

# Instance Independence of Single Layer Quantum Approximate Optimization Algorithm on Mixed-Spin Models at Infinite Size

Jahan Claes<sup>1,2</sup> and Wim van Dam<sup>1,3,4</sup>

<sup>1</sup>QC Ware Corporation, Palo Alto, CA USA

<sup>2</sup>Department of Physics and Institute for Condensed Matter Theory, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA

<sup>3</sup>Department of Computer Science, University of California, Santa Barbara, CA USA

<sup>4</sup>Department of Physics, University of California, Santa Barbara, CA USA

This paper studies the application of the Quantum Approximate Optimization Algorithm (QAOA) to spin-glass models with random multi-body couplings in the limit of a large number of spins. We show that for such mixed-spin models the performance of depth 1 QAOA is independent of the specific instance in the limit of infinite sized systems and we give an explicit formula for the expected performance. We also give explicit expressions for the higher moments of the expected energy, thereby proving that the expected performance of QAOA concentrates.

arXiv:2102.12043v3 [quant-ph] 7 Sep 2021

## Contents

<b>1</b>	<b>Summary</b>	<b>1</b>
<b>2</b>	<b>Preliminaries and Notation</b>	<b>2</b>
<b>3</b>	<b>Current Results and Prior Work</b>	<b>3</b>
3.1	Our Results on Mixed-Spin Sherrington-Kirkpatrick Model . . . . .	3
3.2	Numerical Results and Relation to Prior Results . . . . .	4
<b>4</b>	<b>Derivation of Main Result</b>	<b>6</b>
4.1	Moment Generating Function . . . . .	7
4.2	Expectation when Couplings are Normal Distributed Variables . . . . .	7
4.3	General Form of Moments . . . . .	8
4.4	Evaluating Sums Over the Sketches $(n_{++}, n_{+-}, n_{-+}, n_{--})$ . . . . .	9
4.5	Large $n$ Limit of $A_t^a$ , $B_t^b$ , and $T_\xi^{ab}$ . . . . .	9
4.6	Large $n$ properties of $f_q$ . . . . .	10
4.7	Evaluating the First Moment . . . . .	11
4.8	Evaluating the Second Moment . . . . .	11
4.9	Evaluating Higher Moments . . . . .	11
4.10	Alternative expression for $\mathbb{E}[\langle H/n \rangle]$ . . . . .	12
<b>5</b>	<b>Discussion and Conclusion</b>	<b>12</b>
<b>1</b>	<b>Summary</b>	

The Quantum Approximate Optimization Algorithm (QAOA) is a variational quantum algorithm designed to give approximate solutions to unconstrained binary optimization problems [1]. While QAOA can be proven to give the optimal answer in the limit where the number of QAOA layers  $p$  goes to infinity, rigorous results on the performance of QAOA with finite  $p$  are difficult to obtain. In a recent paper, Farhi et al. [2] studied the application of the QAOA to the Sherrington-Kirkpatrick (SK) model, a spin-glass model with random all-to-all two-body couplings, in the limit of a large number of spins. Their paper demonstrated that for fixed  $p$ , the performance of the QAOA is independent of the specific instance of the SK model and can be predicted by explicit formulas. The paper also showed that the approximation ratio of the QAOA at  $p = 11$  outperforms a large class of classical optimization algorithms (although not the best classical algorithm [3]). In the current paper, we generalize the result of Farhi et al. to mixed-spin SK models, which extends the two-body couplings of standard SK to random all-to-all  $q$ -body couplings. We demonstrate that for  $p = 1$ , the performance of the QAOA is again independent of the specific instance, and we provide an explicit formula for the expected performance. Our work provides a potential avenue to demonstrating the advantage of QAOA over classical algorithms, as the best known classical algorithms for mixed-spin SK models have an approximation ratio that is bounded away from 1 [4, 5].

Jahan Claes: [jclaes2@illinois.edu](mailto:jclaes2@illinois.edu)

## 2 Preliminaries and Notation

The Quantum Approximate Optimization Algorithm (QAOA) [1] is a heuristic quantum algorithm for binary optimization. Given a cost function of  $n$  binary variables (spins)  $H(z_1, \dots, z_n)$ , QAOA seeks to produce a string  $z := (z_1, \dots, z_n)$  close to the minimum of  $H$ . Commonly we view  $H$  as a Hamiltonian operator that is diagonal in the  $Z$ -basis. A depth- $p$  QAOA circuit then consists of  $p$  repetitions of alternatively applying the Hamiltonian  $H$  and the mixing Hamiltonian  $B = X_1 + \dots + X_n$  to a uniform superposition as initial state, which is the product of  $+X$  single particle eigenstates. Explicitly, the depth- $p$  QAOA state is given by

$$|\beta_1, \dots, \beta_p; \gamma_1, \dots, \gamma_p\rangle := e^{-i\beta_p B} e^{-i\gamma_p H} \dots e^{-i\beta_1 B} e^{-i\gamma_1 H} \cdot \frac{1}{\sqrt{2^n}} \sum_{z \in \{\pm 1\}^n} |z\rangle. \quad (1)$$

A depth- $p$  QAOA circuit is parameterized by the  $2p$  angles  $\{\gamma_i\}, \{\beta_i\}$ . For a given problem, these angles should be optimized so that measuring in the  $Z$ -basis gives strings that make  $H$  as small as possible. In practice, this is typically done by minimizing the expectation value of the energy

$$\langle H \rangle := \langle \beta_1, \dots, \beta_p; \gamma_1, \dots, \gamma_p | H | \beta_1, \dots, \beta_p; \gamma_1, \dots, \gamma_p \rangle. \quad (2)$$

For some problems, this minimization may be done analytically on a classical computer [1, 6–8]. Otherwise, the minimization can be performed by running the QAOA on a quantum computer repeatedly for a fixed set of angles, estimating the expectation value, and updating the angles according to classical minimization algorithms [1, 9, 10]. We note that minimizing the energy expectation value is only one possible definition of “best” angles; in general minimizing the expectation value and maximizing the probability of finding the optimal  $z$  (or maximizing the probability of  $H(z_1, \dots, z_n)$  falling below a certain threshold) do not coincide.

It was recently demonstrated in [11] that in local optimization problems with cost functions drawn from realistic random distributions (e.g. MAXCUT on random 3-regular graphs), the expectation value per spin is *instance independent* as  $n \rightarrow \infty$ . That is, for fixed angles  $\{\gamma_i\}$  and  $\{\beta_i\}$   $\langle H/n \rangle$  is the same for all problem instances. This implies that the angles  $\{\gamma_i\}$  and  $\{\beta_i\}$  do not need to be optimized for every problem instance, but can be optimized once and reused for every problem drawn from the same distribution. The methods of [11] can also be used to derive *concentration of measure* results for local optimization problems. (While [11] did not explicitly address concentration of measure, it can be easily derived from their methodology.) That is, in the limit as  $n \rightarrow \infty$ , the variance in the energy per spin goes to zero:

$$\langle (H/n)^2 \rangle - \langle H/n \rangle^2 \rightarrow 0 \quad (3)$$

Concentration of measure means that for large  $n$ , every measurement of a QAOA state in the  $Z$ -basis gives strings with the same energy per spin. In total, instance independence implies the QAOA angles do not need to be optimized from instance to instance in order to minimize the expectation value, and concentration of measure implies that expectation value is the correct measure of the “best” angles.

While instance independence (and concentration of measure) were initially derived for local cost functions, similar results have also been derived for the Sherrington-Kirkpatrick (SK) model [12], a physics-inspired optimization problem with cost function

$$H = \sum_{1 \leq j < k \leq n} J_{jk} Z_j Z_k \quad (4)$$

with the  $J_{jk}$  are independently drawn from a Gaussian distribution with mean 0. The SK model is the “most nonlocal” two-body cost function, and serves as a model for realistic nonlocal two-body optimization problems. In the limit  $n \rightarrow \infty$ , the ground state energy per spin is known to be independent of the instance and can be computed explicitly [13, 14]. Recently, Montanari derived a classical algorithm that produces strings  $z$  with energy within  $(1 - \epsilon)$  times the optimum; this method has a complexity of  $C(\epsilon)n^2$ , where  $C(\epsilon)$  is a polynomial in  $1/\epsilon$  [3].

Ref. [2] proved both instance independence and concentration of measure for depth- $p$  QAOA applied to the SK model. In addition, [2] provided an explicit formula for  $\langle H/n \rangle$  in the limit  $n \rightarrow \infty$  for  $p = 1$ , and provided a computer algorithm to generate  $\langle H/n \rangle$  for any fixed depth  $p > 1$ . Therefore, the QAOA angles for the SK model can be chosen on a classical computer, and there are fixed performance guarantees in the limit  $n \rightarrow \infty$ . Ref. [2] demonstrated that at  $p = 11$ , QAOA outperforms semidefinite programming for the SK model, but could not show that the QAOA matches the performance of the Montanari algorithm.

In this work, we study a generalization of the SK model, the mixed-spin SK model, that allows for polynomials of degree  $d$  in the binary variables instead of only degree-2 terms [15, 16]. This can serve as a model for nonlocal optimization problems with higher-order terms. The mixed-spin SK model is also known to have a ground state energy per spin that is independent of the instance and can be computed explicitly [15–18]. For a mixed-spin SK model with degree  $d = 3$ , the generalization of the Montanari algorithm [4] approaches a fixed approximation ratio of  $\sim (0.9843 \pm 0.0003)$  times the optimum value, rather than the optimal value [5]. Thus, the mixed-spin SK model is a potential avenue for establishing the advantage of the QAOA over classical algorithms.

In this paper, we generalize the work of [2] to prove instance independence and concentration of measure for  $p = 1$  QAOA applied to mixed-spin SK models. As part of our work, we derive an explicit formula for  $\langle H/n \rangle$  in the limit  $n \rightarrow \infty$ , implying that the QAOA angles for the mixed-spin SK models can be chosen on a classical computer. Our work can likely be generalized to depth  $p \geq 1$  using the same methods as [2].

### 3 Current Results and Prior Work

#### 3.1 Our Results on Mixed-Spin Sherrington-Kirkpatrick Model

The mixed-spin SK model (often called the mixed  $p$ -spin model, although we will not use this terminology to avoid confusion with the  $p$  of QAOA) is given by [15, 16]

$$\begin{aligned}
 H &= \sum_{j=1}^n J_j Z_j + \frac{1}{\sqrt{n}} \sum_{1 \leq j < k \leq n} J_{jk} Z_j Z_k + \frac{1}{n} \sum_{1 \leq j < k < \ell \leq n} J_{jkl} Z_j Z_k Z_\ell + \cdots + \frac{1}{n^{\frac{d-1}{2}}} \sum_{1 \leq j_1 < \cdots < j_d \leq n} J_{j_1 \cdots j_d} Z_{j_1} \cdots Z_{j_d} \\
 &= \sum_{q=1}^d n^{\frac{1-q}{2}} \sum_{\substack{S \subseteq \{1, \dots, n\} \\ |S|=q}} J_S Z_S
 \end{aligned} \tag{5}$$

where in the last line we denote the product  $\prod_{i \in S} Z_i$  by  $Z_S$ . We assume each  $J_S$  is sampled independently from a Gaussian distribution  $N(0, \sigma_{|S|}^2)$  that only depends on  $|S|$ , and we will let  $\sigma_q$  be the standard deviation of the coupling constants  $J_S$  with  $|S| = q$ , for  $q = 1, \dots, d$ .

The ground state of the mixed-spin SK Hamiltonian is known to have a fixed energy per spin in the limit  $n \rightarrow \infty$  [18]. In fact we can also allow arbitrarily high orders in the mixed-spin SK model ( $d = \infty$ ), provided the variances decrease quickly enough to make  $\sum_q 2^q \sigma_q^2$  finite, and the ground state model will still have a fixed energy per spin as  $n \rightarrow \infty$  [18]. However, for simplicity we consider some finite bound  $d$  on the degree.

Our main result is as follows. Define the  $n$ -spin model by Eq. 5 with  $J_S \sim N(0, \sigma_{|S|}^2)$ . Then, using depth 1 QAOA with angles  $\beta, \gamma$ , the expectation of the energy per spin equals in the large  $n$  limit is given by

$$\lim_{n \rightarrow \infty} \mathbb{E}_{J \sim N} [\langle H/n \rangle] = 2\gamma \sum_{q=1}^d \sigma_q^2 \sum_{\substack{a \text{ odd} \\ a \leq q}} \frac{(-1)^{\frac{a-1}{2}}}{a!(q-a)!} \exp\left(-2\gamma^2 a \sum_{q'=1}^d \frac{\sigma_{q'}^2}{(q'-1)!}\right) \sin^a(2\beta) \cos^{q-a}(2\beta), \tag{6}$$

If we define the variables  $c_q := \sigma_q/\sqrt{q!}$  and the polynomial  $\xi(x) := \sum_{q=1}^d c_q^2 x^q$ , as is common in discussions of the mixed-spin model (see Section 3.2), we can also write this as

$$\lim_{n \rightarrow \infty} \mathbb{E}_{J \sim N} [\langle H/n \rangle] = 2\gamma \Im \left[ \xi \left( \cos(2\beta) + i \sin(2\beta) \exp(-2\gamma^2 \xi'(1)) \right) \right]. \tag{7}$$

Furthermore, the second moment of the energy per spin equals

$$\lim_{n \rightarrow \infty} \mathbb{E}_{J \sim N} [\langle (H/n)^2 \rangle] = \lim_{n \rightarrow \infty} \mathbb{E}_{J \sim N} [\langle H/n \rangle]^2. \tag{8}$$

Our second result, Eq. 8, allows us to prove that  $p = 1$  QAOA applied to the mixed-spin SK model has both concentration of measure and instance independence. To see this, we note that we may write

$$\mathbb{E}_{J \sim N} [\langle (H/n)^2 \rangle] - \mathbb{E}_{J \sim N} [\langle H/n \rangle]^2 = \left( \mathbb{E}_{J \sim N} [\langle (H/n)^2 \rangle - \langle H/n \rangle^2] \right) + \left( \mathbb{E}_{J \sim N} [\langle H/n \rangle^2] - \mathbb{E}_{J \sim N} [\langle H/n \rangle]^2 \right) \tag{9}$$

When applying QAOA to mixed-spin SK models, the measured value of  $(H/n)$  varies for two reasons. First, it varies because the bonds  $J_S$  vary from instance to instance (the  $\mathbb{E}_J[\cdot]$  expectation). Second, it varies because the QAOA state  $|\gamma, \beta\rangle$  is not an eigenstate of the  $Z$ -operators, so that the measurement outcomes have randomness even for fixed  $J_S$  (the  $\langle \cdot \rangle$  expectation). The left hand side of Eq. 9 represents the total variance in  $(H/n)$  due to both sources of randomness. The right hand side demonstrates that the total variance can be decomposed into two terms. The first term is the average over  $J_S$  of the variance due only to the measurement randomness. The second term is the variance in the expected value  $\langle H/n \rangle$  due to the randomness in the bonds  $J_S$ . Since both of these terms are non-negative, they must both tend to zero as  $n \rightarrow \infty$  as well.

The fact that the first variance approaches zero gives us concentration of measure: it says that for typical couplings  $J_S$  the variance in the measurement outcomes vanishes, and thus we always measure a string with energy per spin equal to the expectation value (note that the term inside the  $\mathbb{E}[\cdot]$  is always positive, so that for the average over  $J_S$  to go to zero the magnitude of the typical value must also go to zero). The second term approaching zero clearly gives instance independence of the expectation value, since it shows that the variance in the expectation value due to different couplings vanishes.

Finally, we note that the methods we use also suffice to derive a formula for all higher moments of the energy per spin

$$\lim_{n \rightarrow \infty} \mathbb{E}_{J \sim N} [\langle (H/n)^m \rangle] = \lim_{n \rightarrow \infty} \mathbb{E}_{J \sim N} [\langle H/n \rangle]^m, \tag{10}$$

although unlike the  $m = 2$  result, the  $m > 2$  result does not have any obvious implication for the performance of the QAOA.

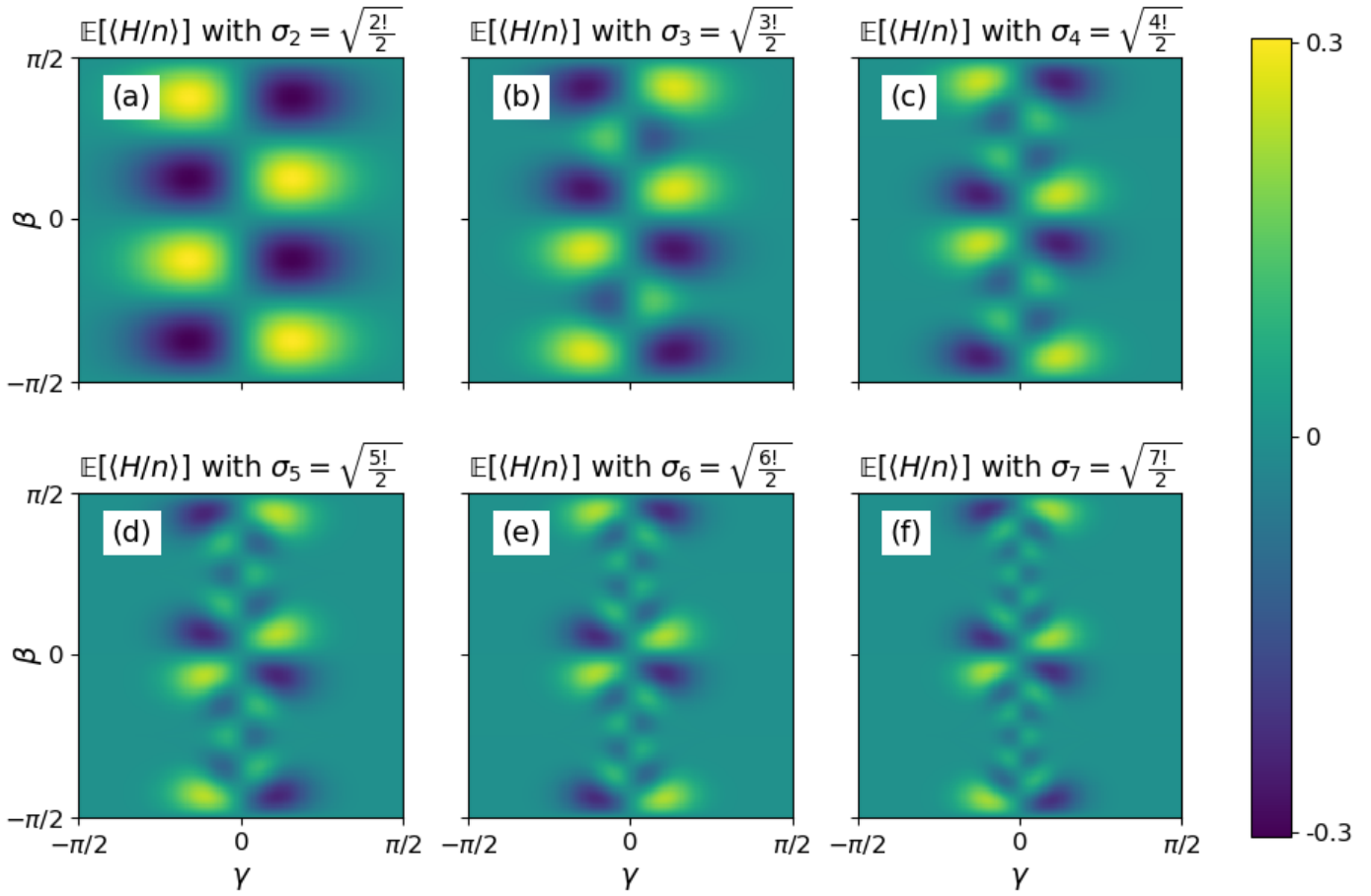


Figure 1: The energy landscapes for the pure  $d$ -spin models with  $\sigma_q = \delta_{d,q} \sqrt{d!/2}$ , in the limit as  $n \rightarrow \infty$ .

### 3.2 Numerical Results and Relation to Prior Results

As an example of using our Eqs. 6 and 7, we can compute the optimal angles for pure  $d$ -spin models given by  $\sigma_q = \delta_{d,q} \sqrt{d!/2}$ . In this case, our central equation reduces to

$$\begin{aligned} \lim_{n \rightarrow \infty} \mathbb{E}_{J \sim N} [\langle H/n \rangle] &= \gamma d! \sum_{\substack{a \text{ odd} \\ a \leq d}} \frac{(-1)^{\frac{a-1}{2}}}{a!(d-a)!} \exp(-2\gamma^2 ad) \sin^a(2\beta) \cos^{d-a}(2\beta) \\ &= \gamma \Im \left[ \left( \cos(2\beta) + i \sin(2\beta) \exp(-2\gamma^2 d) \right)^d \right]. \end{aligned} \quad (11)$$

Examples of this energy landscape for small  $d$  are plotted in Fig. 1, and numerically optimized angles and corresponding energy per spin are plotted in Fig. 2.

As another example application of our formula, we explore the instance independence at finite  $n$  to demonstrate convergence as  $n \rightarrow \infty$ . In Fig. 3a-e, we plot the expected energy per spin  $\langle H/n \rangle$  for five randomly generated mixed-spin optimization problems generated from a distribution with  $(\sigma_1, \sigma_2, \sigma_3) = (1/3, 1/2, 1)$  (and all higher-order terms zero). We see that as we increase  $n$ , the energy landscape of any given problem instance quickly approaches the  $n \rightarrow \infty$  landscape (Fig. 3f), as predicted.

To put Eq. 6 in perspective we explicitly calculate what it implies for two basic cases  $q = 2$  and  $q = 3$  and compare it with the literature on mixed-spin models.

The standard Sherrington-Kirkpatrick model has  $q = 2$ ,  $\sigma_2 = 1$ , and  $\sigma_q = 0$  for all other  $q \neq 2$ . In this case our Eq. 6 tells us that

$$\mathbb{E}_{J \sim N} [\langle H_{SK}/n \rangle] = 2\gamma e^{-2\gamma^2} \sin(2\beta) \cos(2\beta) = \gamma e^{-2\gamma^2} \sin(4\beta), \quad (12)$$

which agrees with previous work in [2] and [8]. This expectation is minimized by the parameters  $\beta_* = \pm\pi/8 \approx \pm 0.392699$  and  $\gamma_* = \mp 0.5$  for which  $\mathbb{E}[\langle H_{SK}/n \rangle] = -1/\sqrt{4e} \approx -0.303265$ . This is the same result we found numerically for  $d = 2$  in Fig. 2.

For the more general case  $q \geq 2$  we will briefly review the notation used by the articles [4, 5, 17, 18]. In the current paper the mixed-spin model is described by the Hamiltonian

$$H = \sum_{q=1}^d \frac{1}{\sqrt{n^{q-1}}} \sum_{1 \leq j_1 < \dots < j_q \leq n} J_{j_1 \dots j_q} Z_{j_1} \dots Z_{j_q}, \quad (13)$$

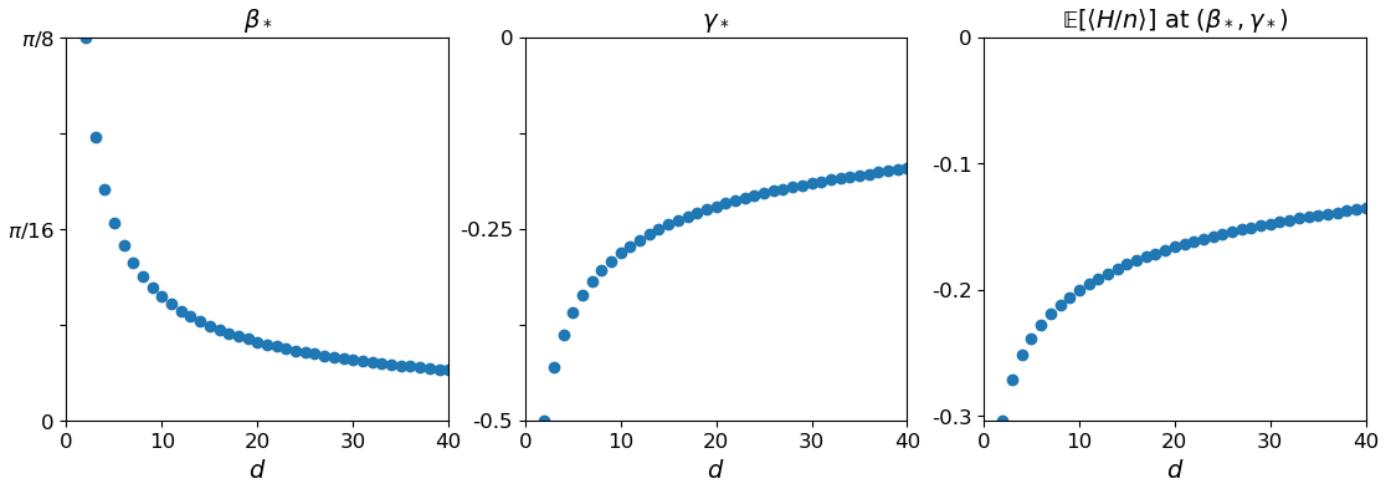


Figure 2: The optimal angles  $(\beta_*, \gamma_*)$  for the pure  $d$ -spin models with  $\sigma_q = \delta_{d,q} \sqrt{d!}/2$  that minimize the energy per spin  $\mathbb{E}[\langle H/n \rangle]$  in the limit  $n \rightarrow \infty$ .

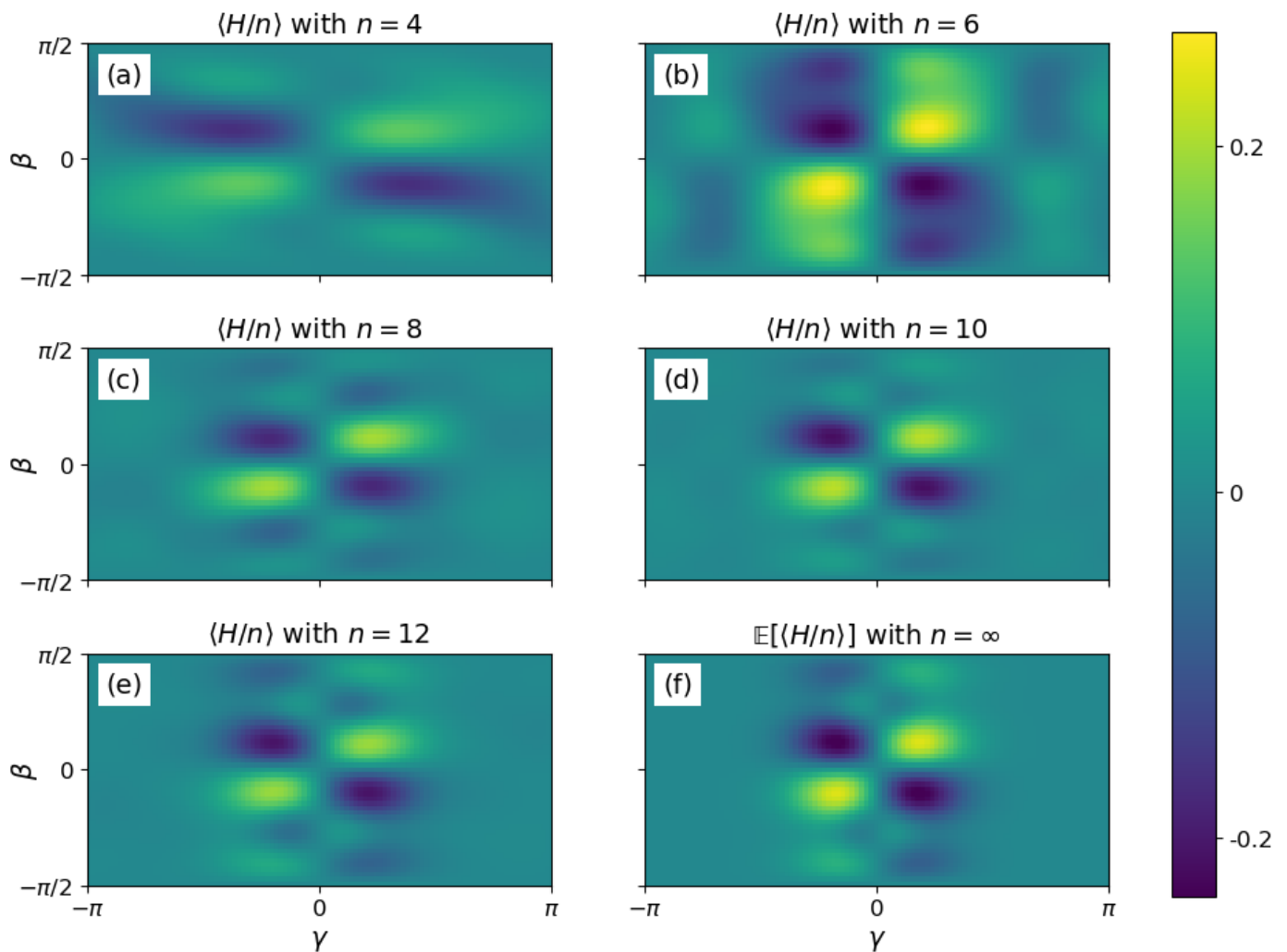


Figure 3: (a-e) The energy landscape of random instances of a mixed-spin optimization problem with  $(\sigma_1, \sigma_2, \sigma_3) = (1/3, 1/2, 1)$ , at varying problem sizes  $n$ . (f) The predicted energy landscape in the limit  $n \rightarrow \infty$ . We see the energy landscape for random instances rapidly converges to the  $n \rightarrow \infty$  landscape, demonstrating instance independence in this limit.

with  $J_{j_1 \dots j_q} \sim N(0, \sigma_q^2)$  and for a fixed  $q$  the summation has  $\binom{n}{q}$  terms. In [5] and elsewhere a different notation is used where the Hamiltonian is defined by

$$H' = \sum_{q=1}^d \frac{c_q}{\sqrt{n^{q-1}}} \sum_{1 \leq j_1, \dots, j_q \leq n} J_{j_1 \dots j_q} Z_{j_1} \cdots Z_{j_q}, \quad (14)$$

with  $J_{j_1 \dots j_q} \sim N(0, 1)$ . Crucially, now the indices  $j_1, \dots, j_q$  can have repeated values and permutations are treated distinctly, hence for each  $q$  the summation involves  $n^q$  terms.

In the limit of large  $n$  one can show that the  $q$ -plets  $(j_1, \dots, j_q)$  with repeated values can be ignored as their relative contribution decreases as a function  $n$ . For distinct indices there are  $q!$  permutations to consider in  $H'$ , hence we have a sum of  $q!$  standard normal distributions:  $J'_{j_1 \dots j_q} + \dots + J'_{j_q \dots j_1}$ , which is identical to a single normal distribution with variance  $q!$ . As a result we re-express  $H'$  as

$$H' = \sum_{q=1}^d \frac{c_q \sqrt{q!}}{\sqrt{n^{q-1}}} \sum_{1 \leq j_1 < \dots < j_q \leq n} J_{j_1 \dots j_q} Z_{j_1} \cdots Z_{j_q}, \quad (15)$$

with  $J \sim N(0, 1)$ . Thus when  $\sigma_q = c_q \sqrt{q!}$  we have that  $H$  and  $H'$  describe the same model. It is common to capture the different  $c_q$  coefficients by the *mixture* function  $\xi(x) = \sum_q c_q^2 x^q$ , such that the standard SK model has  $\xi(x) = x^2/2$  (i.e.  $c_2 = 1/\sqrt{2}$  and hence  $\sigma_2 = 1$ ).

For the 3-body model  $H_3$ , [4] analyzes the mixture  $\xi(x) = c_3^2 x^3 = x^3/2$ , such that  $c_3 = 1/\sqrt{2}$ . It was shown that the expected ground state energy-per-spin of this model equals  $-0.8132 \pm 0.0001$ . In the notation of the current article this model is equivalent to  $\sigma_3 = \sqrt{3}$  and with this our Eq. 6 gives the expectation

$$\mathbb{E}_{J \sim N} [\langle H_3/n \rangle] = 3\gamma e^{-3\gamma^2} \sin(2\beta) \cos^2(2\beta) - \gamma e^{-9\gamma^2} \sin^3(2\beta). \quad (16)$$

This expectation is minimized to  $\mathbb{E}_J[\langle H_3/n \rangle] \approx -0.270638$  by the angles  $(\beta_*, \gamma_*) \approx (\pm 0.290003, \mp 0.430091)$ , which are the solutions to

$$\beta_* = -\frac{1}{4} \arccos\left(1 - \frac{1}{9\gamma_*^2}\right) \quad (17)$$

$$\exp(-6\gamma_*^2) = 18\gamma_*^2 - 3 \quad (18)$$

$$\mathbb{E}_J[\langle H_3/n \rangle] = \sqrt{4\gamma_*^2 - \frac{2}{3}}. \quad (19)$$

This is the same result we found numerically for  $d = 3$  in Fig. 2. We thus see how depth  $p = 1$  QAOA can approximate the ground state energy by a factor of 0.332806. This 0.33 approximation factor had been reported earlier by Zhou et al.[19]

## 4 Derivation of Main Result

The proof in this section follows to large extent the framework of the earlier result by Farhi et al. [2], which relied on manipulating the moment generating function  $\mathbb{E}_J [e^{i\lambda H/n}]$  to extract expressions for the first and second moments of  $(H/n)$ . We use their method of simplifying the moment generating function, and their reorganization of the sum over  $z$ -strings into a sum over sketches (see Section 4.2). We extend their proof technique by generalizing their form of the moment generating function to higher-spin models and demonstrating that it can still be written as a sum over sketches, developing careful power-counting methods to allow us to extract the relevant terms in the  $n \rightarrow \infty$  limit, and deriving identities that allow us to explicitly evaluate the relevant sums.

In our proof, we will use the following conventions:

- A  $Z$ -basis state  $|z\rangle$  is specified by a string  $z = (z_1, z_2, \dots, z_n)$  of  $\pm 1$ s. We will use the shorthand  $z \in \{\pm 1\}^n$  for this.
- The XOR of two bitstrings  $z$  and  $z'$  is given by the componentwise product  $zz' := (z_1 z'_1, z_2 z'_2, \dots, z_n z'_n) \in \{\pm 1\}^n$ .
- For a set  $S \subseteq \{1, \dots, n\}$ , we denote the product of the bits in  $S$  as  $\prod_{i \in S} z_i =: z_S$ ; with this convention we thus have  $Z_S |z\rangle = z_S |z\rangle$ .
- The uniform superposition over all strings  $z$  is denoted by

$$|\Sigma\rangle := \frac{1}{\sqrt{2^n}} \sum_{z \in \{\pm 1\}^n} |z\rangle. \quad (20)$$

This is in contrast to the usual convention in quantum information, in which a  $Z$ -basis state is specified by a string  $z$  of 0s and 1s and the XOR of two strings is given by componentwise addition modulo 2. We choose our notation to be consistent with [2] and to simplify certain expressions in our derivation.

## 4.1 Moment Generating Function

Following [2], to evaluate  $\mathbb{E}_J[\langle H/n \rangle]$  and  $\mathbb{E}_J[\langle (H/n)^2 \rangle]$  we use the moment-generating function  $\mathbb{E}_J[\langle e^{i\lambda H/n} \rangle]$ . From the moment-generating function, we can find the moments via

$$\mathbb{E}_{J \sim N} [\langle (H/n)^m \rangle] = (-i\partial_\lambda)^m \mathbb{E}_{J \sim N} [\langle e^{i\lambda H/n} \rangle] \Big|_{\lambda=0} \quad (21)$$

We can simplify the expectation inside the moment-generating function by inserting three complete sets of states:

$$\begin{aligned} \langle e^{i\lambda H/n} \rangle &= \langle \Sigma | e^{i\gamma H} e^{i\beta B} e^{i\lambda H/n} e^{-i\beta B} e^{-i\gamma H} | \Sigma \rangle \\ &= \sum_{z, z', z'' \in \{\pm 1\}^n} \langle \Sigma | z \rangle \langle z | e^{i\gamma H} e^{i\beta B} | z'' \rangle \langle z'' | e^{i\lambda H/n} e^{-i\beta B} e^{-i\gamma H} | z' \rangle \langle z' | \Sigma \rangle \\ &= \frac{1}{2^n} \sum_{z, z', z'' \in \{\pm 1\}^n} \exp \left( i\gamma [H(z) - H(z')] + i \frac{\lambda H(z'')}{n} \right) \langle z | e^{i\beta B} | z'' \rangle \langle z'' | e^{-i\beta B} | z' \rangle \\ &= \frac{1}{2^n} \sum_{z, z', z'' \in \{\pm 1\}^n} \exp \left( i \sum_{q=1}^d n^{\frac{1-q}{2}} \sum_{|S|=q} J_S \left[ \gamma(z_S - z'_S) + \frac{\lambda z''_S}{n} \right] \right) \langle z | e^{i\beta B} | z'' \rangle \langle z'' | e^{-i\beta B} | z' \rangle \\ &= \frac{1}{2^n} \sum_{z, z', z'' \in \{\pm 1\}^n} \exp \left( i \sum_{q=1}^d n^{\frac{1-q}{2}} \sum_{|S|=q} z''_S J_S \left[ \gamma(z_S - z'_S) + \frac{\lambda}{n} \right] \right) \langle z | e^{i\beta B} | (+1)^n \rangle \langle (+1)^n | e^{-i\beta B} | z' \rangle, \end{aligned} \quad (22)$$

where in the last step we used that  $\langle z | e^{i\beta B} | z' \rangle = \langle z z' | e^{i\beta B} | (+1)^n \rangle$  and made the replacements  $z \mapsto z z''$  and  $z' \mapsto z' z''$ .

## 4.2 Expectation when Couplings are Normal Distributed Variables

We will now treat the  $J_S$  couplings as a random variable and consider the expectation  $\mathbb{E}_J$  of the energy. We assume that the distribution is symmetric with respect to  $J \leftrightarrow -J$ , such that we can replace  $z''_S J_S$  by  $J_S$  to get

$$\begin{aligned} \mathbb{E}_J [\langle e^{i\lambda H/n} \rangle] &= \frac{1}{2^n} \sum_{z, z', z'' \in \{\pm 1\}^n} \mathbb{E}_{J \sim N} \left[ \exp \left( i \sum_{q=1}^d n^{\frac{1-q}{2}} \sum_{|S|=q} z''_S J_S \left[ \gamma(z_S - z'_S) + \frac{\lambda}{n} \right] \right) \right] \langle z | e^{i\beta B} | (+1)^n \rangle \langle (+1)^n | e^{-i\beta B} | z' \rangle \\ &= \sum_{z, z' \in \{\pm 1\}^n} \mathbb{E}_{J \sim N} \left[ \exp \left( i \sum_{q=1}^d n^{\frac{1-q}{2}} \sum_{|S|=q} J_S \left[ \gamma(z_S - z'_S) + \frac{\lambda}{n} \right] \right) \right] \langle z | e^{i\beta B} | (+1)^n \rangle \langle (+1)^n | e^{-i\beta B} | z' \rangle. \end{aligned} \quad (23)$$

When we further assume that the  $J_S$  variables are independent between different  $S$  we can continue by

$$\mathbb{E}_J [\langle e^{i\lambda H/n} \rangle] = \sum_{z, z' \in \{\pm 1\}^n} \prod_{q=1}^d \prod_{|S|=q} \mathbb{E}_{J_S} \left[ \exp \left( i \cdot n^{\frac{1-q}{2}} J_S \left[ \gamma(z_S - z'_S) + \frac{\lambda}{n} \right] \right) \right] \langle z | e^{i\beta B} | (+1)^n \rangle \langle (+1)^n | e^{-i\beta B} | z' \rangle. \quad (24)$$

Next we assume that the  $J_S$  are normally distributed with a standard deviation that is the same for all sets  $S$  of the same size, that is  $J_S \sim N(0, \sigma_{|S|}^2)$ . We note that taking the expectation value of a Gaussian random variable  $J$  with standard deviation  $\sigma$  gives

$$\mathbb{E}_{J \sim N(0, \sigma^2)} [e^{icJ}] = \frac{1}{\sigma\sqrt{2\pi}} \int e^{-J^2/2\sigma^2 + icJ} = e^{-c^2\sigma^2/2} \quad (25)$$

so that our overall expression becomes

$$\mathbb{E}_{J \sim N} [\langle e^{i\lambda H/n} \rangle] = \sum_{z, z' \in \{\pm 1\}^n} \exp \left( - \sum_{q=1}^d \frac{\sigma_q^2}{2n^{q-1}} \sum_{|S|=q} \left[ \gamma^2(z_S - z'_S)^2 + \frac{2\gamma\lambda(z_S - z'_S)}{n} + \frac{\lambda^2}{n^2} \right] \right) \langle z | e^{i\beta B} | (+1)^n \rangle \langle (+1)^n | e^{-i\beta B} | z' \rangle. \quad (26)$$

To do the sum over  $z, z'$ , we claim that the summand does not depend on all  $2n$  spin values of  $z$  and  $z'$ . Instead, it is only a function of the four integer values  $(n_{++}, n_{+-}, n_{-+}, n_{--})$ , where  $n_{ss'}$  is defined to be the number of positions  $k \in \{1, \dots, n\}$  with  $(z_k, z'_k) = (s, s')$ . Note that only three of these variables are actually independent, as we always have  $n_{++} + n_{+-} + n_{-+} + n_{--} = n$ . As these numbers summarize the crucial information of the strings, we will refer to  $(n_{++}, n_{+-}, n_{-+}, n_{--})$  as the *sketch* of  $(z, z')$ . Writing the summand in terms of the sketch rather than  $(z, z')$  was introduced in [2]; here we establish that we can still write the summand in terms of the sketch for the mixed-spin SK model. To start, it is straightforward to verify that

$$\langle z | e^{i\beta B} | (+1)^n \rangle \langle (+1)^n | e^{-i\beta B} | z' \rangle = \prod_{ss' \in \{\pm\}^2} Q_{ss'}^{n_{ss'}}, \quad \text{with} \quad \begin{cases} Q_{++} := \cos^2(\beta) \\ Q_{--} := \sin^2(\beta) \\ Q_{-+} := i \sin(\beta) \cos(\beta) \\ Q_{+-} := -i \sin(\beta) \cos(\beta). \end{cases} \quad (27)$$

We can also write explicit combinatorial formulas for the sums in the exponential:

$$\begin{aligned} \sum_{\substack{S \subseteq \{1, \dots, n\} \\ |S|=q}} (z_S - z'_S) &= \sum_{k=0}^q (-1)^{q-k} \left[ \binom{n_{++} + n_{+-}}{k} \binom{n_{-+} + n_{--}}{q-k} - \binom{n_{++} + n_{-+}}{k} \binom{n_{+-} + n_{--}}{q-k} \right] \\ &= \sum_{k=0}^q (-1)^{q-k} \left[ \binom{\frac{n+(n_{+-}-n_{-+})+(n_{++}-n_{--})}{2}}{k} \binom{\frac{n-(n_{+-}-n_{-+})-(n_{++}-n_{--})}{2}}{q-k} \right. \\ &\quad \left. - \binom{\frac{n-(n_{+-}-n_{-+})+(n_{++}-n_{--})}{2}}{k} \binom{\frac{n+(n_{+-}-n_{-+})-(n_{++}-n_{--})}{2}}{q-k} \right] \end{aligned} \quad (28)$$

$$\begin{aligned} \sum_{\substack{S \subseteq \{1, \dots, n\} \\ |S|=q}} (z_S - z'_S)^2 &= 4 \sum_{k \text{ odd}} \binom{n_{+-} + n_{-+}}{k} \binom{n_{++} + n_{--}}{q-k} \\ &= 4 \sum_{k \text{ odd}} \binom{n_{+-} + n_{-+}}{k} \binom{n - (n_{+-} + n_{-+})}{q-k} \\ &=: g_q(n_{++}, n_{+-}, n_{-+}, n_{--}). \end{aligned} \quad (29)$$

Therefore, the summand indeed depends only on  $(n_{++}, n_{+-}, n_{-+}, n_{--})$  and the number of ways to assign the  $n$  positions into four groups of these sizes is the multinomial  $n!/(n_{++}!n_{+-}!n_{-+}!n_{--}!)$ . To condense our notation we will use  $\{n_*\}$  to denote the set of sketches  $n_* = (n_{++}, n_{+-}, n_{-+}, n_{--})$ , allowing us to use the shorthand

$$\sum_{\{n_*\}} \binom{n}{n_*} F(n_*) := \sum_{\substack{(n_{++}, n_{+-}, n_{-+}, n_{--}) \in \mathbb{N}^4 \\ n_{++} + n_{+-} + n_{-+} + n_{--} = n}} \binom{n}{n_{++}, n_{+-}, n_{-+}, n_{--}} F(n_{++}, n_{+-}, n_{-+}, n_{--}). \quad (30)$$

Note that this summation has  $\binom{n+3}{3}$  terms. We thus have

$$\mathbb{E}_{J \sim N} [\langle e^{i\lambda H/n} \rangle] = \sum_{\{n_*\}} \binom{n}{n_*} \exp \left( - \sum_{q=1}^d \left[ \gamma^2 g_q(n_*) + 2\gamma\lambda \frac{f_q(n_*)}{n} + \lambda^2 \frac{\binom{n}{q}}{n^2} \right] \frac{\sigma_q^2}{2n^{q-1}} \right) \prod_{ss' \in \{\pm\}^2} Q_{ss'}^{n_{ss'}}. \quad (31)$$

Eq. 31 is the form of the moment-generating function we will use to evaluate  $\mathbb{E}_J[\langle H/n \rangle]$  and  $\mathbb{E}_J[\langle (H/n)^2 \rangle]$ .

### 4.3 General Form of Moments

Using Eq. 21 combined with the form of the moment-generating function given in Eq. 31, we can write the first moment as:

$$\mathbb{E}_{J \sim N} [\langle H/n \rangle] = i\gamma \sum_{q=1}^d \frac{\sigma_q^2}{n^q} \sum_{\{n_*\}} \binom{n}{n_*} \exp \left( - \sum_{q'=1}^d \gamma^2 g_{q'}(n_*) \frac{\sigma_{q'}^2}{2n^{q'-1}} \right) f_q(n_*) \prod_{ss' \in \{\pm\}^2} Q_{ss'}^{n_{ss'}}, \quad (32)$$

and the second moment as:

$$\mathbb{E}_{J \sim N} [\langle (H/n)^2 \rangle] = \sum_{q=1}^d \binom{n}{q} \frac{\sigma_q^2}{n^{q+1}} - \gamma^2 \sum_{q, q'=1}^d \frac{\sigma_q^2 \sigma_{q'}^2}{n^{q+q'}} \sum_{\{n_*\}} \binom{n}{n_*} \exp \left( - \sum_{q''=1}^d \gamma^2 g_{q''}(n_*) \frac{\sigma_{q''}^2}{2n^{q''-1}} \right) f_q(n_*) f_{q'}(n_*) \prod_{ss' \in \{\pm\}^2} Q_{ss'}^{n_{ss'}}. \quad (33)$$

The explicit expression for  $f_q$  (Eq. 28) shows that  $f_q$  is a degree- $q$  polynomial in the variables  $(n_{+-} - n_{-+})$ ,  $(n_{++} - n_{--})$ , and  $n$ , so we can expand  $f_q$  as

$$f_q(n_*) = \sum_{a+b+c \leq q} f_q^{abc} (n_{+-} - n_{-+})^a (n_{++} - n_{--})^b n^c \quad (34)$$

where the  $f_q^{abc}$  are constants independent of  $n$ . In terms of this expansion, we have

$$\mathbb{E}_{J \sim N} [\langle H/n \rangle] = i\gamma \sum_{q=1}^d \sigma_q^2 \sum_{a+b+c \leq q} \frac{f_q^{abc}}{n^{q-c}} \sum_{\{n_*\}} \binom{n}{n_*} \exp \left( - \sum_{q'=1}^d g_{q'}(n_*) \frac{\gamma^2 \sigma_{q'}^2}{2n^{q'-1}} \right) (n_{+-} - n_{-+})^a (n_{++} - n_{--})^b \prod_{ss' \in \{\pm\}^2} Q_{ss'}^{n_{ss'}} \quad (35)$$



and

$$\mathbb{E}_{J \sim N} [\langle (H/n)^2 \rangle] = \sum_{q=1}^d \binom{n}{q} \frac{\sigma_q^2}{n^{q+1}} - \gamma^2 \sum_{q,q'=1}^d \sigma_q^2 \sigma_{q'}^2 \sum_{\substack{a+b+c \leq q \\ a'+b'+c' \leq q'}} \frac{f_q^{abc} f_{q'}^{a'b'c'}}{n^{q-c} n^{q'-c'}} \sum_{\{n_*\}} \binom{n}{n_*} \exp \left( - \sum_{q''=1}^d g_{q''}(n_*) \frac{\gamma^2 \sigma_{q''}^2}{2n^{q''-1}} \right) \times (n_{+-} - n_{-+})^{a+a'} (n_{++} - n_{--})^{b+b'} \prod_{ss' \in \{\pm\}^2} Q_{ss'}^{n_{ss'}}. \quad (36)$$

Ref. [2] could explicitly evaluate these terms for the small values of  $a$  and  $b$  relevant for the two-body SK model, using concise expressions for  $f_2$  and  $g_2$ . However, to get explicit formulas beyond  $q = 2$  requires carefully counting powers of  $n$  to establish which terms survive in the  $n \rightarrow \infty$  limit, and using the general expressions for  $f_q$  and  $g_q$  (Eqs. 28 and 29) to derive explicit forms of the leading-order terms. To tame this sum our derivation thus uses techniques that go beyond a simple generalization of [2].

#### 4.4 Evaluating Sums Over the Sketches $(n_{++}, n_{+-}, n_{-+}, n_{--})$

We see that evaluating both moments reduces to repeatedly evaluating terms of the form

$$T_\xi^{ab} := \frac{1}{n^\xi} \sum_{\{n_*\}} \binom{n}{n_*} \exp \left( - \sum_{q'=1}^d \gamma^2 g_{q'}(n_*) \frac{\sigma_{q'}^2}{2n^{q'-1}} \right) (n_{+-} - n_{-+})^a (n_{++} - n_{--})^b \prod_{ss' \in \{\pm\}^2} Q_{ss'}^{n_{ss'}} \quad (37)$$

with  $\xi \geq a+b$  playing the role of  $q-c$  in Eq. 35. For a fixed  $n$ , note that  $g_{q'}(n_*)$  only depends on  $n_{+-} + n_{-+}$  (Eq. 29) and with this symmetry we define the single variable function  $\bar{g}_{q'}(n_{+-} + n_{-+}) := g_{q'}(n_{++}, n_{+-}, n_{-+}, n_{--})$  and use  $\sum_t$  with  $t = n_{+-} + n_{-+}$ . The remaining two degrees of freedom are further reorganized by summing over  $n_{+-}$  using  $\sum_{n_{+-} + n_{-+} = t}$ , and over  $n_{++}$  using  $\sum_{n_{++} + n_{--} = n-t}$ . We thus obtain

$$T_\xi^{ab} = \frac{1}{n^\xi} \sum_{t=0}^n \binom{n}{t} \exp \left( - \sum_{q'=1}^d \gamma^2 \bar{g}_{q'}(t) \frac{\sigma_{q'}^2}{2n^{q'-1}} \right) \underbrace{\left[ \sum_{n_{+-} + n_{-+} = t} \binom{t}{n_{+-}} (n_{+-} - n_{-+})^a Q_{+-}^{n_{+-}} Q_{-+}^{n_{-+}} \right]}_{A_t^a} \times \underbrace{\left[ \sum_{n_{++} + n_{--} = n-t} \binom{n-t}{n_{++}} (n_{++} - n_{--})^b Q_{++}^{n_{++}} Q_{--}^{n_{--}} \right]}_{B_t^b}, \quad (38)$$

where we have defined  $A_t^a$  and  $B_t^b$  to be analyzed next in the limit of large  $n$ .

#### 4.5 Large $n$ Limit of $A_t^a$ , $B_t^b$ , and $T_\xi^{ab}$

As we plan to take the limit  $n \rightarrow \infty$ , we will use big  $O(\cdot)$  notation in  $n$  to keep track of the significant terms in  $A_t^a$  and  $B_t^b$  for constant  $a, b$ , and  $t$ . The summations for  $A_t^a$  and  $B_t^b$  can be expressed exactly via the following identity (which is a generalization of the identities given in [2] for the case of  $a = 1, 2$  and which can be proven by induction on  $a$ ):

$$\sum_{n_1 + n_2 = u} \binom{u}{n_1} (n_1 - n_2)^a Q_1^{n_1} Q_2^{n_2} = (x\partial_x - y\partial_y)^a (x+y)^u \Big|_{\substack{x=Q_1 \\ y=Q_2}} \quad (39)$$

For  $A_t^a$ , we thus have

$$A_t^a = (x\partial_x - y\partial_y)^a (x+y)^t \Big|_{\substack{x=Q_{+-} \\ y=Q_{-+}}} \quad (40)$$

Using the operator equality  $\partial_x \cdot x = 1 + x\partial_x$  we see that expanding  $(x\partial_x - y\partial_y)^a$  gives a sum of differential operators  $x^k y^\ell \partial_x^k \partial_y^\ell$  with  $k + \ell \leq a$ . Because  $(Q_{+-} + Q_{-+}) = 0$  (Eq. 27) we see that only the operators  $\partial_x^k \partial_y^\ell$  with  $k + \ell = t$  give nonzero contributions to  $A_t^a$ . For  $t > a$  we therefore have  $A_t^a = 0$ . In the particular case of  $t = a$ , the only terms in  $(x\partial_x - y\partial_y)^a$  that give a nonzero result are  $\sum_{k+\ell=a} \binom{a}{k} x^k (-y)^\ell \partial_x^k \partial_y^\ell$ , hence

$$A_{a,a} = a! \sum_{k+\ell=a} \binom{a}{k} Q_{+-}^k (-Q_{-+})^\ell = a! (Q_{+-} - Q_{-+})^a = a! (-i)^a \sin^a(2\beta). \quad (41)$$

For  $t < a$  it will be enough to observe that the order of  $A_t^a$  is independent of  $n$ . Summarizing, we have

$$A_t^a = \begin{cases} 0 & \text{if } t > a, \\ a! (-i)^a \sin^a(2\beta) & \text{if } t = a, \\ O(1) & \text{if } t < a. \end{cases} \quad (42)$$

For  $B_t^b$ , we use

$$B_t^b = (x\partial_x - y\partial_y)^b (x+y)^{n-t} \Big|_{\substack{x=Q_{++} \\ y=Q_{--}}} . \quad (43)$$

The highest-order terms in  $n$  come from the terms  $\sum_{k+\ell=b} \binom{b}{k} x^k (-y)^\ell \partial_x^k \partial_y^\ell$  in  $(x\partial_x - y\partial_y)^b$  and thus

$$\begin{aligned} B_t^b &= \frac{(n-t)!}{(n-t-b)!} \sum_{k+\ell=b} \binom{b}{k} Q_{++}^k (-Q_{--})^\ell (Q_{++} + Q_{--})^{n-t-b} + O(n^{b-1}) \\ &= n^b (Q_{++} - Q_{--})^b (Q_{++} + Q_{--})^{n-t-b} + O(n^{b-1}) \\ &= n^b \cos^b(2\beta) + O(n^{b-1}). \end{aligned} \quad (44)$$

Plugging our Eqs. 42 and 44 for  $A_t^a$  and  $B_t^b$  into Eq. 38 for  $T_\xi^{ab}$  hence gives

$$\begin{aligned} T_\xi^{ab} &= \frac{1}{n^\xi} \binom{n}{a} \exp\left(-\sum_{q'=1}^d \gamma^2 \bar{g}_{q'}(a) \frac{\sigma_{q'}^2}{2n^{q'-1}}\right) \cdot a! (-i)^a \sin^a(2\beta) \cdot [n^b \cos^b(2\beta) + O(n^{b-1})] \\ &\quad + \frac{1}{n^\xi} \sum_{t=0}^{a-1} \binom{n}{t} \exp\left(-\sum_{q'=1}^d \gamma^2 \bar{g}_{q'}(t) \frac{\sigma_{q'}^2}{2n^{q'-1}}\right) \cdot O(1) \cdot [n^b \cos^b(2\beta) + O(n^{b-1})]. \end{aligned} \quad (45)$$

As  $\binom{n}{a} = n^a/a! + O(n^{a-1})$  and similarly  $\binom{n}{t} = O(n^t)$  this simplifies to

$$T_\xi^{ab} = \frac{n^{a+b-\xi}}{a!} \exp\left(-\sum_{q'=1}^d \gamma^2 \bar{g}_{q'}(a) \frac{\sigma_{q'}^2}{2n^{q'-1}}\right) \cdot a! (-i)^a \sin^a(2\beta) \cdot \cos^b(2\beta) + O(n^{a+b-\xi-1}). \quad (46)$$

After reminding ourselves that  $a+b \leq \xi$  we see that, in the  $n \rightarrow \infty$  limit,  $T_\xi^{ab}$  further condenses to

$$T_\xi^{ab} \rightarrow \begin{cases} (-i)^a \exp\left(-\sum_{q'=1}^d \gamma^2 \sigma_{q'}^2 \lim_{n \rightarrow \infty} \left[\frac{\bar{g}_{q'}(a)}{2n^{q'-1}}\right]\right) \sin^a(2\beta) \cos^b(2\beta) & \text{if } a+b = \xi \\ 0 & \text{if } a+b < \xi. \end{cases} \quad (47)$$

As  $\bar{g}_{q'}(a)$  has an implicit  $n$  dependency we have to determine its relevant terms as  $n \rightarrow \infty$ . Eq. 47 shows that the only relevant terms in  $\bar{g}_{q'}$  are those of degree at least  $(q'-1)$  in  $n$ . From the definition of  $g_q$  (Eq. 29), elementary algebra gives

$$\bar{g}_{q'}(n_{+-} + n_{-+}) = \frac{4(n_{+-} + n_{-+})n^{q'-1}}{(q'-1)!} + O(n^{q'-2}). \quad (48)$$

Therefore, we have the large  $n$  limit

$$T_\xi^{ab} \rightarrow \begin{cases} (-i)^a \exp\left(-2a\gamma^2 \sum_{q'=1}^d \frac{\sigma_{q'}^2}{(q'-1)!}\right) \sin^a(2\beta) \cos^b(2\beta) & \text{if } a+b = \xi \\ 0 & \text{if } a+b < \xi. \end{cases} \quad (49)$$

## 4.6 Large $n$ properties of $f_q$

To complete the evaluation of the moments, we to determine the large  $n$  dependency of  $f_q$ , its  $f_q^{abc}$  coefficients (Eq. 34) and their role in the first and second moment expressions of Eqs. 35 and 36. We see from Eq. 49 that the only coefficients that matter in the  $n \rightarrow \infty$  limit are those with  $a+b = \xi$ , i.e. those with  $a+b+c = q$ . We can evaluate these  $f_q^{abc}$  from the explicit formula for  $f_q$  (Eq. 28) by keeping only terms of total degree  $q$  in the terms  $(n_{+-} - n_{-+})$ ,  $(n_{++} - n_{--})$ , and  $n$ . This gives

$$f_q(n_{++}, n_{+-}, n_{-+}, n_{--}) = \sum_{\substack{a \text{ odd} \\ a \leq q}} \frac{2}{a!(q-a)!} (n_{+-} - n_{-+})^a (n_{++} - n_{--})^{q-a} + (\text{lower degree terms}) \quad (50)$$

Note that there are no terms  $(n_{+-} - n_{-+})^a (n_{++} - n_{--})^b n^c$  with  $c \neq 0$  and  $(a+b+c) = q$  in  $f_q$ . Comparing Eq. 50 with the definition of  $f_q^{abc}$  in Eq. 34, we see for  $n \rightarrow \infty$  after some algebra that

$$f_q^{abc} = \begin{cases} \frac{2}{a!b!} & \text{if } a+b = q, c = 0, \text{ and } a \text{ odd} \\ 0 & \text{if } a+b+c = q \text{ with } a \text{ even or } c > 0. \end{cases} \quad (51)$$

Having established these properties of  $T_{q-c}^{ab}$  and  $f_q$ , we now have sufficient information to evaluate our moments.

## 4.7 Evaluating the First Moment

To evaluate the first moment, we simplify Eq. 35 using the definition of  $T_\xi^{ab}$  (Eq. 37) and the results of Eqs. 49 and 51 to get

$$\begin{aligned} \mathbb{E}_{J \sim N} [(H/n)] &= i\gamma \sum_{q=1}^d \sigma_q^2 \sum_{a+b+c \leq q} f_q^{abc} T_{q-c}^{ab} \\ &\rightarrow i\gamma \sum_{q=1}^d \sigma_q^2 \sum_{\substack{a \text{ odd} \\ a \leq q}} \frac{2}{a!(q-a)!} (-i)^a \exp\left(-2\gamma^2 a \sum_{q'=1}^d \frac{\sigma_{q'}^2}{(q'-1)!}\right) \sin^a(2\beta) \cos^{q-a}(2\beta) \\ &= 2\gamma \sum_{q=1}^d \sigma_q^2 \sum_{\substack{a \text{ odd} \\ a \leq q}} \frac{(-1)^{\frac{a-1}{2}}}{a!(q-a)!} \exp\left(-2\gamma^2 a \sum_{q'=1}^d \frac{\sigma_{q'}^2}{(q'-1)!}\right) \sin^a(2\beta) \cos^{q-a}(2\beta), \end{aligned} \quad (52)$$

which is precisely what we claimed in Eq. 6.

## 4.8 Evaluating the Second Moment

To evaluate the second moment, we simplify Eq. 36 using the definition of  $T_\xi^{ab}$  (Eq. 37) and the results of Eqs. 49 and 51 to get

$$\begin{aligned} \mathbb{E}_{J \sim N} [(H/n)^2] &= \sum_{q=1}^d \binom{n}{q} \frac{\sigma_q^2}{n^{q+1}} - \gamma^2 \sum_{q, q'=1}^d \sigma_q^2 \sigma_{q'}^2 \sum_{\substack{a+b+c \leq q \\ a'+b'+c' \leq q'}} f_q^{abc} f_{q'}^{a'b'c'} T_{q-c+q'-c'}^{(a+a')(b+b')} \\ &\rightarrow -\gamma^2 \sum_{q, q'=1}^d \sigma_q^2 \sigma_{q'}^2 \sum_{\substack{a, a' \text{ odd} \\ a \leq q \\ a' \leq q'}} \frac{2}{a!(q-a)!} \frac{2}{a'!(q'-a')!} (-i)^{a+a'} \\ &\quad \times \exp\left(-2\gamma^2(a+a') \sum_{q''=1}^d \frac{\sigma_{q''}^2}{(q''-1)!}\right) \sin^{a+a'}(2\beta) \cos^{q-a+q'-a'}(2\beta) \\ &= \left[ i\gamma \sum_{q=1}^d \sigma_q^2 \sum_{\substack{a \text{ odd} \\ a \leq q}} \frac{2}{a!(q-a)!} (-i)^a \exp\left(-2\gamma^2 a \sum_{q''=1}^d \frac{\sigma_{q''}^2}{(q''-1)!}\right) \sin^a(2\beta) \cos^{q-a}(2\beta) \right]^2 \\ &= \lim_{n \rightarrow \infty} \mathbb{E}_{J \sim N} [(H/n)]^2, \end{aligned} \quad (53)$$

which is precisely what we claimed in Eq. 8.

## 4.9 Evaluating Higher Moments

The proof technique used above also applies to higher moments. When computing the  $m$ th moment by taking derivatives of the moment-generating function according to Eq. 21, the only terms that survive in the limit  $n \rightarrow \infty$  are the terms where all derivatives hit the  $2\gamma\lambda f_q/n$  term in the exponential, so that the expression for the moment becomes

$$\mathbb{E}_{J \sim N} [(H/n)^m] = (i\gamma)^m \sum_{q_1, \dots, q_m=1}^d \sigma_{q_1}^2 \cdots \sigma_{q_m}^2 \sum_{\substack{a_1+b_1+c_1 \leq q_1 \\ \dots \\ a_m+b_m+c_m \leq q_m}} f_{q_1}^{a_1 b_1 c_1} \cdots f_{q_m}^{a_m b_m c_m} T_{(q_1-c_1)+\dots+(q_m-c_m)}^{(a_1+\dots+a_m)(b_1+\dots+b_m)} + O\left(\frac{1}{n}\right). \quad (54)$$

Noting from Eq. 49 that

$$T_{(q_1-c_1)+\dots+(q_m-c_m)}^{(a_1+\dots+a_m)(b_1+\dots+b_m)} = T_{q_1-c_1}^{a_1 b_1} \cdots T_{q_m-c_m}^{a_m b_m} + O\left(\frac{1}{n}\right), \quad (55)$$

we can write the  $m$ th moment as

$$\begin{aligned} \mathbb{E}_{J \sim N} [(H/n)^m] &= (i\gamma)^m \sum_{q_1, \dots, q_m=1}^d \sigma_{q_1}^2 \cdots \sigma_{q_m}^2 \sum_{\substack{a_1+b_1+c_1 \leq q_1 \\ \dots \\ a_m+b_m+c_m \leq q_m}} f_{q_1}^{a_1 b_1 c_1} \cdots f_{q_m}^{a_m b_m c_m} T_{q_1-c_1}^{a_1 b_1} \cdots T_{q_m-c_m}^{a_m b_m} + O\left(\frac{1}{n}\right) \\ &= \left[ i\gamma \sum_{q=1}^d \sigma_q^2 \sum_{a+b+c \leq q} f_q^{abc} T_{q-c}^a \right]^m + O\left(\frac{1}{n}\right) \\ &= \mathbb{E}_{J \sim N} [(H/n)]^m + O\left(\frac{1}{n}\right). \end{aligned} \quad (56)$$

Thus, we find that for all  $m$ ,

$$\lim_{n \rightarrow \infty} \mathbb{E}_{J \sim N} [\langle (H/n)^m \rangle] = \lim_{n \rightarrow \infty} \mathbb{E}_{J \sim N} [\langle H/n \rangle]^m. \quad (57)$$

#### 4.10 Alternative expression for $\mathbb{E}[\langle H/n \rangle]$

Here, we present an alternative expression for  $\mathbb{E}[\langle H/n \rangle]$  that makes contact with the notation used in [5] and elsewhere (see Section 3.2). The current paper's central result (Eq. 6) reads:

$$\lim_{n \rightarrow \infty} \mathbb{E}_{J \sim N} [\langle H/n \rangle] = 2\gamma \sum_{q=1}^d \sigma_q^2 \sum_{\substack{a \text{ odd} \\ a \leq q}} \frac{(-1)^{\frac{a-1}{2}}}{a!(q-a)!} \exp\left(-2\gamma^2 a \sum_{q'=1}^d \frac{\sigma_{q'}^2}{(q'-1)!}\right) \sin^a(2\beta) \cos^{q-a}(2\beta). \quad (58)$$

If we use the substitution from Section 3.2,  $\sigma_q^2 = c_q^2 q!$ , this becomes

$$\begin{aligned} \lim_{n \rightarrow \infty} \mathbb{E}_{J \sim N} [\langle H/n \rangle] &= 2\gamma \sum_{q=1}^d c_q^2 q! \sum_{\substack{a \text{ odd} \\ a \leq q}} \frac{(-1)^{\frac{a-1}{2}}}{a!(q-a)!} \exp\left(-2\gamma^2 a \sum_{q'=1}^d q' c_{q'}^2\right) \sin^a(2\beta) \cos^{q-a}(2\beta) \\ &= 2i\gamma \sum_{q=1}^d c_q^2 \sum_{\substack{a \text{ odd} \\ 0 \leq a \leq q}} \binom{q}{a} \left(-i \sin(2\beta) \exp\left(-2\gamma^2 \sum_{q'=1}^d q' c_{q'}^2\right)\right)^a (\cos(2\beta))^{q-a}. \end{aligned} \quad (59)$$

Using the identity

$$2i \sum_{a \text{ odd}} \binom{q}{a} (-iA)^a B^{q-a} = -i((iA+B)^q - (-iA+B)^q) = 2\Im[(iA+B)^q], \quad (60)$$

we can rewrite this as

$$\lim_{n \rightarrow \infty} \mathbb{E}_{J \sim N} [\langle H/n \rangle] = 2\gamma \sum_{q=1}^d c_q^2 \Im \left[ \left( i \sin(2\beta) \exp\left(-2\gamma^2 \sum_{q'=1}^d q' c_{q'}^2\right) + \cos(2\beta) \right)^q \right]. \quad (61)$$

Finally, using the definition  $\xi(x) := \sum_q c_q^2 x^q$  and hence  $\xi'(1) = \sum_q q c_q^2$ , we get

$$\begin{aligned} \lim_{n \rightarrow \infty} \mathbb{E}_{J \sim N} [\langle H/n \rangle] &= 2\gamma \sum_{q=1}^d c_q^2 \Im \left[ \left( \cos(2\beta) + i \sin(2\beta) \exp(-2\gamma^2 \xi'(1)) \right)^q \right] \\ &= 2\gamma \Im \left[ \xi \left( \cos(2\beta) + i \sin(2\beta) \exp(-2\gamma^2 \xi'(1)) \right) \right], \end{aligned} \quad (62)$$

which is precisely what we claimed in Eq. 7. For a pure  $d$ -spin model with only  $c_d \neq 0$  the above further simplifies to

$$\lim_{n \rightarrow \infty} \mathbb{E}_{J \sim N} [\langle H/n \rangle] = 2\gamma c_d^2 \cdot \Im \left[ \left( \cos(2\beta) + i \sin(2\beta) \exp(-2dc_d^2 \gamma^2) \right)^d \right]. \quad (63)$$

## 5 Discussion and Conclusion

In this work, we have derived explicit formulas to quantify the performance of  $p = 1$  QAOA on mixed-spin models in the  $n \rightarrow \infty$  limit. We demonstrated both concentration of measure and instance independence for arbitrary mixed-spin models, which imply that the expectation value of the energy per spin is independent of the specific model specification and that measurements of the QAOA state are guaranteed to give energies close to the expectation value. Our explicit formula for the expectation value of the energy for arbitrary mixed-spin models allows us to find the optimal angles on a classical computer.

There are two obvious open questions raised by this work. First, the approach of this paper can probably be combined with the methods of [2] to generalize our work to depth  $p > 1$  QAOA. Such a result would allow one to prove instance independence and concentration of measure, and derive a computer algorithm to generate formulas for the expectation value per spin, at arbitrary depth  $p$ . This is a particularly interesting route of research, since it is known that in the cubic case with  $\sigma_q \propto \delta_{q,3}$ , Montanari's classical algorithm does not approach the optimal solution [5], so that at sufficient depth  $p$  the QAOA has a chance of outperforming the best known classical algorithm. Higher-spin models with  $q > 3$  may even provide a more direct route towards finding a setting where QAOA at depth  $p = 1$  outperforms classical optimization such as Montanari's algorithm [4, 5]. While the generalization to higher  $p$  is likely possible, it is a nontrivial extension of this paper, and we leave it for future work.

Second, it remains an open question what to what extent results on the random models can be used to find optimal angles for realistic binary optimization problems. One hypothetical approach to finding QAOA angles for a single instance of an  $n$ -spin optimization problem would be the procedure:

1. For all  $q = 1, \dots, d$ , compute the standard deviations  $\sigma_q$  of the  $q$ -spin couplings in the problem.
2. Use Eq. 6 for the mixed-spin model as an estimate for the expectation value of the energy per spin.
3. Run the QAOA at the optimal angles for the corresponding mixed-spin model (or use these angles as starting points for a numerical optimization of the angles)

This procedure assumes that the means of the couplings are zero:  $\mu_q = 0$  for all  $q = 1, \dots, d$ . If this method can be shown to work, it would greatly simplify finding QAOA angles for general optimization problems.

## References

- [1] Edward Farhi, Jeffrey Goldstone, and Sam Gutmann. A quantum approximate optimization algorithm. *arXiv preprint arXiv:1411.4028*, 2014. URL <https://arxiv.org/abs/1411.4028>.
- [2] Edward Farhi, Jeffrey Goldstone, Sam Gutmann, and Leo Zhou. The quantum approximate optimization algorithm and the Sherrington-Kirkpatrick model at infinite size. *arXiv preprint arXiv:1910.08187*, 2019. URL <https://arxiv.org/abs/1910.08187>.
- [3] Andrea Montanari. Optimization of the Sherrington–Kirkpatrick hamiltonian. *SIAM Journal on Computing*, 0(0):FOCS19–1, 2021. DOI: [10.1137/20M132016X](https://doi.org/10.1137/20M132016X).
- [4] Ahmed El Alaoui, Andrea Montanari, and Mark Sellke. Optimization of mean-field spin glasses. *arXiv preprint arXiv:2001.00904*, 2020. URL <https://arxiv.org/abs/2001.00904>.
- [5] Ahmed El Alaoui and Andrea Montanari. Algorithmic thresholds in mean field spin glasses. *arXiv preprint arXiv:2009.11481*, 2020. URL <https://arxiv.org/abs/2009.11481>.
- [6] Zhang Jiang, Eleanor G Rieffel, and Zhihui Wang. Near-optimal quantum circuit for Grover’s unstructured search using a transverse field. *Physical Review A*, 95(6):062317, 2017. DOI: [10.1103/PhysRevA.95.062317](https://doi.org/10.1103/PhysRevA.95.062317).
- [7] Zhihui Wang, Stuart Hadfield, Zhang Jiang, and Eleanor G Rieffel. Quantum approximate optimization algorithm for maxcut: A fermionic view. *Physical Review A*, 97(2):022304, 2018. DOI: [10.1103/PhysRevA.97.022304](https://doi.org/10.1103/PhysRevA.97.022304).
- [8] Asier Ozaeta, Wim van Dam, and Peter L. McMahon. Expectation values from the single-layer quantum approximate optimization algorithm on Ising problems. *arXiv preprint arXiv:2012.03421*, 2020. URL <https://arxiv.org/abs/2012.03421>.
- [9] Gavin E Crooks. Performance of the quantum approximate optimization algorithm on the maximum cut problem. *arXiv preprint arXiv:1811.08419*, 2018. URL <https://arxiv.org/abs/1811.08419>.
- [10] Leo Zhou, Sheng-Tao Wang, Soonwon Choi, Hannes Pichler, and Mikhail D Lukin. Quantum approximate optimization algorithm: Performance, mechanism, and implementation on near-term devices. *Physical Review X*, 10(2):021067, 2020. DOI: [10.1103/PhysRevX.10.021067](https://doi.org/10.1103/PhysRevX.10.021067).
- [11] Fernando GSL Brandao, Michael Broughton, Edward Farhi, Sam Gutmann, and Hartmut Neven. For fixed control parameters the quantum approximate optimization algorithm’s objective function value concentrates for typical instances. *arXiv preprint arXiv:1812.04170*, 2018. URL <https://arxiv.org/abs/1812.04170>.
- [12] David Sherrington and Scott Kirkpatrick. Solvable model of a spin-glass. *Physical Review Letters*, 35(26):1792, 1975. DOI: [10.1103/PhysRevLett.35.1792](https://doi.org/10.1103/PhysRevLett.35.1792).
- [13] Giorgio Parisi. A sequence of approximated solutions to the SK model for spin glasses. *Journal of Physics A: Mathematical and General*, 13(4):L115, 1980. DOI: [10.1088/0305-4470/13/4/009](https://doi.org/10.1088/0305-4470/13/4/009).
- [14] Michel Talagrand. The Parisi formula. *Annals of mathematics*, pages 221–263, 2006. DOI: [10.4007/annals.2006.163.221](https://doi.org/10.4007/annals.2006.163.221).
- [15] Dmitry Panchenko. The Sherrington-Kirkpatrick model: an overview. *Journal of Statistical Physics*, 149(2):362–383, 2012. DOI: [10.1007/s10955-012-0586-7](https://doi.org/10.1007/s10955-012-0586-7).
- [16] Dmitry Panchenko. *The Sherrington-Kirkpatrick model*. Springer Science & Business Media, 2013. DOI: [10.1007/978-1-4614-6289-7](https://doi.org/10.1007/978-1-4614-6289-7).
- [17] Antonio Auffinger, Wei-Kuo Chen, et al. Parisi formula for the ground state energy in the mixed  $p$ -spin model. *Annals of Probability*, 45(6B):4617–4631, 2017. DOI: [10.1214/16-AOP1173](https://doi.org/10.1214/16-AOP1173).
- [18] Dmitry Panchenko et al. The Parisi formula for mixed  $p$ -spin models. *Annals of Probability*, 42(3):946–958, 2014. DOI: [10.1214/12-AOP800](https://doi.org/10.1214/12-AOP800).
- [19] Leo Zhou, Edward Farhi, Jeffrey Goldstone, and Sam Gutmann. The quantum approximate optimization algorithm and the Sherrington-Kirkpatrick model at infinite size. Talk presented at QIP, 2021. URL <https://youtu.be/UP-Zuke7IUg>.