum algorithms in a new regime where the whole graph is seen.

Index Terms—optimization; quantum computing; quantum algorithm; spin glass

I. INTRODUCTION

Quantum computers are widely believed to be more powerful than classical computers, in part due to Shor’s seminal quantum algorithm for solving the classically intractable problem of integer factorization [1]. As quantum computers begin to come online, an important open question is whether we can harness their power to achieve a computational advantage on optimization problems with widespread real-world applications. The Quantum Approximate Optimization Algorithm (QAOA) is a leading quantum algorithm designed to find approximate solutions of combinatorial optimization problems (COPs) [2]. The QAOA is computationally universal [3], and its generalizations can capture other powerful algorithms such as the quantum singular value transformation [4]. Although the QAOA can find the optimum when its level\(^1\) (number of layers) $p$ goes to infinity [2], we have a limited knowledge of its behavior at finite $p$ due to the challenges in analyzing quantum many-body dynamics on classical computers. Even at the lowest level $p = 1$, the QAOA has output distributions that cannot be efficiently simulated on any classical device under reasonable complexity-theoretic assumptions [5], similar to algorithms implemented in recent “quantum supremacy” experiments [6], [7]. Experimental tests of quantum optimization algorithms have largely been restricted to only modest-sized problems and short runtimes due to noisy quantum hardware limitations [8], although hints of a polynomial speedup over simulated annealing are recently observed in some cases [9]. Nevertheless, it is difficult to conclude from experimental observations whether there is a definitive quantum advantage for approximate optimization without a convincing picture of the quantum algorithms’ asymptotic behavior at large problem sizes and long runtimes.

To address this question, we are in need of rigorous theoretical studies of quantum optimization algorithms. Recent work have taken steps in this direction by analyzing the QAOA and obtaining provable performance guarantees. Early results look at MaxCut, where it was shown the QAOA at $p = 1$ has a guaranteed worst-case approximation ratio that beats random guessing but not the best known guarantee achieved by the classical semi-definite programming (SDP) algorithm [2]. Since the known methods for proving the worst-case guarantees require a computation that scales doubly exponentially with $p$ and are thus limited to extremely small $p$ [10], later results turn to the more tractable analysis of average-case performance. It was shown in [11] that the QAOA at $p = 11$ outperforms the standard SDP on typical instances of the Sherrington-Kirkpatrick spin glass problem. This result was extended to MaxCut on large random regular graphs in [12] where the QAOA beats the best unconditionally proven performance of any known classical algorithm. Nevertheless, pending a widely believed conjecture that these two problems exhibit no overlap gap property (OGP), an approximate message passing (AMP) algorithm [13], [14] can get arbitrarily close to the optimum.

Moreover, recent theoretical results show that the QAOA’s level $p$ needs to grow at least logarithmically with problem size $n$ for some COPs on graphs exhibiting locally tree-like

\(^1\)In some literature, $p$ is also referred to as the QAOA depth. Here we call $p$ as the QAOA’s level to avoid confusion with the quantum circuit-depth or runtime, which for the QAOA scales roughly as $p \times \text{(max graph degree)}$.\n
structures [15]–[18]. The practical relevance of this limitation on the QAOA is yet to be understood, and furthermore these results do not apply to models of graphs exhibiting full connectivity. Many classical algorithms including AMP are similarly limited, and provably fail to reach optimality if the problem exhibits OGP even when there is full connectivity. The possibility for a quantum advantage, however, even in the constant-\(p\) regime, was thus left open prior to this work.

In this paper, we analyze the power and limitation of the QAOA applied to general models of random COPs in the infinite size limit. An example we consider is the \(q\)-spin model, which describes ensembles of COPs with random all-to-all \(q\)-body couplings. This problem is provably difficult for many classical algorithms, including AMP [19], low-degree polynomials [20] and Boolean circuits [21], when \(q \geq 4\) and is even, because it exhibits the OGP. On the other hand, the power of quantum algorithms for this model is not known. Our first contribution is a formalism to calculate the average performance of the QAOA at any fixed level \(p\) in the infinite size limit for various ensembles with i.i.d. random couplings (Theorem 1), using a strong generalization of the multinomial theorem (Proposition IV.1). Although the proof of the latter result is mathematically involved, the result itself has a simple interpretation as a saddle-point approximation. We also establish concentration properties for these problems, where we show that measurement outputs from the QAOA applied to a typical instance concentrate at the calculated average. We note that these concentration results do not follow from general concentration bounds applicable to classical algorithms, and instead we establish the concentration property by showing vanishing variance. Our result substantially generalizes previous analyses, which were limited to either two-body couplings \((q=2)\) [11], or the lowest QAOA level \(p=1\) for arbitrary \(q\)-body couplings [22], [23].

For our second main result, we show that the performance of the QAOA on \(q\)-spin models matches asymptotically the one for Max-\(q\)-XORSAT on sparse random hypergraphs. This asymptotic equivalence was first shown via explicit formulas for \(q=2\) in [12], [23], and we generalize it to arbitrary \(q\) in the current paper (Theorem 2 and 3). Our proof method for the asymptotic equivalence of QAOA performance on these models differs from the approach usually employed in the classical context. Classical results on the equivalence of dense and sparse models, such as [24], [25], use Lindeberg’s argument or its variants to establish universality properties of free energy of random Hamiltonians. Unfortunately, Lindeberg’s argument appears powerless in our setting, and we use other methods to establish this correspondence. We leave it as an interesting challenge to develop general purpose methods establishing Lindeberg-type universality in the quantum setting.

Lastly, we show that the QAOA at any constant \(p\) cannot approximate arbitrarily well the ground state values of \(q\)-spin models in the average-case when \(q \geq 4\) and is even (Theorem 4). Previously, this limitation was only shown for some COPs on sparse hypergraphs, via arguments exploiting the OGP and the locality of the QAOA that prevents it from seeing the whole graph at sufficiently low depth [16], [18]. Importantly, for the fully connected \(q\)-spin models we consider, the locality-based arguments do not apply, and no limitation of this kind was known. Instead, our result is obtained from a “dense-from-sparse” reduction where we use the previous obstruction-by-OGP result on sparse random hypergraphs to prove limitation of the QAOA on dense spin glass models. Since we have shown the QAOA’s performance on the two types of models match asymptotically, its ability to find near optimums for \(q\)-spin models would contradict its failure to reach near optimality on sparse ones. This proof idea is unusual and novel: many results for sparse random hypergraphs are obtained by establishing them for complete graphs \((q\text{-spin})\) models and then employing the asymptotical equivalence of the two graph structures [24]–[27], including the OGP statement established in [27]. However, the usage of results for sparse random graphs in order to obtain ramification for complete graphs is new.

Our work clarifies paths forward in understanding quantum advantages in approximate optimization. Although it was previously known that local quantum algorithms such as the QAOA are limited in the low circuit-depth regime where they do not see the whole graph, our result shows for the first time that significant barriers remain even when the whole graph is seen. One natural path forward is to compare the energy achieved by the constant-\(p\) QAOA to that by the AMP algorithm [13], [28], which holds the current record on the \(q\)-spin models among polynomial-time classical algorithms [29]. It would be very interesting to see whether the QAOA can achieve a better energy than AMP. Moreover, there is currently no good methods for analyzing the QAOA when the level \(p\) grows faster than say, \(2 \log n\), even for sparse random hypergraphs. In this regime, none of the currently known limitations apply, and it remains an open question how fast \(p\) needs to grow to achieve arbitrarily good approximation.

II. BACKGROUND: THE QAOA, SPIN GLASSES, AND THE OVERLAP GAP PROPERTY

Review of the QAOA.— The QAOA is a quantum algorithm introduced by [2] for finding approximate solutions to combinatorial optimization problems. The goal is to maximize a cost function, which counts the number (or total weight) of clauses satisfied by an input bit string. Given a cost function \(C(z)\) on bit strings \(z \in \{+1, -1\}^n\), we can define a corresponding quantum operator \(C\), diagonal in the computational basis, as \(C(z) = C(z)|z\rangle\). Moreover, define the operator \(B = \sum_{j=1}^n X_j\), where \(X_j\) is the Pauli \(X\) operator acting on qubit \(j\). Given a set of parameters \(\gamma = (\gamma_1, \gamma_2, \ldots, \gamma_p) \in \mathbb{R}^p\) and \(\beta = (\beta_1, \beta_2, \ldots, \beta_p) \in \mathbb{R}^p\), the QAOA initializes the system of qubits in the state \(|s\rangle = 2^{-n/2} \sum_z |z\rangle\) and applies \(p\) alternating layers of unitary operations \(e^{-i \gamma_t C}\) and \(e^{-i \beta_t B}\) to prepare the state

\[
|\gamma, \beta\rangle = e^{-i \beta_1 B} e^{-i \gamma_1 C} \ldots e^{-i \beta_p B} e^{-i \gamma_p C} |s\rangle.
\]

(1)

For a given cost function \(C\), measuring \(|\gamma, \beta\rangle\) in the computational basis enough times will yield a bit string \(z\) whose
value $C(z)$ is near the quantum expectation $\langle \gamma, \beta | C | \gamma, \beta \rangle$ or better. Heuristic strategies have been proposed to optimize $\langle \gamma, \beta | C | \gamma, \beta \rangle$ with respect to parameters $(\gamma, \beta)$ using a good initial guess [30].

**Defining ensembles of random COPs.**— We consider a general combinatorial optimization problem (COP) on $n$ bits where each clause involves at most $q_{\text{max}}$ bits. This problem can be understood as maximizing a cost function over $z \in \{\pm 1\}^n$ that takes the form

$$
C_J(z) = \sum_{q=1}^{q_{\text{max}}} c_q \sum_{i_1, \ldots, i_q=1}^n J_{i_1i_2 \cdots i_q} z_{i_1} z_{i_2} \cdots z_{i_q} = c_1 \sum_{i=1}^n J_{i} z_i + c_2 \sum_{i,j=1}^n J_{i,j} z_i z_j + \cdots
$$

(2)

where each problem instance is specified by a choice of tensors $J = \{(J_{i_1, \ldots, i_q})_{i_1, \ldots, i_q \in [n]} : q \in [q_{\text{max}}]\}$. We study the application of the QAOA to an ensemble of random COPs that takes the above form. For example, we consider

- $G_{d, \alpha}(n)$ — mixed spin model. This ensemble is defined by any sequence of $c_q \in \mathbb{R}$ and randomly chosen $J_{i_1, \ldots, i_q} \sim_{\text{id}} \mathcal{N}(0, 1/n^{q-1})$ as normal distribution, for each $q \in [q_{\text{max}}]$ and $1 \leq i, \ldots, q \leq [n]$. $G_d(n)$ — pure $q$-spin model. This is a special case of the mixed spin model where $c_1 = 0$ and $c_q = 0$ for all $q' \neq q$. Note the Sherrington-Kirkpatrick (SK) model is the pure 2-spin model.

- $G_{d, \alpha}(R)$ — Max-$q$-XORSAT on a random Erdős-Rényi directed multi-hypergraph [18]. Here, a random directed multi-hypergraph on $n$ vertices is obtained by first choosing the number of edges $m \sim \text{Poisson}(dn)$, and then choosing hyperedges $e^1, e^2, \ldots, e^m$ i.i.d. uniformly at random from the set $[n]^q$ of all vertex $q$-tuples (some hyperedges could potentially be identical). Each hyperedge associates a random weight $w(e^j) \sim_{\text{Unif}} (\{-1/\sqrt{d}\})$. The cost function is defined as $C_{\text{ER}, q}(z) = \sum_{j=1}^{m} w(e^j) z_{i_1} \cdots z_{i_q}$. This model has an equivalent description in the form (2). More specifically, the cost function $C_{\text{ER}, q}(z)$ (as a function) has the same distribution as $C_J(z) = \sum_{i_1, \ldots, i_q=1}^n J_{i_1, \ldots, i_q} z_{i_1} \cdots z_{i_q}$, where $J_{i_1, \ldots, i_q} \sim_{\text{id}} \mathcal{N}(0, 1/n^{q-1})$, with $J_{i_1, \ldots, i_q} = J_{i_1, \ldots, i_q}$ and a sequence $\{\delta(d)\}_{d \geq 1}$ with the following property. For every $\epsilon > 0$ there exist sufficiently

there exists a constant $\eta_{\text{OPT}}(G_{d, \alpha}(\text{mixed}))$ which is the asymptotic optimum value of the associated optimization problem in the sense

$$
\lim_{n \to \infty} \max_{z \in \{\pm 1\}^n} C_J(z) = \eta_{\text{OPT}}(G_{d, \alpha}(\text{mixed})), \quad \text{in probability.}
$$

(3)

The existence of the limit (3) follows from a simple and clever subadditivity argument of Guerra and Toninelli [31], and the actual value of this limit was obtained as a result of a very impressive development starting from a non-rigorous physicist-style argument by Parisi [32], and then rigorously verified by Talagrand [33]. See also Panchenko [34] for a book reference for this and related results.

For the special case of pure $q$-spin model $J \sim G_q(n)$ (recall that this corresponds to $c_q = 1$ and $c_{q'} = 0$ for all $q' \neq q$ in $G_{d, \alpha}(\text{mixed})$), we denote the right hand side of (3) as $\gamma_{\text{OPT}}(G_q)$. It is known that $\gamma_{\text{OPT}}(G_q)$ also describes the ground energy density of random sparse model $G_{d, \alpha}(R)$ in the large degree limit:

$$
\lim_{d, \alpha \to \infty} \frac{1}{n} \mathbb{E}_{J \sim G_{d, \alpha}(R)} \max_{z \in \{\pm 1\}^n} C_J(z) = \gamma_{\text{OPT}}(G_q).
$$

(4)

This result was established first in [24] for the case of graphs (that is $q = 2$), and then extended to the case of Erdős-Rényi hypergraphs in [25], [27]. While the results were restricted to certain types of distributions of $J$, the proof approach developed in [25], which is based on the Lindeberg-type argument, reveals a universality property: the limit depends on the distribution only through the first and second moment, and furthermore applies to the setting of $J$ with non-zero mean upon centering.

**The overlap gap property and algorithmic thresholds.**— While the above results give us a statistical prediction of the typical-case optimal energy density $\gamma_{\text{OPT}}$ for these random COPs, they are non-constructive and thus do not yield efficient algorithms to find near-optimal solutions $z$ such that $C_J(z)/n \approx \gamma_{\text{OPT}}$. As it turns out, there is a provable obstacle preventing many algorithms to reach optimality in certain ensembles of problems, which is described as a property in the solution space geometry of the problem. This is the overlap gap property (OGP), which roughly says that for certain choices of the disorder $J$, specifically in the case of the pure $q$-spin model with $q \geq 4$ even, there is a gap in the set of possible pairwise overlaps of near-optimal solutions.

The use of OGP to show obstruction for quantum algorithms, specifically the QAOA, was initiated in [16] and subsequently extended in [18]. Both work prove limitation of local quantum algorithms when the COPs can be embedded on a sparse hypergraph. We now describe the result in [18] formally. This result forms the basis of our negative result on the limitation of the performance of the QAOA in the setting of the fully connected spin glass models.

**Theorem CLSS21** (Modified version of [18] Corollary 4.4). Let $J \sim G_{d, \alpha}(R)$. For every even $q \geq 4$, there exists $\eta_{\text{OGP}}(G_q) < \eta_{\text{OPT}}(G_q)$ and a sequence $\{\delta(d)\}_{d \geq 1}$ with the following property. For every $\epsilon > 0$ there exist sufficiently
large $d_0$ such that for every $d \geq d_0$, every $p \leq d(d) \log n$ and an arbitrary choice of parameters $\gamma, \beta$, with probability converging to 1 as $n \to \infty$, the performance of the QAOA at level $p$ satisfies $(\gamma, \beta|C_J/n|\gamma, \beta) \leq \eta_{\text{OGP}}(G) + \epsilon$.

The proof idea first introduced in [16] and then extended to the hypergraph setting in [18] uses the effective locality of the QAOA at logarithmic depth (or level) which prevents it from overcoming the OGP barrier. The implementation of this idea in the context of classical algorithm was introduced in [35] and then extended to a broad class of other classical algorithms; see [36] for a survey.

When considering COPs embedded on dense hypergraphs, however, such as the case of the model $G_{e, \text{mixed}}$ or $G_q$, the techniques in [16] and [18] do not apply since they rely crucially on the locality of the algorithm and the sparsity of the hypergraph, so that the whole graph is not seen by the QAOA at sufficiently low level. In contrast, for these dense models the QAOA sees the whole graph at any level. Thus, a new idea is needed to prove the obstruction in this non-local setting. This is one of the main goals of this paper, and it is achieved by a “dense-from-sparse” argument which uses an asymptotic equivalence of the algorithm’s performance on the dense $G_q$ and the sparse $G_{e, \text{mixed}}$ models. It is noteworthy that while there is a large literature on using the “sparse-from-dense” reduction for the purposes of establishing the results on sparse graphs from known results on dense graphs, such as [24]–[27], the converse direction we undertake in this paper is novel.

III. MAIN RESULTS

A. Performance of the QAOA on random COPs

Our first main result is a constructive method to evaluate the QAOA’s performance applied to a generic ensemble of random COPs in the form of (2) that satisfies the following assumption:

**Assumption 1.** For every $q \in \{q_{\text{max}}\}$, $J_1, \ldots, J_q$ are i.i.d. following some mean zero symmetric distribution with finite second moment. Assume that $E[e^{\lambda J_1 + \cdots + J_q}]$ is real positive for large $n$ and denote $g_{q,n}(\lambda) = n^{q-1} \log E[e^{\lambda J_1 + \cdots + J_q}]$. For any fixed $\lambda$, we assume that (1) $\lim_{n \to \infty} g_{q,n}(\lambda)/n = 0$; (2) $\lim_{n \to \infty} g_{q,n}(\lambda) = g_q(\lambda)$ exists and $g_q$ is differentiable; (3) $\lim_{n \to \infty} g_{q,n}(\lambda) = g_q(\lambda)$.

Note the ensembles $G_{e, \text{mixed}}, G_q, G_{e, \text{QR}}$ defined earlier satisfy the assumption above. In the theorem below, we establish the limiting performance of QAOA of at any constant level $p$ for any ensemble satisfying Assumption 1.

**Theorem 1.** Suppose $C_J$ is a random COP of form (2) drawn from an ensemble $G$ that satisfies Assumption 1. Then for any $p$ and any parameters $(\gamma, \beta) \in \mathbb{R}^{2p}$, we have

$$\lim_{n \to \infty} E_{J \sim G(n)} [\langle \gamma, \beta | C_J/n | \gamma, \beta \rangle] = V_p(G, \gamma, \beta)$$

and

$$\lim_{n \to \infty} E_{J \sim G(n)} [\langle \gamma, \beta | (C_J/n) | \gamma, \beta \rangle] = [V_p(G, \gamma, \beta)]^2,$$

where the limit $V_p(G, \gamma, \beta)$ has a formula that we define explicitly in (14). A corollary of the above theorem is that the value produced by the QAOA satisfy concentration properties. Specifically, it means that with probability tending to 1 as $n \to \infty$, measurements of the QAOA applied to a typical instance of the random COP ensemble will yield a string $z$ whose value $C_J(x)/n$ concentrate at the quantum expectation $\langle \gamma, \beta | C_J/n | \gamma, \beta \rangle$, which itself is close to the ensemble-averaged value $V_p(G, \gamma, \beta)$. This follows from the fact that the second moment is equal to the first moment squared in the $n \to \infty$ limit. To see this, note that

$$E_J[\langle \gamma, \beta | (C_J/n) | \gamma, \beta \rangle] - E_J^2[\langle \gamma, \beta | C_J/n | \gamma, \beta \rangle] = \text{Var}(\text{instance}) + \text{Var}(\text{measurement})$$

is the combined variance over instances and measurements, where

$$\text{Var}(\text{instance}) = E_J[\langle \gamma, \beta | C_J/n | \gamma, \beta \rangle^2] - E_J^2[\langle \gamma, \beta | C_J/n | \gamma, \beta \rangle],$$

$$\text{Var}(\text{measurement}) = E_J[\langle \gamma, \beta | (C_J/n)^2 | \gamma, \beta \rangle - \langle \gamma, \beta | C_J/n | \gamma, \beta \rangle^2].$$

Since both are non-negative, Theorem 1 implies both $\text{Var}(\text{instance})$ and $\text{Var}(\text{measurement})$ tend to 0 as $n \to \infty$. By Chebyshev’s inequality, this means the QAOA is concentrated over both instances and measurements.

**Formula for $V_p(G, \gamma, \beta)$.** We now describe our formula of $V_p$. First, we denote

$$A := \{a_1, a_2, \ldots, a_2, a_{-p}, \ldots, a_{-2}, a_{-1} : a_{ \pm j} \in \{\pm 1\}\}$$

as the set of 2$p$-bit strings. Given QAOA parameters $(\gamma, \beta) \in \mathbb{R}^{2p}$, we define for any $a \in A$

$$Q_a := \prod_{r=1}^{p} (\cos \beta_r)^{1+(a_r+a_{-r})/2} (\sin \beta_r)^{1-(a_r+a_{-r})/2} \times (i)^{(a_r-a_{-r})/2},$$

$$\Phi_a := \sum_{r=1}^{p} \gamma_r (a_r a_{r+1} \cdots a_p - a_{-p} \cdots a_{-r-1} a_{-r}).$$

We also denote $ab \in A$ as the bit-wise product of $a, b \in A$. Then we define $\{W_a\}_{a \in A}$ to be the unique solution to the following self-consistent equation: for all $a \in A$,

$$W_a = Q_a e^{\sum_{b=1}^{2p} g_q(\lambda)} e^{\Phi_a} W_{a_1} \cdots W_{a_{-1}}.$$
The performance of QAOA for the SK model at any constant level \( p \) has been previously derived in [11]. Subsequently, the QAOA’s performance at level \( p = 1 \) has been derived for the mixed spin model in [22] and sparse random hypergraph model in [23] respectively. Our Theorem 1 encompasses these results by deriving the performance for any such ensemble of COPs with i.i.d. couplings at any level \( p \). In particular it can be verified that, when \( c_2 = 1/\sqrt{2} \), \( c_0 = 0 \) for all \( q \neq 2 \), and \( g_2 = -x^2/2 \), our formula (14) coincides with the formula of the QAOA’s performance for the SK model in [11]. Our result also coincides with the ones in [22], [23] for various models at \( p = 1 \).

B. Equivalence of the performance of QAOA on dense and sparse graphs

As (4) states, the global optimum of the dense model \( G_q \) is asymptotically identical to the global optimum of the sparse model \( G_{d,q} \). This was established using Lindeberg’s universality type arguments in [24], [25]. It is conceivable that the performance of the QAOA will also be similar on these two models of disorder. We establish that a more general version of this universality is true at constant levels in the following theorem.

**Theorem 2 (Universality).** Let \( G_{d,q}(n) \) be a generic ensemble of COPs with only \( q \)-body couplings satisfying Assumption 1, with characteristic function \( g_{d,n}^{(d)}(\lambda) = n^{q-1} \log \mathbb{E}[e^{\lambda J^{(d)}_{d,n}}] \) for \( J^{(d)} \sim G_{d,q}(n) \). Moreover, suppose that \( \forall \lambda \in \mathbb{R} \),

\[
\lim_{d \to \infty} \lim_{n \to \infty} \left( g_{d,n}^{(d)}(\lambda), g_{d,q}^{(d)}(\lambda) \right) = \left( -\lambda^2/2, -\lambda \right). 
\]

Then, the asymptotic performance of the level-\( p \) QAOA on \( G_{d,q} \) is the same as on \( G_q \) (the pure q-spin model):

\[
V_p(G_q, \gamma, \beta) = \lim_{d \to \infty} V_p(G_{d,q}, \gamma, \beta),
\]

at any parameters \((\gamma, \beta)\). In particular, this applies to the sparse model \( G_{d,q}^{ER} \), i.e.,

\[
V_p(G_q, \gamma, \beta) = \lim_{d \to \infty} V_p(G_{d,q}^{ER}, \gamma, \beta).
\]

We remark that this identity is by explicit computation via the formula (14) above. Unfortunately, the Lindeberg type argument appears powerless in this setting.

Moreover, the performance of the QAOA for Max-\( q \)-XORSAT on large-girth \( d \)-regular hypergraphs was derived in [12] using an approach more direct than our Theorem 1. There, the authors computed the expected performance of QAOA which was shown to be identical among all \( d \)-regular \( q \)-uniform hypergraphs with girth > \( 2p + 1 \). They gave an explicit formula for the performance and denoted it as \( \nu_p^{(d)}(d, \gamma, \beta) \). It was also shown in [12] that the formula of \( \lim_{d \to \infty} \nu_p^{(d)}(d, \gamma, \beta) \) for large-girth regular graphs matches \( V_p(G_2, \gamma, \beta) \), the analogous formula for the SK model. We generalize this correspondence to arbitrary \( q \) in the following theorem, which shows that the QAOA’s performance for the \( q \)-spin model \( G_q \) is also equivalent to that for Max-\( q \)-XORSAT on any large girth \( d \)-regular hypergraphs in the \( d \to \infty \) limit.

**Theorem 3.** Let \( \nu_p^{(d)}(d, \gamma, \beta) \) be the performance of QAOA on any instance of Max-\( q \)-XORSAT on any \( d \)-regular \( q \)-uniform hypergraph with girth > \( 2p + 1 \) given in [12]. Then for any \( p \) and any parameters \((\gamma, \beta)\), we have

\[
V_p(G_q, \gamma, \beta) = \sqrt{2} \lim_{d \to \infty} \nu_p^{(d)}(d, \sqrt{q} \gamma, \beta). 
\]

We note that [12] provides a more succinct formula for \( \lim_{d \to \infty} \nu_p^{(d)}(d, \gamma, \beta) \), and has evaluated it up to \( p \leq 20 \) with an \( O(p^2 \rho) \)-time iteration on a classical computer. By the equality (18), this also gives a faster procedure to evaluate \( V_p(G_q, \gamma, \beta) \) than the \( O(\rho^{\max}) \)-time procedure for the more generic case described in Theorem 1.

**Remark III.2.** Although \( d \)-regular hypergraphs are similar to Erdős-Rényi hypergraphs, our Theorem 1 does not apply to \( d \)-regular hypergraphs since we need independence structure of the tensor \( J \) in the cost function \( C_J \) (c.f. Eq. (2)). On the other hand, the technique in [12] is algebraic and is specifically for \( d \)-regular hypergraphs, and their technique does not apply to \( q \)-spin models and Erdős-Rényi hypergraphs.

**C. Limitation of the QAOA on dense hypergraphs**

We now turn to our last main result, where we show that the QAOA’s performance is obstructed even in a regime when the whole graph is seen. This is in contrast to all known proven limitations of the QAOA that apply to sparse graphs when the QAOA does not see the whole graph [15]–[18]. Recall the value \( \eta_{\text{QGP}}(G_q) < \eta_{\text{OPT}}(G_q) \) from Theorem CLSS21.

**Theorem 4.** For any fixed \( p \), parameters \((\gamma, \beta)\), and any even \( q \geq 4 \), we have

\[
V_p(G_q, \gamma, \beta) = \lim_{n \to \infty} \mathbb{E}_{J \sim G_q(n)}[\langle \gamma, \beta \rangle C_J/n \gamma, \beta]\]

\[
\leq \eta_{\text{QGP}}(G_q). 
\]

This implies that constant-\( p \) QAOA is not able to find a near-global optimizer of the \( q \)-spin model when \( q \geq 4 \) and is even.

The proof of this obstruction theorem exploits the equivalence of the QAOA’s performance on dense and sparse hypergraphs established in Theorem 2 above, together with Theorem CLSS21 established in [18]. See Section IV-B where we give a short proof.
IV. TECHNICAL OVERVIEW

A. A generalized multinomial theorem motivated by the QAOA

We now explain the key technical idea behind this paper, where we provide a mathematical framework to study the performance of the QAOA for a general ensemble of random COPs. The goal is to evaluate the quantum expectation of the operator $C_J$ which yields the average value produced by the algorithm. Using techniques introduced in [11], we can insert complete sets of $Z$-basis states between unitary operations in (1) tracking the path of every qubit, and write this expectation explicitly for any ensemble $G$ satisfying Assumption 1 as the following sum over paths:

$$E_{J,G(n)} \left[ (\gamma, \beta) \right| \frac{C_J}{n} \right] = \sum_{\{n_a\}} \left( \frac{n}{n_a} \right) \prod_{a \in A} Q_a^{n_a} \times \exp \left[ \frac{n}{2} \sum_{q=1}^{g_{\text{max}}} \sum_{a_1, \ldots, a_q \in A} g_{q,n}(c_q \Phi_{a_1} \cdots \Phi_{a_q}) \frac{n_{a_1} \cdots n_{a_q}}{n^q} \right] \times \left( - \frac{i c_q}{\sum_{b_1, \ldots, b_q \in A} g_{q,n}^{t}(c_q \Phi_{b_1} \cdots \Phi_{b_q}) \frac{n_{b_1} \cdots n_{b_q}}{n^q} \right),$$

where $A = \{ \pm 1 \}^P$, $Q_a$ and $\Phi_a$ are defined earlier in (11) and (12), and the sum is over all sets of non-negative integers $\{n_a : a \in A\}$ that add up to $n$. Here, each $n_a$ counts the number of qubits whose path matches a given bit string $a$. We may also consider higher powers $(C_J/n)^k$ to obtain concentration properties of the algorithm, but we will focus on $k = 1$ here to explain the essentials.

While the above expression can be evaluated explicitly by summing over all $O(n!^{|A|}) = O(n^P)$ terms, this double exponential scaling quickly becomes intractable even for the modest case of $p = 2$. The fact that we have a polynomial of $\{n_a\}$ inside the exponential in (20) also prevents us from applying the multinomial theorem when its degree $g_{\text{max}} > 1$. In the $n \rightarrow \infty$ limit, one may be tempted to treat the $\{n_a\}$ as random variables from a multinomial distribution with $\{Q_a\}$ as probabilities, so that we can approximate the sum as a Gaussian integral and then apply Laplace’s method. However, since the $Q_a$’s are generally complex numbers, this approach does not apply.

To overcome this difficulty, our main technical contribution in this paper is a generalized multinomial theorem that enables evaluation of sums like the one in (20) in the $n \rightarrow \infty$ limit. We state the informal theorem here and defer its formal version to Proposition 4.1 (Formal) in the full version of this paper [37], where we also provide the full proof.

**Proposition IV.1** (Informal). Suppose we are given a finite set $A$ and a set of complex numbers $\{Q_a\}_{a \in A}$ where $\sum_{a \in A} Q_a = 1$. Also suppose $A$ has a subset $A_0$ such that $\{Q_a\}_{a \in A_0} \subseteq [0, 1]$, and the remaining elements in $A \setminus A_0$ can be decomposed into pairs of $(a, \bar{a})$ such that $Q_a + Q_{\bar{a}} = 0$. Then for any sequence of bounded-degree polynomials $f_n(\{\omega_a\}_{a \in A})$ and “well-played” (defined in Definition B.5 in [37]) polynomials $P_n(\{\omega_a\}_{a \in A})$ with $\lim_{n \rightarrow \infty} (f_n, P_n) = (f, P)$, we have

$$\lim_{n \rightarrow \infty} \sum_{\{n_a\}} \left( \frac{n}{n_a} \right) \left( \prod_{b \in A} Q_b^{n_b} \right) e^{n P_n(\{n_a/n\})} f_n(\{n_a/n\}) = f([W_a]),$$

where $\{W_a\}_{a \in A}$ is given as the unique solution to

$$W_a = Q_a \exp \left[ \frac{\partial P([W_b]_{b \in A})}{\partial W_a} \right], \quad \forall a \in A. \quad (21)$$

Although motivated by the desire to analyze the QAOA, this generalized multinomial theorem may be of independent interest for other endeavors. The proof is rather cumbersome, but the result is surprisingly consistent with the answer obtained from a simple but non-rigorous application of the saddle-point method. Specifically, in the $n \rightarrow \infty$ limit, one may define continuous variables $\omega_a := n_a/n$, and approximate the sum over paths in (21) as an integral:

$$\int \left( \prod_{a \in A} d\omega_a \right) e^{S(\{\omega_a\})} f(\{\omega_a\})$$

where $S(\{\omega_a\}) = -\sum_{a \in A} \omega_a \log(\omega_a/Q_a) + P(\{\omega_a\})$. Then (21) may be understood as a saddle-point approximation of the above integral as $n \rightarrow \infty$, where it is dominated by the saddle point of $S(\{\omega_a\})$ subject to the constraint that $\sum_a \omega_a = 1$. This saddle point turns out to be the unique solution to (22). See [37] for more details.

However, it is challenging to make this saddle-point approximation rigorous directly. Instead, we prove Proposition IV.1 by making use of the combinatorial structure that emerges in the summation when the coefficients of $P_n$ satisfy a property that we call “well-played”. This “well-played” property manifests after pairing up the variables $(n_a, n_{\bar{a}})$ associated with the cancelling pairs of complex numbers $(Q_a, \bar{Q}_a)$, and then transforming the polynomial $P_n(\{n_a/n\})$ into a “canonical representation” of the dual variables $t_a = n_a + n_{\bar{a}}$ and $d_a = n_a - n_{\bar{a}}$. In this canonical representation, we find that the limit as $n \rightarrow \infty$ exists if all the terms of $P_n$ are at least linear in $t_a$, along with some additional constraints. This property also enables (22) to be exactly solved with an iterative procedure and allows for an explicit evaluation of the limiting value (21).

B. Proof (sketches) of main theorems

In this section, we sketch the proofs of Theorems 1, 2 and 3, and provide the full proof of Theorem 4. The missing details are given in the full version of this paper [37].

Proof sketch of Theorem 1 (the QAOA’s performance on i.i.d. ensembles).— We start by deriving (20) using similar techniques as in [11]. In order to evaluate the more general cases considered in this paper, we show that the polynomial in the exponential of (20) satisfies the “well-played” property whenever $g_{q,n}(\lambda)$ is an even function. Then applying Proposition IV.1, we get the expected performance of the QAOA.
as in Theorem 1. The second moment is obtained similarly. In order for the proofs to go through easily, we require some simple technical conditions on $g_{d,n}$ (that is, on the distribution of $J$), as stated in Assumption 1.

**Proof sketch of Theorem 2 and Theorem 3 (dense-sparse correspondence).**— We apply the formula $V_p(G, \gamma, \beta)$ given in Theorem 1 to the pure $q$-spin model $G_q$ and a generic ensemble $G_{d,q}$ satisfying the stated assumptions. In particular, we show that the aforementioned Erdős–Rényi ensemble $G_{ER}$ can be transformed to an equivalent description in the form of $G_{d,q}$ using the Poisson splitting trick. The $V_p$ formulas for these ensembles are then shown to match asymptotically, yielding Theorem 2.

With the formula for $V_p(G_q, \gamma, \beta)$ in hand, Theorem 3 is then straightforwardly proved by algebraically transforming the formula for $\lim_{d, n \to \infty} V_p(G_{ER}, \gamma, \beta)$ given in [12] using the notations of this paper for all $q$. We explicitly show the two formulas match, similar to a proof in [12] which had obtained the analogous result at $q = 2$.

**Proof of Theorem 4 (limitation of the QAOA on $G_q$).**— Here we give the short but complete proof of Theorem 4, which is easily implied by Theorem 2 and Theorem CLSS21. Indeed, for any fixed $p \in \mathbb{Z}_{>0}$ and fixed $(\gamma, \beta)$, Theorem CLSS21 implies that

$$\lim_{d \to \infty} V_p(G_{d,q}^{ER}, \gamma, \beta) = \lim_{d \to \infty} \lim_{n \to \infty} \mathbb{E}_{J \sim G_{d,q}^{ER}(n)} [\langle \gamma, \beta | C J \rangle] \leq \eta_{OGP}(G_q).$$

(24)

In the equation above, we use the fact that $\langle \gamma, \beta | C J \rangle$ concentrates around its expectation which is implied by Theorem 1, and thus the high probability bound in Theorem CLSS21 extends to the expectation bound.

Furthermore, (17) in Theorem 2 gives

$$V_p(G_q, \gamma, \beta) = \lim_{d \to \infty} V_p(G_{d,q}^{ER}, \gamma, \beta).$$

Combining these two equations the proof is completed.

**V. DISCUSSION AND OUTLOOK**

In this paper we have considered the performance of the QAOA for the problem of finding a near ground state of spin glass models when the algorithm is applied at a level (number of layers) that does not grow with problem size. We have derived an analytic formula of the value produced by the quantum algorithm as a function of its parameters in the limit as the number of spins diverges to infinity. Using this formula we have established that this value is asymptotically the same for the pure $q$-spin model and for Max-$q$-XORSAT on a sparse random hypergraph model. This extends recent results for the case of 2-spin models at any level [11], [12], [23] and for the case of $q$-spin models at level 1 [22], [23]. We have also established a concentration result showing that this value is concentrated around the instance-independent average with high probability as the system size diverges to infinity.

Using this correspondence, we prove that the value produced by the QAOA is bounded away from optimality by a multiplicative constant for the case of $q$-spin models with $q \geq 4$ and even. This is obtained as a corollary of a recent result [18] that the value of the QAOA is bounded away from optimality when the algorithm is implemented on sparse random hypergraphs. The latter result relied on locality of the algorithm and was restricted to sparse hypergraphs, much like prior negative results [15]–[17] for the QAOA in regimes where it does not see the whole graph. In this paper, we extend the limitation to the $q$-spin models, where the QAOA sees the whole graph at any level.

Our proof approach for this limitation uses a novel idea of "dense-from-sparse" reduction. While many results in the past have used the "sparse-from-dense" reduction where properties of sparse random hypergraphs are established from the corresponding properties of the $q$-spin model, the reversed direction implemented in this paper is new.

There is a large scope of problems which remain open. Our method of proof for the concentration result, which follows [11], is rather unconventional and is based on explicitly computing the second moment of the value produced by the algorithm. This contrasts sharply with approaches in classical settings where concentration bounds follow rather directly by application of standard techniques such as McDiarmid’s or Azuma’s inequalities. We note these concentration inequalities give stronger (exponential) convergence than what can be obtained from our explicit calculation. The quantum setting considered in this paper prevents the implementation of the more standard methods, and in general the concentration properties in quantum systems represent a general scope of rather interesting open problems.

Similarly, the “dense-from-sparse” reduction in our paper is obtained from a rather bulky explicit calculation of the asymptotic performance of the QAOA. In the classical settings such equivalence results follow from a broader universality type argument based on Lindeberg’s approach. The direct application of Lindeberg’s argument to the quantum setting appears to fail, and finding a workable quantum counterpart for such universality argument is an interesting open problem.

It is surprising to us that the result of our complicated calculation can be understood simply as a saddle-point approximation. The latter is a tool commonly used in physics calculations, often non-rigorously, e.g., in Parisi’s formula of the SK model [32]. Nevertheless, rigorous verification of the saddle-point approximation’s predictions can sometimes require indirect and sophisticated methods, e.g., in Talagrand’s proof of the Parisi formula [33]. Here, our generalized multinomial theorem serves as an indirect proof that the saddle-point approximation gives correct predictions of the QAOA’s behavior for many spin glass models. Following the appearance of this work, [38] is able to directly apply the saddle-point method to analyze the QAOA for random $k$-SAT in certain regimes. It would be interesting to understand more broadly when the saddle-point method can be applied to yield simple and accurate analysis of quantum algorithms and many-body dynamics.

Although we have proven a limitation of the QAOA at any
constant level $p$, our work still leaves open a few possibilities of a quantum advantage in this regime. For example, it would be very interesting to compare the constant-$p$ QAOA's performance on the $q$-spin models to the state-of-the-art classical algorithm which is the AMP algorithm [13], [28]. This algorithm provably finds $(1 - \epsilon)$-approximate optimums when there is no OGP (conjectured for $q = 2$) after $p_{\text{AMP}}(\epsilon)$ number of iterations for any $\epsilon > 0$. Here $p_{\text{AMP}}(\epsilon)$ is a constant independent of problem size. Nevertheless, AMP faces an algorithmic threshold bounded away from optimality when $q \geq 4$ is even [19], the setting where OGP is known to exist. The maximum value achievable by AMP can be obtained numerically via an extended Parisi formula. Whether the QAOA can match or possibly even beat the performance of the AMP algorithm remains an interesting open challenge.

Another interesting challenge regards improving our analysis and obtaining explicit numerical values achieved by the constant-$p$ QAOA at large $p$. Presently, we only know explicit values up to $p \leq 20$ from [12] due to the $O(p^{2\Delta P})$-complexity of evaluating the current formula. Going beyond and obtaining these values at higher $p$ can shed light on the performance of the QAOA with the performance of the AMP algorithm, mentioned earlier.

Recently, it was shown in [29] that no algorithms satisfying an “overlap concentration property” can obtain a value better than AMP on the mixed $q$-spin models. This was done using a variant of the overlap gap property, called the branching-OGP. It would be interesting to see if this limitation extends to the QAOA at constant levels. This would imply in particular that the QAOA at constant levels does not surpass the value achieved by the AMP algorithm.

Our proof method is limited to the QAOA with a constant level $p$. It is of interest to extend it to the QAOA with $p$ that grows with problem size $n$. At the current stage we don’t have the techniques to approach this. Since the QAOA provably reaches optimality when no bound on $p$ is placed, it is in particular important to understand whether this can be achieved at $p$ which is only polynomially large, so that the QAOA remains within the class of polynomial-time algorithms. This would provide a definitive evidence of a quantum advantage in optimization.

ACKNOWLEDGMENTS

We thank Sergio Boixo, Edward Farhi, Sam Gutmann, Jarrod R. McClean, and Benjamin Villalonga for helpful comments. D.G. is supported in part by NSF grant DMS-2015517. S.M. is supported in part by NSF grant DMS-2210827.

REFERENCES

[1] P. W. Shor, “Algorithms for quantum computation: discrete logarithms and factoring,” in Proceedings of 35th Annual Symposium on Foundations of Computer Science (FOCS), 1994, pp. 124–134.

[2] E. Farhi, J. Goldstone, and S. Gutmann, “A Quantum Approximate Optimization Algorithm,” arXiv preprint arXiv:1411.4028, 2014. [Online]. Available: https://arxiv.org/abs/1411.4028

[3] S. Lloyd, “Quantum approximate optimization is computationally universal,” arXiv preprint arXiv:1812.11075, 2018. [Online]. Available: https://arxiv.org/abs/1812.11075

[4] S. Lloyd, B. T. Kiani, D. R. Arvidsson-Shukur, S. Bosch, G. De Palma, W. M. Kaminsky, Z.-W. Liu, and M. Marvian, “Hamiltonian singular value transformation and inverse block encoding,” arXiv preprint arXiv:2104.01410, 2021. [Online]. Available: https://arxiv.org/abs/2104.01410

[5] E. Farhi and A. W. Harrow, “Quantum supremacy through the Quantum Approximate Optimization Algorithm,” arXiv preprint arXiv:1602.07674, 2016. [Online]. Available: https://arxiv.org/abs/1602.07674

[6] F. Arute, K. Arya, R. Babbush, D. Bacon, J. C. Bardin, R. Barends, R. Biswas, S. Boixo, F. G. S. L. Brandao, D. A. Buell, B. Burkett, Y. Chen, Z. Chen, B. Chiari, R. Collins, W. Courtney, A. Dunsworth, E. Farhi, B. Foxen, A. Fowler, C. Gidney, M. Giustina, R. Graff, K. Guerin, S. Habecker, M. P. Harrigan, M. J. Hartmann, A. Ho, M. Hoffmann, T. Huang, T. S. Humble, S. V. Isakov, E. Jeffrey, Z. Jiang, D. Kafri, K. Kechedzhi, J. Kelly, P. V. Klimov, S. Knysy, A. Korotkov, F. Kostritsa, D. Landhuis, M. Lindmark, E. Lucero, D. Lyakh, S. Mandrà, J. R. McClean, M. McEwen, A. Megrant, X. Mi, K. Michielsen, M. Mohseni, J. Mutus, O. Naaman, M. Neeley, C. NEILL, M. Y. Niu, E. Osby, A. Petukhov, J. Platt, C. Quintana, E. G. Rieffel, P. Roushan, N. C. Rubin, D. Sank, K. J. Satzinger, V. Smelyanskiy, J. K. Sung, M. D. Tishovich, A. Vainsencher, B. Villalonga, T. White, Z. J. Yao, P. Yeh, A. Zalcman, H. Neven, and J. M. Martinis, “Quantum supremacy using a programmable superconducting processor,” Nature, vol. 545, no. 7687, pp. 505–510, 2019.

[7] H.-S. Zhong, H. Wang, Y.-H. Deng, M.-C. Chen, L.-C. Peng, Y.-H. Luo, J. Qin, D. Wu, X. Ding, Y. Hu, P. Hu, X.-Y. Yang, W.-J. Zhang, H. Li, Y. Li, X. Jiang, L. Gan, G. Yang, L. You, Z. Wang, L. Li, N.-L. Liu, C.-Y. Lu, and J.-W. Pan, “Quantum computation advantage using photons,” Science, vol. 370, no. 6523, pp. 1460–1463, 2020.

[8] M. P. Harrigan, K. J. Sung, M. Neeley, K. J. Satzinger, F. Arute, K. Arya, J. Atalaya, J. C. Bardin, R. Barends, S. Boixo, M. Broughton, B. B. Buckley, D. A. Buell, B. Burkett, N. Bushnell, Y. Chen, Z. Chen, B. Chiari, R. Collins, W. Courtney, S. Demura, A. Dunsworth, D. Eppens, A. Fowler, B. Foxen, C. Gidney, M. Giustina, R. Graff, S. Habecker, A. Ho, S. Hong, T. Huang, L. B. Ioffe, S. V. Isakov, E. Jeffrey, Z. Jiang, C. Jones, D. Kafri, K. Kechedzhi, J. Kelly, S. Kim, P. V. Klimov, A. N. Korotkov, F. Kostritsa, D. Landhuis, P. Latpve, M. Lindmark, M. Leib, O. Martin, J. M. Martinis, J. R. McClean, M. McEwen, A. Megrant, X. Mi, M. Mohseni, W. Mruczkiewicz, J. Mutus, O. Naaman, C. Neill, F. Neukart, M. Y. Niu, T. E. O’Brien, B. O’Gorman, E. Osby, A. Petukhov, H. Putterman, C. Quintana, P. Roushan, N. C. Rubin, D. Sank, A. Skoltik, V. Smelyanskiy, D. Strain, M. Tame, Z. J. Yao, P. A. Vainsencher, T. White, Z. J. Yao, P. Yeh, A. Zalcman, L. Zhou, H. Neven, D. Bacon, E. Lucero, E. Farhi, and R. Babbush, “Quantum approximate optimization of non-planar graph problems on a superconducting processor,” Nature Physics, vol. 17, no. 3, pp. 332–336, 2021.
[14] A. E. Alaoui, A. Montanari, and M. Sellke, “Local algorithms for Maximum Cut and Minimum Bisection on locally tree-like regular graphs of large degree,” arXiv preprint arXiv:2111.06813, 2021. [Online]. Available: https://arxiv.org/abs/2111.06813

[15] S. Bravyi, A. Kliessch, R. Koening, and E. Tang, “Obstacles to variational quantum optimization from symmetry protection,” Physical review letters, vol. 125, no. 26, p. 260505, 2020.

[16] E. Farhi, D. Gamarnik, and S. Gutmann, “The Quantum Approximate Optimization Algorithm Needs to See the Whole Graph: A Typical Case,” arXiv preprint arXiv:2004.09002, 2020. [Online]. Available: https://arxiv.org/abs/2004.09002

[17] ——, “The Quantum Approximate Optimization Algorithm Needs to See the Whole Graph: Worst Case Examples,” arXiv preprint arXiv:2005.08747, 2020. [Online]. Available: https://arxiv.org/abs/2005.08747

[18] C.-N. Chou, P. J. Love, J. S. Sandhu, and J. Shi, “Limitations of Local Quantum Algorithms on Random Max-k-XOR and Beyond,” arXiv preprint arXiv:2108.06049, 2021. [Online]. Available: https://arxiv.org/abs/2108.06049

[19] D. Gamarnik and A. Jagannath, “The overlap gap property and approximate message passing algorithms for p-spin models,” The Annals of Probability, vol. 49, no. 1, pp. 180–205, 2021.

[20] D. Gamarnik, A. Jagannath, and A. S. Wein, “Low-degree hardness of random optimization problems,” in Proceedings of 61st Annual Symposium on Foundations of Computer Science (FOCS), 2020, pp. 131–140.

[21] ——, “Circuit lower bounds for the p-spin optimization problem,” arXiv preprint arXiv:2109.01342, 2021. [Online]. Available: https://arxiv.org/abs/2109.01342

[22] J. Claes and W. v. Dam, “Instance Independence of Single Layer Quantum Approximate Optimization Algorithm on Mixed-Spin Models at Infinite Size,” Quantum, vol. 5, p. 542, Sep 2021. [Online]. Available: http://dx.doi.org/10.22331/q-2021-09-15-542

[23] S. Boulebnane and A. Montanaro, “Predicting parameters for the Quantum Approximate Optimization Algorithm for MAX-CUT from the infinite-size limit,” arXiv preprint arXiv:2110.10685, 2021. [Online]. Available: https://arxiv.org/abs/2110.10685

[24] A. Dembo, A. Montanari, and S. Sen, “Extremal cuts of sparse random graphs,” The Annals of Probability, vol. 45, no. 2, pp. 1190–1217, 2017. [Online]. Available: https://arxiv.org/abs/1503.03923

[25] S. Sen, “Optimization on sparse random hypergraphs and spin glasses,” Random Structures & Algorithms, vol. 53, no. 3, pp. 504–536, 2018.

[26] D. Panchenko, “On the k-sat model with large number of clauses,” Random Structures & Algorithms, vol. 52, no. 3, pp. 536–542, 2018.

[27] W.-K. Chen, D. Gamarnik, D. Panchenko, and M. Rahman, “Suboptimality of local algorithms for a class of max-cut problems,” The Annals of Probability, vol. 47, no. 3, pp. 1587–1618, 2019.

[28] A. E. Alaoui, A. Montanari, and M. Sellke, “Optimization of mean-field spin glasses,” The Annals of Probability, vol. 49, no. 6, pp. 2922–2960, 2021.

[29] B. Huang and M. Sellke, “Tight lipschitz hardness for optimizing mean field spin glasses,” arXiv preprint arXiv:2110.07847, 2021. [Online]. Available: https://arxiv.org/abs/2110.07847

[30] L. Zhou, S.-T. Wang, S. Choi, H. Pichler, and M. D. Lukin, “Quantum Approximate Optimization Algorithm: Performance, Mechanism, and Implementation on Near-Term Devices.” Phys. Rev. X, vol. 10, p. 021067, 2020. [Online]. Available: https://arxiv.org/abs/1812.01041

[31] F. Guerra and F. Toninelli, “The thermodynamic limit in mean field spin glass models,” Commun. Math. Phys., vol. 230, pp. 71–79, 2002.

[32] G. Parisi, “A sequence of approximated solutions to the SK model for spin glasses,” Journal of Physics A: Mathematical and General, vol. 13, no. 4, p. L115, 1980. [Online]. Available: https://doi.org/10.1088/0305-4470/13/4/009

[33] M. Talagrand, “The Parisi formula,” Annals of mathematics, pp. 221–263, 2006.

[34] D. Panchenko, The Sherrington-Kirkpatrick model. Springer Science & Business Media, 2013. [Online]. Available: https://doi.org/10.1007/978-1-4614-6289-7

[35] D. Gamarnik and M. Sudan, “Limits of local algorithms over sparse random graphs,” in Proceedings of the 5th Conference on Innovations in Theoretical Computer Science (ITCS), 2014, pp. 369–376. [Online]. Available: https://doi.org/10.1145/2554797.2554831

[36] D. Gamarnik, “The overlap gap property: A topological barrier to optimizing over random structures,” Proceedings of the National Academy of Sciences, vol. 118, no. 41, 2021. [Online]. Available: https://arxiv.org/abs/2109.14409

[37] J. Basso, D. Gamarnik, S. Mei, and L. Zhou, “Performance and limitations of the QAOA at constant levels on large sparse hypergraphs and spin glass models,” arXiv preprint arXiv:2204.10306, 2022. [Online]. Available: https://arxiv.org/abs/2204.10306

[38] S. Boulebnane and A. Montanaro, “Solving boolean satisfiability problems with the quantum approximate optimization algorithm,” arXiv preprint arXiv:2208.06909, 2022. [Online]. Available: https://arxiv.org/abs/2208.06909