Maximum-Likelihood Detection with QAOA for Massive MIMO and Sherrington-Kirkpatrick Model with Local Field at Infinite Size

Burhan Gulbahar, *Senior Member, IEEE*

***Abstract*—Quantum-approximate optimization algorithm (QAOA) is promising in Noisy Intermediate-Scale Quantum (NISQ) computers with applications for NP-hard combinatorial optimization problems. It is recently utilized for NP-hard maximum-likelihood (ML) detection problem with fundamental challenges of optimization, simulation and performance analysis for $n \times n$ multiple-input multiple output (MIMO) systems with large n. QAOA is recently applied by Farhi et al. on infinite size limit of Sherrington-Kirkpatrick (SK) model with a cost model including only quadratic terms. In this article, we extend application of QAOA on SK model by including also linear terms and then realize SK modeling of massive MIMO ML detection by ensuring independence from specific problem instance and size n while preserving computational complexity of $O(16^p)$ designed by Farhi et al. We provide both optimized and extrapolated angles for $p \in [1, 14]$ and signal-to-noise (SNR) < 12 dB achieving near-optimum ML performance for 25×25 MIMO system with $p \geq 4$ in extensive simulations where 236500 different QAOA circuits are simulated. We present two conjectures about the concentration properties of QAOA and its near-optimum performance for massive MIMO systems with large sizes covering $n < 300$ promising significant performance for next generation massive MIMO systems.**

***Index Terms*—Quantum approximate optimization, massive MIMO, Sherrington-Kirkpatrick model, maximum-likelihood detection**

I. Introduction

Quantum-approximate optimization algorithm (QAOA) is a future promising quantum algorithm to be utilized on Noisy Intermediate-Scale Quantum (NISQ) devices [1] to provide high-performance solutions for NP-hard combinatorial optimization problems [2]. It is applied on a rich set of engineering problems with extensive theoretical and performance simulation studies especially for Max-Cut problems in graphs while promising quantum computational advantage (quantum supremacy) on NISQ computers [3]. QAOA is a hybrid method where the angle parameters of the layers of its quantum circuit with a depth count denoted by $p \geq 1$ are optimized in a classical computer while the solution is obtained among the measurement results of QAOA circuit running on a quantum computer. There are extensive theoretical and simulation studies for QAOA circuits with $p = 1$ calculating the optimum angles and expected cost explicitly while it becomes challenging to calculate for $p \geq 2$ [4]. Besides that, Sherrington-Kirkpatrick (SK) model denotes a spin system with all-to-all couplings among the spins while defining a combinatorial search problem to find the value of the lowest energy with a recent classical solution [5], [6]. Application of QAOA for general SK models including only quadratic terms is introduced by Farhi et al. in [5] where an algorithm for expectation calculation is presented with the computational complexity of $O(16^p)$ for the infinite size limit of SK model. In [7], a generalized multinomial theorem is provided to calculate expectations at the infinite size limit for various random ensembles including mixed spin models consisting of a cost function not only including local and quadratic terms but also higher order terms until $q \geq 2$ including clauses with more than two terms. The computational complexity is shown to be either $O(4^{q\,p})$ or $O(p^2\,4^p)$. In this article, we extend the algorithm in [5] for SK models which include linear terms in addition to quadratic terms and where the variances of quadratic and linear cost coefficients increase as $O(n)$ and $O(n^2)$, respectively, with respect to problem size n. Then, we apply QAOA with the extended SK model on NP-hard optimal maximum likelihood (ML) detection problem [8]. We observe near-optimum performance in extensive simulation studies with promising applications of QAOA in NISQ devices.

QAOA is recently applied for optimal ML detection problem especially for multiple-input-multiple-output (MIMO) communication systems as a low complexity alternative for the optimum ML detector for which the complexity increases exponentially with respect to symbol size n for a general $n \times n$ MIMO system with large number of users and transmitters [8]. Massive MIMO systems with optimum ML detection capability are highly promising to increase the performance of next generation communication systems compared with conventional solutions including minimum mean square error (MMSE) detection with low error performance and sphere decoding resulting in high complexity for large systems [9]. The number of studies analyzing the application of QAOA for ML detection is highly limited. In [8], ML detection problem is modeled as a quadratic unconstrained binary optimization (QUBO) problem and QAOA implementation is proposed as a solution for NP-hard ML detection problem with application for binary-phase-shift-keying (BPSK) MIMO systems. On the other hand, computational complexity of finding optimal angle values for QAOA is presented as $O(m\,p^5\,2^{3\,n})$ where n is the number of transmit antennas, p is the QAOA depth parameter and m is the number of iterations. Theoretical analysis is performed for $p \leq 3$, while simulations are performed for

This work was supported by TUBITAK (The Scientific and Technical Research Council of Turkey) under Grant #119E584. Burhan Gulbahar is with the Department of Electrical and Electronics Engineering, Yaşar University, 35100 Izmir, Turkey, e-mail: burhan.gulbahar@yasar.edu.tr.

significantly small scale systems with $n \leq 3$ and $p \leq 20$. It is not clear how to calculate expectation of QAOA circuit, to optimize angle parameters and to simulate performance for large n and high depth p as a significant challenge for the utilization of QAOA in practical ML detection problems. Similarly, in [10], proposed design is generalized for higher order modulation schemes as an extension of BPSK model. In [11], a theoretical analysis of QAOA expectation for $p = 1$ for channel decoding with arbitrary binary linear codes is presented and simulations are performed. In [12], QAOA is utilized for turbo detection in coded MIMO systems by defining a learning algorithm with recurrent neural networks to optimize parameters with simulations showing high performance.

On the other hand, there are high number of studies analyzing the application of quantum annealing (QA) and Ising machines for ML detection problems. In [9], Coherent Ising Machines (CIMs) are presented for performing near-optimal MIMO detection for MIMO systems as large as 64×256. The importance of realizing massive MIMO systems with near-optimal performance is emphasized for allowing both large number of users and antennas at the same time. In [13], a multi-stage version is presented for high order modulations and large systems. In [14], QA is utilized for vector perturbation precoding (VPP) in MIMO systems reaching sizes of 12×12. There are also additional studies including D-Wave based experimental tests for MIMO ML decoding in [15]–[17]. On the other hand, in [18], Grover search based quadratic speed-up is utilized in MIMO ML detection.

A. Contributions

In this article, we firstly include linear terms for the local fields in addition to quadratic terms of SK model by extending the algorithm in [5]. Then, we apply QAOA on SK modeling of ML detection problems for $n \times n$ massive MIMO systems. We present an optimization algorithm of QAOA angle parameters with significant advantages of independence from specific problem instance and n. The proposed algorithm is an extended version of the expectation calculation algorithm in [5] designed by Farhi et al. while preserving the computational complexity of $O(16^p)$. We present optimized angles for $p \in [1, 7]$ achieving near-optimum performance with $p \geq 4$ for 25×25 MIMO systems and SNR < 12 dB. Besides that, we propose a heuristic extrapolation algorithm for $p \in [8, 14]$ achieving high performance in simulation studies with the ratio of expected cost of QAOA to the optimum cost, i.e., denoted as approximation ratio conventionally, reaching ≈ 0.9364 [19]. In addition, we define additional approximation ratio performance metrics with respect to minimum measured cost among QAOA circuit outputs and observing almost unity ratios achieving ML performance with $p \geq 4$. We present two conjectures modeling the concentration properties and performance of QAOA for ML problems. Conjecture 1 is supported by extensive simulation studies for $n \leq 28$ while requiring theoretical and simulation based verifications for larger n values as an open issue. Conjecture 2 promises that presented angles for $p \in [1, 14]$ achieve near-optimum performance for massive MIMO symbol length $n < 300$.

We provide extensive simulation studies for QAOA in IBM Quantum Lab by analyzing statistical properties of QAOA measurement outputs for 236500 different QAOA circuits corresponding to random problem instances for varying n, p and SNR [20]. We share statistical properties for each problem instance in a data repository in [21] by creating an extensive data set for further analysis and verification by researchers. We theoretically analyze and simulate problem instance based mean and variance of measured minimum cost of QAOA circuits as an extension to the statistical properties of expected cost for SK models in [5]. Proposed solution promises widespread industrial utilization of QAOA in NISQ devices with error mitigation approaches [22]–[24] for near future massive MIMO wireless communications.

The remainder of the paper is organized as follows. Next, in Sections II and III, background about QAOA and application of QAOA for ML problems are presented. In Section IV, SK model of ML problem is given. In Section V, application of QAOA on SK model is extended by including linear terms for the local field. In Section VI, concentration properties of QAOA measurements and problem based expectations are discussed. Then, extensive simulation studies for QAOA circuits solving ML problem are presented in Section VII. Finally, in Section VIII, conclusions are presented.

II. Background about QAOA

QAOA is a promising to be utilized on NISQ devices for addressing combinatorial optimization problems [1], [2]. QAOA prepares a quantum state approximating the ground state of a problem Hamiltonian which corresponds to the optimal solution of a classical cost function $C(\boldsymbol{z})$ defined on n-bit strings $\boldsymbol{z} = (z_1, z_2, \ldots, z_n) \in \{+1, -1\}^n$ [5]. A typical cost function in a combinatorial optimization problem is represented in Ising form as follows [4]:

$$C(\boldsymbol{z}) = \sum_{j<k} J_{jk}\, z_j\, z_k \; + \; \sum_j h_j\, z_j \tag{1}$$

where $J_{j,k}$ values are the quadratic terms (interactions) and h_j terms are the linear terms (local fields). Besides that, SK model generalizes the Ising model of a classical spin system with couplings between all n spins while J_{jk} and h_j are chosen as independent Gaussian random variables [5]. Cost function is mapped onto Ising model Hamiltonian H_C defined as follows:

$$H_C = \sum_{j<k} J_{j,k}\, Z_j\, Z_k \; + \; \sum_j h_j\, Z_j \tag{2}$$

where Z_j are Pauli-Z operators acting on the j-th qubit. QAOA prepares a quantum state $|\Psi(\boldsymbol{\gamma}, \boldsymbol{\beta})\rangle$ such that expectation value of the problem Hamiltonian with respect to the quantum state, i.e., $\langle C \rangle$ or $\langle C_{QAOA} \rangle$, is minimized where the expectation of the cost is defined as follows:

$$\langle C \rangle = \langle \Psi(\boldsymbol{\gamma}, \boldsymbol{\beta}) | H_C | \Psi(\boldsymbol{\gamma}, \boldsymbol{\beta}) \rangle \tag{3}$$

State $|\Psi(\boldsymbol{\gamma}, \boldsymbol{\beta})\rangle \equiv |\Psi\rangle$ minimizing $\langle C \rangle$ is prepared as follows:

$$|\Psi\rangle = \prod_{r=1}^{p} U(B, \beta_r)\, U(C, \gamma_r)\; |+\rangle^{\otimes n} \tag{4}$$

$$= U(B, \beta_p)\, U(C, \gamma_p) \cdots U(B, \beta_1)\, U(C, \gamma_1)\; |+\rangle^{\otimes n} \tag{5}$$

where $B = \sum_{j=1}^{n} X_j$ is a mixing Hamiltonian, X_j is the Pauli-X operator acting on jth qubit, initial state $|+\rangle^{\otimes n} = (1/\sqrt{2^n}) \sum_{\boldsymbol{z}} |\boldsymbol{z}\rangle$ is the equal superposition state, and $U(B,\beta_r) = e^{-\imath \beta_r B}$ and $U(C,\gamma_r) = e^{-\imath \gamma_r H_C}$ are unitary operators parametrized by the optimized angle sets $\boldsymbol{\gamma} = (\gamma_1, \gamma_2, \ldots, \gamma_p)$ and $\boldsymbol{\beta} = (\beta_1, \beta_2, \ldots, \beta_p)$ while p is the layer depth of QAOA circuit. Angle parameters $\boldsymbol{\gamma}$ and $\boldsymbol{\beta}$ are optimized classically to maximize the expectation of the problem Hamiltonian [25]. The measurement of $|\Psi(\boldsymbol{\gamma},\boldsymbol{\beta})\rangle$ results in a string $\boldsymbol{z}$ such that $C(\boldsymbol{z})$ is close to $\langle C \rangle$.

III. Application of QAOA for ML Problems

On the other hand, QAOA is recently applied for optimum ML detection in MIMO systems in [8], [10] by also proposing QUBO model for the problem. Assume that $\boldsymbol{H}$ represents real valued $n \times n$ channel gain matrix known by the receiver for a specific instance of the estimated channel, $H(k,l)$ is the element of $\boldsymbol{H}$ at the kth row and lth column, $\boldsymbol{s} = [s_1 \ s_2 \ \ldots \ s_n]^T$ is the transmitted symbol of length n, $s_i \in \{-1, 1\}$ and $\boldsymbol{n} = [n_1 \ n_2 \ \ldots \ n_n]^T$ is the noise vector. Then, the received vector $\boldsymbol{y}$ is given as follows:

$$\boldsymbol{y} = \boldsymbol{H}\,\boldsymbol{s} + \boldsymbol{n} \tag{6}$$

Then, ML detection problem is defined as follows:

$$\underset{\boldsymbol{z}}{\text{minimize}} \quad ||\boldsymbol{y} - \mathbf{H}\,\boldsymbol{z}||^2 \tag{7}$$

In [8], the cost function is transformed into the following form:

$$||\boldsymbol{y} - \boldsymbol{H}\,\boldsymbol{z}||^2 = \sum_{j<k} J_{jk}\, z_j\, z_k + \sum_j h_j\, z_j + A \tag{8}$$

where J_{jk}, h_j and constant A are defined as follows:

$$J_{j,k} = 2 \sum_{l=1}^{n} H(l,j)\, H(l,k) \tag{9}$$

$$h_j = -\sum_{k=1}^{n} s_k\, J_{j,k} - 2 \sum_{l=1}^{n} n_l\, H(l,j) \tag{10}$$

$$A = \boldsymbol{y}^T \boldsymbol{y} + \sum_{l=1}^{n} \sum_{j=1}^{n} H^2(l,j) \tag{11}$$

and the corresponding quadratic term represented with the matrix $\boldsymbol{J}$ including the elements $J_{i,j}$ at ith row and jth column, and linear term represented with the column vector $\boldsymbol{h}$ with elements h_j at the jth row are defined as follows:

$$\boldsymbol{J} = 2\,\boldsymbol{H}^T \boldsymbol{H} \tag{12}$$

$$\boldsymbol{h} = -2\,\boldsymbol{H}^T \boldsymbol{y} \tag{13}$$

The performance of QAOA is analyzed in [8] for small length symbols, i.e., $n \leq 3$, and for specific instances of $\mathbf{H}$ by showing promising results for the application of QAOA for large problem instances which is one of the the main targets in this article. We not only provide optimization for large n but also provide instance independent solution working for all instances once the optimization is achieved classically for the general model of the problem based on statistical properties of channel matrix $\boldsymbol{H}$ and noise $\boldsymbol{n}$, SNR and layer depth p.

Classical calculation of $\langle C \rangle$ for large values of n is required to optimize $\boldsymbol{\gamma}$ and $\boldsymbol{\beta}$ without requiring a quantum computer to calculate the expectation. The classical calculation is provided explicitly for $p = 1$ in [4] while the complexity of the calculation is not clearly defined for $p \geq 2$ and for large values of n in the fully coupled model of the cost function in (8). Next, SK model of ML problem is defined which allows to calculate $\langle C \rangle$ for large values of n and large depth p for the typical instances of $\mathbf{H}$ with the elements of the matrix $\mathbf{H}$ represented as independent and identically distributed Gaussian random variables with zero mean.

IV. SK Model of ML Detection Problem

Conventional SK model as discussed in [5] includes the classical cost function with a scaled $J_{j,k}$ as follows:

$$C(\boldsymbol{z}) = \sum_{j<k} (J_{jk}/\sqrt{n})\, z_j\, z_k \tag{14}$$

where J_{jk} is a normal random variable. In [5], the expected value of $\langle C/n \rangle$ and $\langle (C/n)^2 \rangle$ as $n \to \infty$ with respect to the instances of $J_{j,k}$ are calculated as follows:

$$\lim_{n\to\infty} \mathbb{E}_J[\langle \Psi(\boldsymbol{\gamma},\boldsymbol{\beta})|\,(C/n)\,|\Psi(\boldsymbol{\gamma},\boldsymbol{\beta})\rangle] = V_p(\boldsymbol{\gamma},\boldsymbol{\beta}) \tag{15}$$

$$\lim_{n\to\infty} \mathbb{E}_J\left[\langle \Psi(\boldsymbol{\gamma},\boldsymbol{\beta})|\,(C/n)^2\,|\Psi(\boldsymbol{\gamma},\boldsymbol{\beta})\rangle\right] = (V_p(\boldsymbol{\gamma},\boldsymbol{\beta}))^2 \tag{16}$$

where $\langle C^a/n^a \rangle \equiv \langle \Psi(\boldsymbol{\gamma},\boldsymbol{\beta})|\,(C^a/n^a)\,|\Psi(\boldsymbol{\gamma},\boldsymbol{\beta})\rangle$ for $a \in [1,2]$, $\mathbb{E}_J[.]$ denotes the statistical expectation with respect to the instances of J_{jk} and explicit formula for $V_p(\boldsymbol{\gamma},\boldsymbol{\beta})$ is provided. It is conjectured that the variance of (C/n) goes to zero as $n \to \infty$ showing an important concentration property of the QAOA output strings. In other words, measurements of QAOA circuit output for large values of n will result in a string $\boldsymbol{z}$ with $C(\boldsymbol{z})/n$ highly close to $\langle \Psi(\boldsymbol{\gamma},\boldsymbol{\beta})|\,(C/n)\,|\Psi(\boldsymbol{\gamma},\boldsymbol{\beta})\rangle$ which is approximated as $V_p(\boldsymbol{\gamma},\boldsymbol{\beta})$ calculated explicitly. In this article, we exploit the SK model of ML problem in order to calculate statistical expectation of $\langle \Psi(\boldsymbol{\gamma},\boldsymbol{\beta})|\,C\,|\Psi(\boldsymbol{\gamma},\boldsymbol{\beta})\rangle$ for typical instances of ML problem. Besides that, in [7], a generalization of multinomial theorem is presented to calculate expectations at the infinite size limit for various ensembles of random combinatorial optimization problems including mixed spin models defined with the following cost function:

$$C(\boldsymbol{z}) \equiv \sum_{q=1}^{q_{max}} c_q \sum_{i_1,\ldots,i_q=1}^{n} J_{i_1,i_2,\ldots,i_q}\, z_{i_1} z_{i_2} \ldots z_{i_q} \tag{17}$$

where c_q for $q \in [1, q_{max}]$ are real variables and $J_{i_1,\ldots,i_q}$ is the cost tensor multiplying the expression $z_{i_1} z_{i_2} \ldots z_{i_q}$. The model generalizes [5] by including not only linear terms but also higher order terms compared with quadratic terms.

Since the elements of $\mathbf{H}$ are independent and identically distributed Gaussian random variables with zero mean, $J_{j,k}$ and h_j are approximated also as Gaussian random variables due to law of large numbers for large n. As a result, ML problem is modeled with SK model with the difference that it also includes the linear terms $h_j\, z_j$ compared with (14) and without direct normalization of $J_{j,k}$ and h_j as described next.

Assume that variances of real $H(l,k)$ for $k,l \in [1,n]$ and noise n_k for $k \in [1,n]$ are σ_H^2 and σ_n^2, respectively, and their

mean values are zero denoted with $\mu_H = \mu_n = 0$. SNR for the ℓth element of $\boldsymbol{y}$, i.e., y_ℓ, is calculated as follows:

$$SNR = \mathbb{E}_{\mathbf{H}}[\big(\sum_{k=1}^{n} H(\ell,k)s_k\big)^2] \,/\, \sigma_n^2 \;=\; n\,\sigma_H^2 \,/\, \sigma_n^2 \quad (18)$$

As proved in Section V, SK model with independent $J_{j,k}$ and h_j allows to calculate $\langle C\rangle$ with $O(16^p)$ complexity based on the extended version of the algorithm presented in [5]. The empirical mean and variance of $J_{j,k}$ and h_j are approximated as follows for $n \gg 1$ and typical transmitted BPSK symbols $\boldsymbol{s}$ with equally probable individual bits $s_i \in \{-1,+1\}$ for $i \in [1,n]$ as proved in Appendix A:

$$\mu_J = \mu_h \;\approx\; 0 \quad (19)$$

$$\sigma_J^2 \;\approx\; n\,(4\,\sigma_H^4) \equiv n\,\widetilde{\sigma}_J^2 \quad (20)$$

$$\sigma_h^2 \;\approx\; n^2\,\widetilde{\sigma}_J^2\left(2 - \frac{1}{n} + \frac{n-1}{n\,SNR}\right) \equiv n^2\,\widetilde{\sigma}_h^2 \quad (21)$$

where $\mu_J = \mathbb{E}_{\mathbf{H}}[J_{j,k}]$, $\mu_h = \mathbb{E}_{\mathbf{H},\boldsymbol{n}}[h_j]$, $\sigma_J^2 = \mathbb{E}_{\mathbf{H}}[(J_{j,k} - \mu_J)^2]$, $\sigma_h^2 = \mathbb{E}_{\mathbf{H},\boldsymbol{n}}[(h_j - \mu_h)^2]$, $\widetilde{\sigma}_J^2 = 4\,\sigma_H^4$ and $\widetilde{\sigma}_h^2 = 4\,\sigma_H^4\left(2 - 1\,/\,n + (n-1)\,/\,(n\,SNR)\right)$ with the observation that σ_J^2 and σ_h^2 increase as $O(n)$ and $O(n^2)$, respectively, with increasing $n \gg 1$. It is an open issue to extend the algorithmic framework in [5] to extend for correlated $\boldsymbol{J}$ and $\boldsymbol{h}$. The correlations are neglected in this article to exploit the described heuristic framework showing high performance.

A. QAOA Solution with SK Model

QAOA solution is simply realized by creating $\boldsymbol{J}$ and $\boldsymbol{h}$ for a given problem instance with known values of $\boldsymbol{H}$ and received vector $\boldsymbol{y}$ by using (12) and (13), respectively, and then realize p-layer QAOA circuit with angle parameters $\boldsymbol{\gamma}$ and $\boldsymbol{\beta}$ optimized with respect to SK model formed with zero mean $J_{j,k}$ and h_j having variances $n\,\widetilde{\sigma}_J^2$ and $n^2\,\widetilde{\sigma}_h^2$ as defined in (20) and (21), respectively. Then, we perform measurements of the implemented QAOA circuit outputs. We calculate the minimum cost corresponding to measured strings, i.e., denoted as $\min\{C_{QAOA}(\boldsymbol{z})\}$ as defined in the next section, to obtain the $\tilde{\boldsymbol{z}}$QAOA solution string $\boldsymbol{z}$ giving the minimum cost. In the next sections, we interchangeably use $\langle C^k\rangle$ or $\langle C^k_{QAOA}\rangle$ for $k \in [1,2]$ and $\mathbb{E}_{\boldsymbol{J},\boldsymbol{h}}[.]$ denotes expectation with respect to varying problem instances with different $\boldsymbol{J}$ and $\boldsymbol{h}$.

V. QAOA for SK Model with Linear Term

We extend the moment calculations of SK model at infinite size limit in [5] for the cost function in (1) including the local field term h_j in addition to the quadratic term $J_{j,k}$. The algorithm in [5] allows extension with a minor modification for the following differences of the problem model. We follow a similar method in [5], i.e., extended version of the equation (76) in [5], while with the following fundamental differences:

1) $\langle \lambda\, C\,/\,n^2\rangle$ and $\langle \lambda\, C^2\,/\,n^4\rangle$ are calculated instead of $\langle \lambda\, C\,/\,n\rangle$ and $\langle \lambda\, C^2\,/\,n^2\rangle$.
2) Linear h_j terms are included in the cost function with the assumption that $J_{j,k}$ and h_j are zero mean linearly independent and identically distributed Gaussian random variables with variances σ_J^2 and σ_h^2 increasing as $O(n)$ and $O(n^2)$ with respect to n, respectively.
3) Optimized angles $\boldsymbol{\gamma}$ decrease with $1\,/\,n$ as $\boldsymbol{\gamma} = \widetilde{\boldsymbol{\gamma}}\,/\,n$ for optimized $\widetilde{\boldsymbol{\gamma}}$ with elements $\widetilde{\gamma}_{p,r} \in [0, 2\,\pi]$ for $r \in [1,p]$. We include additional term p in the subscripts of γ_r, $\widetilde{\gamma}_r$ and β_r to better analyze in simulation studies.

Algorithm 1 Calculation of $\widetilde{W}_{\boldsymbol{u}}$ for $\boldsymbol{u} \in A$ (Local field included extension of the algorithm in [5] for SK model)

1: Define $A = \{+1,-1\}^{2p}$
2: Partition A into $A = A_{p+1} \cup D \cup D^c$ based on grouping definition such that $\boldsymbol{b} \in D$ and $\overline{\boldsymbol{b}} \in D^c$.
3: Index the strings in $D = \{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_{|D|}\}$ and $D^c = \{\overline{\boldsymbol{u}}_1, \overline{\boldsymbol{u}}_2, \ldots, \overline{\boldsymbol{u}}_{|D|}\}$ where the number of elements in D is $|D| = (2^{2\,p} - 2^p)\,/\,2$ such that if the group index of $\boldsymbol{u}_j$ is denoted with ℓ_j, then $\ell_j \le \ell_k$ if $j \le k$.
4: **for** $j \leftarrow |D|$ to 1 (starting with $j = |D|$ and decreasing to $j = 1$) **do**
5: Calculate $\widetilde{X}_{\boldsymbol{u_j}}$ and $\widetilde{\Delta}_{\boldsymbol{u_k}.\boldsymbol{u_j}}$ for $k > j$ with (27, 28).
6: $\widetilde{W}_{\boldsymbol{u_j}} = \widetilde{X}_{\boldsymbol{u_j}} \exp\left(\sum_{k>j} \widetilde{W}_{\boldsymbol{u_k}}\,\widetilde{\Delta}_{\boldsymbol{u_k}.\boldsymbol{u_j}}\right)$
7: $\widetilde{W}_{\overline{\boldsymbol{u}}_j} = -\widetilde{W}_{\boldsymbol{u_j}}$
8: **end for**
9: **for** $\boldsymbol{a} \in A_{p+1}$ **do**
10: $\widetilde{W}_{\boldsymbol{a}} = Q_{\boldsymbol{a}}$
11: **end for**

On the other hand, the generalized algorithm for mixed spin model of order q in [7] also assumes that the variance of $J_{i_1,i_2,\ldots,i_q}$ equals to $1\,/\,n^{q-1}$. This assumption is also different compared with our case that $\sigma_J^2 = n\,\widetilde{\sigma}_J^2$ and $\sigma_h^2 = n^2\,\widetilde{\sigma}_h^2$ increasing as $O(n^{3-q})$ with n. We provide the next theorem as the extended version of [5]:

Theorem 1. *Given the angles $\gamma_{p,r} = \widetilde{\gamma}_{p,r}\,/\,n$ and $\beta_{p,r}$ for $r \in [1,p]$ for QAOA circuit and the problem Hamiltonian $H_C = \sum_{j=1}^{n-1}\sum_{k=j+1}^{n} J_{j,k}\,Z_j\,Z_k + \sum_j h_j\,Z_j$ where $J_{j,k}$ and h_j are independent zero mean Gaussian random variables with variances $\sigma_J^2 = n\,\widetilde{\sigma}_J^2$ and $\sigma_h^2 = n^2\,\widetilde{\sigma}_h^2$ for constant $\widetilde{\sigma}_J^2$ and $\widetilde{\sigma}_h^2$, respectively, the following equalities hold:*

$$\lim_{n\to\infty} \mathbb{E}_{\boldsymbol{J},\boldsymbol{h}}\left[\langle C/n^2\rangle\right] = \widetilde{V}_p(\widetilde{\boldsymbol{\gamma}},\boldsymbol{\beta}) \quad (22)$$

$$\lim_{n\to\infty} \mathbb{E}_{\boldsymbol{J},\boldsymbol{h}}\left[\langle C^2/n^4\rangle\right] = \widetilde{V}_p^2(\widetilde{\boldsymbol{\gamma}},\boldsymbol{\beta}) \quad (23)$$

where $\widetilde{V}_p(\widetilde{\boldsymbol{\gamma}},\boldsymbol{\beta})$ is defined as follows:

$$\widetilde{V}_p(\widetilde{\boldsymbol{\gamma}},\boldsymbol{\beta}) = \frac{\imath\,\widetilde{\sigma_J}^2}{2}\sum_{r=1}^{p}\widetilde{\gamma}_{p,r}\,\Gamma_r^+\,\Gamma_r^- + \imath\,\widetilde{\sigma_h}^2\sum_{r=1}^{p}\widetilde{\gamma}_{p,r}\,\Gamma_r^- \quad (24)$$

where $\Gamma_r^{\pm} \equiv \sum_{\boldsymbol{u}\in A}(u_{r:p} \pm u_{-r:-p})\,\widetilde{W}_{\boldsymbol{u}}$, $A = \{+1,-1\}^{2p} = \{\boldsymbol{a} = (a_1, a_2, \ldots, a_p, a_{-p}, \ldots, a_{-2}, a_{-1}) : a_{\pm j} \in \{+1,-1\}\}$ is the set of $2^{2\,p}$ strings and $\widetilde{W}_{\boldsymbol{u}}$ is calculated in Algorithm 1. Furthermore, $\mathbb{E}_{\boldsymbol{J},\boldsymbol{h}}\left[\langle C\,/\,n^2\rangle\right]$ and $\mathbb{E}_{\boldsymbol{J},\boldsymbol{h}}\left[\langle C^2\,/\,n^4\rangle\right]$ for large n include additional terms decreasing with $O(1\,/\,n)$.

Proof. Proof is presented in Appendix B. □

As described in [5], the configuration basis A is defined as follows such that kth index bit $(z_k^{[1]}, z_k^{[2]} \ldots, z_k^{[p]}, z_k^{[-p]}, \ldots, z_k^{[-2]}, z_k^{[-1]}) \in A$:

$$A = \{+1,-1\}^{2p} = \{\boldsymbol{a} = (a_1, \ldots, a_p, a_{-p}, \ldots, a_{-1})\} \quad (25)$$

where $a_{\pm j} \in \{+1,-1\}$, $\boldsymbol{z}^{[\pm r]} = \left(z_1^{[\pm r]}, z_2^{[\pm r]}, \ldots, z_n^{[\pm r]}\right) \in \{+1,-1\}^n$ denotes n-bit strings corresponding to rth layer for $r \in [1,p]$. It is expressed that $A = A_1 \cup \cdots \cup A_p \cup A_{p+1}$ where A_ℓ for $\ell \in [1,p]$ and A_{p+1} are defined such that A_ℓ includes elements where $a_{-k} = a_k$ for $p-\ell+1 < k \leq p$ and $a_{-p+\ell-1} = -a_{p-\ell+1}$. A_{p+1} is such that $a_{-k} = a_k$ for $1 \leq k \leq p$. On the other hand, the $\overline{\boldsymbol{a}}$ operation is defined in [5] as follows for $\boldsymbol{a} \in A_\ell$ for $1 \leq \ell \leq p$:

$$\bar{a}_{\pm r} = \begin{cases} a_{\pm r}, & r \neq p-\ell+1 \\ -a_{\pm r} & r = p-\ell+1 \end{cases}. \quad (26)$$

where $\overline{\boldsymbol{a}} \in A_\ell$ without changing the group, i.e., both $\boldsymbol{a}$ and $\overline{\boldsymbol{a}}$ are in A_ℓ. Furthermore, a group excluding A_{p+1} is defined as $B \equiv A \backslash A_{p+1}$ and it is partitioned into two sets $D \cup D^c$ distributing $\boldsymbol{b}$ and $\overline{\boldsymbol{b}}$ into D and D^c, respectively. The number of elements in D is equal to $\left(2^{2p} - 2^p\right)/2$. Indices are defined for the strings in $D = \{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_{|D|}\}$ and $D^c = \{\overline{\boldsymbol{u}}_1, \overline{\boldsymbol{u}}_2, \ldots, \overline{\boldsymbol{u}}_{|D|}\}$ such that if the group index of $\boldsymbol{u}_j$ is denoted with ℓ_j, then $\ell_j \leq \ell_k$ if $j \leq k$. Then, $\widetilde{W}_{\boldsymbol{u}}$ for $\boldsymbol{u} \in A$ is calculated in Algorithm 1 based on the iterative formulation in page 26 in Section 6.2 in [5] with the main difference of replacing Q_b with $\widetilde{Q}_{\boldsymbol{b}} = Q_{\boldsymbol{b}}\, g_{\boldsymbol{b}}^{1/2}$ where Q_b and $g_{\boldsymbol{b}}$ are defined in (56) and (60), respectively. Some other variables utilized in Algorithm 1 are described as follows based on the extension of the iterative formulation in [5]:

$$\widetilde{X}_{\boldsymbol{u}_j} = \exp\left(-\frac{\widetilde{\sigma}_J^2}{2} \sum_{\boldsymbol{a} \in A_{p+1}} Q_{\boldsymbol{a}}\, \widetilde{\Phi}_{\boldsymbol{a}.\boldsymbol{u}_j}^2\right) Q_{\boldsymbol{u}_j}\, g_{\boldsymbol{u}_j}^{1/2} \quad (27)$$

$$\widetilde{\Delta}_{\boldsymbol{u}_k.\boldsymbol{u}_j} = \frac{\widetilde{\sigma}_J^2}{2}\left(\widetilde{\Phi}_{\overline{\boldsymbol{u}}_k.\boldsymbol{u}_j}^2 - \widetilde{\Phi}_{\boldsymbol{u}_k.\boldsymbol{u}_j}^2\right) \quad (28)$$

where $\widetilde{\Phi}_{\boldsymbol{a}} = \sum_{r=1}^p \widetilde{\gamma}_{p,r}\left(a_{r:p} - a_{-r:-p}\right)$.

VI. Concentration Properties of QAOA

Theorem 1 in this article and 'Concentration Corollary' in Section 4 in [5] utilizing Chebyshev inequality result in the following concentration properties as $n \to \infty$:

$$\mathbb{P}_{\boldsymbol{J},\boldsymbol{h}}\left(\left|\langle C/n^2\rangle - \widetilde{V}_p(\widetilde{\boldsymbol{\gamma}},\boldsymbol{\beta})\right| > \epsilon\right) \to 0 \quad (29)$$

$$\mathbb{E}_{\boldsymbol{J},\boldsymbol{h}}\left[\mathbb{P}_{\mathrm{Q}}(|C(\boldsymbol{z})/n^2 - \langle \boldsymbol{\gamma},\boldsymbol{\beta}|C/n^2|\boldsymbol{\gamma},\boldsymbol{\beta}\rangle| > \epsilon)\right] \to 0 \quad (30)$$

where $\mathbb{P}_{\mathrm{Q}}$ denotes probability with respect to QAOA measurement, The meaning of (29) is that for typical instances of the SK model with local field and for large values of n, QAOA measurements result in a symbol string $\boldsymbol{z}$ such that $C(\boldsymbol{z})/n^2 \approx \langle C\,/\,n^2\rangle \approx \widetilde{V}_p(\widetilde{\boldsymbol{\gamma}},\boldsymbol{\beta})$. Therefore, Theorem 1 allows the following approximations for large n:

$$\mathbb{E}_{\boldsymbol{J},\boldsymbol{h}}\left[\langle C\rangle\right] \approx n^2\, \widetilde{V}_p(\widetilde{\boldsymbol{\gamma}},\boldsymbol{\beta}) + O(n) \quad (31)$$

$$\mathbb{E}_{\boldsymbol{J},\boldsymbol{h}}\left[\langle C^2\rangle\right] \approx n^4\, \widetilde{V}_p^2(\widetilde{\boldsymbol{\gamma}},\boldsymbol{\beta}) + O(n^3) \quad (32)$$

Expected cost with a measured string at the QAOA circuit for typical instances of SK problem, e.g., in our case ML problem, is very close to $n^2\, \widetilde{V}_p(\widetilde{\boldsymbol{\gamma}},\boldsymbol{\beta})$ independent from the specific instance of the problem. Besides that, $\sigma_{\langle C\rangle} \equiv \sqrt{\mathbb{E}_{\boldsymbol{J},\boldsymbol{h}}\left[\langle C\rangle^2\right] - \left(\mathbb{E}_{\boldsymbol{J},\boldsymbol{h}}\left[\langle C\rangle\right]\right)^2}$ is bounded as follows:

$$\sigma_{\langle C\rangle} \leq \sqrt{\mathbb{E}_{\boldsymbol{J},\boldsymbol{h}}\left[\langle C^2\rangle\right] - \left(\mathbb{E}_{\boldsymbol{J},\boldsymbol{h}}\left[\langle C\rangle\right]\right)^2} \approx O(n^3) \quad (33)$$

where $\langle O\rangle^2 \leq \langle O^2\rangle$ is utilized for operators [5]. This means that as the mean $\mathbb{E}_{\boldsymbol{J},\boldsymbol{h}}\left[\langle C\rangle\right]$ increases as $O(n^2)$, the standard deviation for various instances of $\boldsymbol{J}$ and $\boldsymbol{h}$ increases with a smaller rate of $O(n\sqrt{n})$ resulting in concentration property. On the other hand, we observe in our simulations in Section VII that instance based mean of the distance between measurement mean, i.e., $\langle C_{QAOA}\rangle$, and the minimum measured value, i.e., $\min_{\boldsymbol{z}}\{C_{QAOA}(\boldsymbol{z})\}$, shows $O(g_p(n)\, n^2)$ increase with increasing n as presented in the following conjecture:

Conjecture 1. *Problem instance based expectation of the difference between the measured mean cost and minimum cost denoted by $\langle C_{QAOA}\rangle$ and $\min_{\boldsymbol{z}}\{C_{QAOA}(\boldsymbol{z})\}$, respectively, for QAOA circuit measurements of ML problem increases as $O\left(g_p(n)\, n^2\right)$ with increasing $n \gg 1$:*

$$E_{J,h}\left[< C_{QAOA} > - \min_{\boldsymbol{z}}\{C_{QAOA}(\boldsymbol{z})\}\right] = g_p(n)\, n^2 \quad (34)$$

where $g_p(n) > 1/\sqrt{n}$ is a slowly decreasing function of n for each p. Besides that, instance based variance of $\min_{\boldsymbol{z}}\{C_{QAOA}(\boldsymbol{z})\} \equiv m_Q$ increases as $O(n^3)$ leading to the following equality for some constant c_p depending on p:

$$E_{J,h}\left[m_Q^2\right] - E_{J,h}^2\left[m_Q\right] = c_p\, n^3 \quad (35)$$

The supporting evidence is provided in simulations in Section VII. Although standard deviation of $\min_{\boldsymbol{z}}\{C_{QAOA}(\boldsymbol{z})\}$ with respect to problem instances increases as $O(n\sqrt{n})$, it is conjectured that distance between $\langle C_{QAOA}\rangle$ and $\min_{\boldsymbol{z}}\{C_{QAOA}(\boldsymbol{z})\}$ increases with a larger rate of $O(g_p(n)\, n^2)$ compared with $O(n\sqrt{n})$ satisfying $g_p(n)\, n^2 > n\sqrt{n}$ at least until some large value of n depending on p. We observe in simulations that variances in (33) and (35) increase as $O(n^3)$ supporting Theorem 1 and conjectures. It is an open issue to determine the probability distribution type for ML problems while QAOA measurements are observed to have approximate Boltzmann distribution in [26], [27].

We assume that ML cost and solution cost are close to each other for high SNR as shown in Section VII. Based on our observations in simulation studies, we assume distribution of $\langle C_{QAOA}\rangle$ with respect to varying problem instances has a Gaussian form as a simplified assumption. For simplicity, if $\min_{\boldsymbol{z}}\{C_{QAOA}(\boldsymbol{z})\}$ is assumed to have mean $\langle C_{QAOA}\rangle$ with a variance denoted by $\sigma_{QAOA}^2(p)$ combining the variance of distributions of $\langle C_{QAOA}\rangle$ with respect to $\boldsymbol{J},\boldsymbol{h}$ and distribution of the measurements for an individual problem instance, then probability of $\min_{\boldsymbol{z}}\{C_{QAOA}(\boldsymbol{z})\} < C(\boldsymbol{s})$ with respect to instances $\boldsymbol{J}$ and $\boldsymbol{h}$ becomes proportional to the following:

$$\propto \Phi\left(\mathbb{E}_{\boldsymbol{J},\boldsymbol{h}}\left[C(\boldsymbol{s}) - \langle C_{QAOA}\rangle\right]/\,\sigma_{QAOA}(p)\right) \quad (36)$$

where $\sigma_{QAOA}^2(p)$ increases as $O(n^3)$ with increasing n based on observations in simulations and $\Phi(.)$ is the cumulative

distribution function of a normal random variable. We heuristically define a variance combining the variances with respect to both all the problem instances and individual measurement outcomes of a single QAOA circuit as follows:

$$\sigma^2_{QAOA}(p) \equiv \mathbb{E}_{\boldsymbol{J},\boldsymbol{h}}\left[\langle C_{QAOA}\rangle^2\right] - \mathbb{E}^2_{\boldsymbol{J},\boldsymbol{h}}\left[\langle C_{QAOA}\rangle\right] + \mathbb{E}_{\boldsymbol{J},\boldsymbol{h}}\left[\langle C^2_{QAOA}\rangle - \langle C_{QAOA}\rangle^2\right] \quad (37)$$

$$= \mathbb{E}_{\boldsymbol{J},\boldsymbol{h}}\left[\langle C^2_{QAOA}\rangle\right] - \mathbb{E}^2_{\boldsymbol{J},\boldsymbol{h}}\left[\langle C_{QAOA}\rangle\right] \quad (38)$$

This definition allows us to include the effects of the variance of the mean $\langle C_{QAOA}\rangle$ with respect to different problem instances, i.e., $\mathbb{E}_{\boldsymbol{J},\boldsymbol{h}}\left[\langle C_{QAOA}\rangle^2\right] - \mathbb{E}^2_{\boldsymbol{J},\boldsymbol{h}}\left[\langle C_{QAOA}\rangle\right]$, and variance of the QAOA circuit output measurement for each individual instance, i.e., $\mathbb{E}_{\boldsymbol{J},\boldsymbol{h}}\left[\langle C^2_{QAOA}\rangle - \langle C_{QAOA}\rangle^2\right]$. Then, $\widehat{n}_{max}$ is approximated by setting input to $\Phi(.)$ function in (36) to a constant value $\sqrt{\alpha_2(r^*)}$ depending on the target approximation ratio r^* resulting in the following:

$$\frac{\left|\mathbb{E}_{\boldsymbol{J},\boldsymbol{h}}\left[C(\boldsymbol{s}) - \langle C_{QAOA}\rangle\right]\right|}{\sigma_{QAOA}(p)} = A(p)\sqrt{\widehat{n}_{max}} = \sqrt{\alpha_2(r^*)} \quad (39)$$

where both $C(\boldsymbol{s})$ and $\langle C_{QAOA}\rangle$ increase as $O(n^2)$ and $\sigma_{QAOA}(p)$ increases as $O(n\sqrt{n})$ such that left hand side of (39) is proportional to $O(\sqrt{\widehat{n}_{max}})$, $\sqrt{\alpha_2(r^*)}$ is the performance constant for r^* and $A(p)$ is calculated as follows:

$$A(p) = \frac{\lim_{n\to\infty}\left|\mathbb{E}_{\boldsymbol{J},\boldsymbol{h}}\left[\langle C_{QAOA}\rangle - C(\boldsymbol{s})\right]\right| / n^2}{\lim_{n\to\infty}(\sigma_{QAOA}(p)\,/\,n^{3/2})} \quad (40)$$

Then, $\widehat{n}_{max} \propto \alpha_2(r^*)\,/\,A^2(p)$ is obtained. Substituting (31) and (38) into (39) and including additional constant for fitting result in the following:

Conjecture 2. *The maximum value of n denoted by $\widehat{n}_{max}$ for p-layer QAOA circuits achieving near-optimum ML performance for high SNR (≥ 12 dB) is approximated as follows:*

$$\widehat{n}_{max}(r^*, p) \approx \alpha_1(r^*) + \alpha_2(r^*)\,\chi(p) \quad (41)$$

where $\alpha_1(r^*)$ and $\alpha_2(r^*)$ are empirical approximation constants depending on the target approximation ratio r^* and $\mathbb{E}_{\boldsymbol{J},\boldsymbol{h}}\left[\langle C_{QAOA}\rangle\right]$ and $\mathbb{E}_{\boldsymbol{J},\boldsymbol{h}}\left[\langle C^2_{QAOA}\rangle\right]$ are problem instance based expectations of $\langle C^m_{QAOA}\rangle$ for $m \in [1,2]$ corresponding to first and second moments of the measured costs for p-layer QAOA circuits and $\chi(p)$ is defined as follows:

$$\chi(p) = \frac{\lim_{n\to\infty}\left(\mathbb{E}_{\boldsymbol{J},\boldsymbol{h}}\left[\langle C^2_{QAOA}\rangle\right] - \mathbb{E}^2_{\boldsymbol{J},\boldsymbol{h}}\left[\langle C_{QAOA}\rangle\right]\right) / n^3}{\left(\lim_{n\to\infty}\mathbb{E}_{\boldsymbol{J},\boldsymbol{h}}\left[\langle C_{QAOA}\rangle - C(\boldsymbol{s})\right] / n^2\right)^2} \quad (42)$$

The proofs and verification of both conjectures require to obtain QAOA performance for large n which is beyond our capability with access to public simulators with qubit numbers smaller than 32. Extensive simulation results show that $\widehat{n}_{max} \approx 28$ for $p = 4$ and SNR $= 15 \approx 12$ dB with near-optimum performance.

VII. QAOA Simulations

We provide extensive simulations by using IBM Quantum Lab and "*ibmq_qasm_simulator*" back-end by *IBM Quantum* [20]. Simulation parameters are presented in Table I. The number of instances denoted by M is the number of realizations of the main ML problem in (6) with randomly generated real valued $\mathbf{H}$, symbols $\boldsymbol{s} \in \{-1,+1\}^{\otimes n}$ and noise $\boldsymbol{n}$ in each instance. Elements of $\mathbf{H}$ and $\boldsymbol{n}$ are generated with variances $\sigma^2_H = 1$ and $\sigma^2_n = n\,\sigma^2_H\,/\,SNR = n\,/\,SNR$, respectively, for each instance. Therefore, the number of instances is also equal to the total number of estimated symbols with length n allowing to calculate total probability of bit error with $P_e = N_b\,/\,(M\,n)$ where N_b is the total number of bit errors in all M symbols. $\boldsymbol{J}$ and $\boldsymbol{h}$ are calculated with (9) and (10) and then QAOA circuit output is simulated. The number of measurements (shots) obtained in each simulation instance is chosen as 4096. We simulated a total of 236500 QAOA circuits for varying n, p and SNR by providing an extensive result set for the performance of QAOA for ML problems. We share measurement results for $\langle C_{QAOA}\rangle$, $\langle C^2_{QAOA}\rangle$, $\min_{\boldsymbol{z}}\{C_{QAOA}(\boldsymbol{z})\}$, $C(\boldsymbol{s})$ and n_b for each random problem instance in [21] as a data repository to be utilized in research studies of QAOA where n_b denotes the total number of bit errors in a single problem instance among n bits.

We simulate QAOA performance for varying symbol lengths $n \in [10, 28]$, SNR $\in \{1, 5, 10, 15\}$ and circuit depth $p \in [1, 14]$ as shown in Table I. The maximum value of SNR is set to $15 \approx 12$ dB due to the requirement of very high M to reliably calculate P_e. We achieve $4000 \leq M \leq 5000$ for the cases requiring accurate estimation of P_e such as for the combinations of $p \in [4, 6]$, $SNR = 15$ and $n = 25$ or for varying n to accurately observe the highest performance value $\widehat{n}_{max}$ achieving ML performance in Section VII-E.

Various performance metrics of QAOA are defined. $\langle C\rangle \equiv \langle C_{QAOA}\rangle$, $\langle C^2\rangle \equiv \langle C^2_{QAOA}\rangle$ and $\min_{\boldsymbol{z}}\{C_{QAOA}(\boldsymbol{z})\} \equiv \min_{\boldsymbol{z}}\{C_{QAOA}(\boldsymbol{z})\}$ correspond to expectation of the cost $C(\boldsymbol{z})$, $C^2(\boldsymbol{z})$ and the minimum cost, respectively, based on the observed measurement outputs $\boldsymbol{z}$ of QAOA circuit in each simulation instance. We calculate mean of all these expectations with respect to all M different problem instances with different $\mathbf{H}$, $\boldsymbol{s}$, $\boldsymbol{n}$, $\boldsymbol{J}$ and $\boldsymbol{h}$ by using the expectation operator defined as $E_{\boldsymbol{J},\boldsymbol{h}}[.]$. The values $\mathbb{E}_{\boldsymbol{J},\boldsymbol{h}}\left[\langle C_{QAOA}\rangle^2\right] - \mathbb{E}^2_{\boldsymbol{J},\boldsymbol{h}}\left[\langle C_{QAOA}\rangle\right]$ and $\mathbb{E}_{\boldsymbol{J},\boldsymbol{h}}\left[\left(\min_{\boldsymbol{z}}\{C_{QAOA}(\boldsymbol{z})\}\right)^2\right] - \mathbb{E}^2_{\boldsymbol{J},\boldsymbol{h}}\left[\min_{\boldsymbol{s}}\{C_{QAOA}(\boldsymbol{s})\}\right]$ denote statistical variances of $\langle C_{QAOA}\rangle$ and $\min_{\boldsymbol{z}}\{C_{QAOA}(\boldsymbol{z})\}$. On the other hand, $\mathbb{E}_{\boldsymbol{J},\boldsymbol{h}}\left[\langle C^2_{QAOA}\rangle - \langle C_{QAOA}\rangle^2\right]$ denotes the instance based mean of the variance of C_{QAOA} with respect to measured costs in each problem instance.

$C(\boldsymbol{s})$ corresponds to the solution cost for the problem with the transmitted symbol $\boldsymbol{s}$ while C_{ML} and C_{MMSE} denote the costs for ML and MMSE solutions in each problem instance. We compare QAOA performance with both ML and MSSE solutions. MMSE solution is defined as follows [28]:

$$\widehat{\boldsymbol{s}} = f_{sign}\left((\mathbf{H}^T\,\mathbf{H} + \sigma^2_n\,\mathbf{I})^{-1}\,\mathbf{H}^T\,\boldsymbol{y}\right) \quad (43)$$

where $f_{sign}(.)$ is the sign function since we utilize BPSK symbols with equally probable $\{\pm 1\}$ values. We obtain cost values of ML solution for $n = 25$ with classical simulation for the number of simulation instances being equal to 1500, 2500, 5000 and 5000 for SNR $\in \{1, 5, 10, 15\}$, respectively. P_e is

TABLE I
SIMULATION PARAMETERS OF QAOA (M: THE NUMBER OF PROBLEM INSTANCES)

SNR	p	n	M	SNR	p	n	M	SNR	p	n	M
15	1	$\{10,12,13,14\}$	4000	15	4	$\{18,20\}$	2000	15	$\in[1,3]$	25	1500
		$\{16,18,20,22\}$	2000			$\{14,22,23,24\}$	4000		$\in[4,14]$		5000
		$\{24,26,28\}$	1000			$\{10,12,16,26,28\}$	2500				
SNR	p	n	M	SNR	p	n	M	SNR	p	n	M
15	2	$\{10,12,13,14,15,16,18,19\}$	4000	15	3	$\{14,16,18,19,20,21\}$	4000	$\in\{1,5\}$	$\in[1,7]$	25	1000
		$\{20,22,24\}$	2000			$\{10,12,13,22,23,24\}$	3000	10	$\{1,2,3,7\}$		
		$\{26,28\}$	1000			$\{26,28\}$	1250	10	$\in[4,6]$		5000

TABLE II
INITIAL POINTS FOR ANGLE OPTIMIZATION OF QAOA FOR $p\in[1,6]$

p	Initial point boundaries of $(\widetilde{\gamma}_{p,p},\beta_{p,p})$	Grid dimensions
$[1,3]$	$([0,2\,\pi],[0,2\,\pi])$	100×100
4	$([0,2\,\pi],[0.5\,\beta_{3,3},1.5\,\beta_{3,3}])$	100×25
5	$([0,\pi\,/\,10],[0.85\,\beta_{4,4},1.15\,\beta_{4,4}])$	15×16
6	$([0.9\,\widetilde{\gamma}_{5,5},1.1\,\widetilde{\gamma}_{5,5}],[0.9\,\beta_{5,5},1.1\,\beta_{5,5}])$	20×20

calculated by using 10000 instances for $n\in[10,26]$ and SNR $=15$. MMSE performances are calculated with $M=5000$.

We define various approximation ratios in addition to the conventional approximation ratio r_{ML} as the ratio of the expected cost obtained for a specific problem instance to the minimum possible cost. We revise the definition by using expectations with respect to varying problem instances. We also utilize normalization for a smooth performance comparison as a precaution due to finite number of simulation instances and difference in the number of simulations among various p and SNR values. The following are defined:

$$r \equiv \mathbb{E}_{\boldsymbol{J},\boldsymbol{h}}\left[\langle C_{QAOA}\rangle\,/\,C(\boldsymbol{s})\right] \tag{44}$$

$$r^* \equiv \mathbb{E}_{\boldsymbol{J},\boldsymbol{h}}\left[\min_{\boldsymbol{z}}\{C_{QAOA}(\boldsymbol{z})\}\,/\,C(\boldsymbol{s})\right] \tag{45}$$

$$r_{ML} \equiv \frac{\mathbb{E}_{\boldsymbol{J},\boldsymbol{h}}\left[\langle C_{QAOA}\rangle\,/\,C(\boldsymbol{s})\right]}{\mathbb{E}_{\boldsymbol{J},\boldsymbol{h}}\left[C_{ML}\,/\,C(\boldsymbol{s})\right]} \tag{46}$$

$$r^*_{ML} \equiv \frac{\mathbb{E}_{\boldsymbol{J},\boldsymbol{h}}\left[\min_{\boldsymbol{z}}\{C_{QAOA}(\boldsymbol{z})\}\,/\,C(\boldsymbol{s})\right]}{\mathbb{E}_{\boldsymbol{J},\boldsymbol{h}}\left[C_{ML}\,/\,C(\boldsymbol{s})\right]} \tag{47}$$

where we normalize costs for the definitions of r_{ML} and r^*_{ML} by dividing with $C(\boldsymbol{s})$. r and r^* measure performance of QAOA with respect to the solution cost while r_{ML} and r^*_{ML} measure with respect to the performance of optimum ML solution which is the conventional performance metric. We achieve approximation ratio $r\approx r_{ML}\approx 0.9364$ with $p=14$ for SNR ≈ 12 dB while $r^*\approx 1$ with $p\geq 5$ for $SNR\geq 10$. In addition, QAOA solution is highly close to ML solution with $r^*_{ML}\approx 1$ with $p\geq 5$ for $SNR\geq 5$.

A. Optimization of QAOA angle parameters

We optimize angles $\widetilde{\boldsymbol{\gamma}}=n\,\boldsymbol{\gamma}$ and $\boldsymbol{\beta}$ for $p\in[1,7]$ based on (31) by minimizing $\widetilde{V}_p(\widetilde{\boldsymbol{\gamma}},\boldsymbol{\beta})$ while obtaining $\widetilde{V}_p(\widetilde{\boldsymbol{\gamma}},\boldsymbol{\beta})$ for each $\boldsymbol{\gamma}$ and $\boldsymbol{\beta}$ with substitution of $\widetilde{W}_{\boldsymbol{u}}$ into (24) while $\widetilde{W}_{\boldsymbol{u}}$ is calculated by using Algorithm 1. The optimization is performed in classical computer by providing initial points as shown in Table II and then applying unconstrained nonlinear optimization with quasi-Newton method as a type of gradient based methods available in various numerical packages. $(\widetilde{\gamma}_{1,1},\beta_{1,1})$ region covering an area of $(2\pi\times2\pi)$ is discretized into a 100×100 grid for $p=1$ by determining 100×100 initial points where $\widetilde{\gamma}_{p,j}$ and $\beta_{p,j}$ denote jth angle parameters for p-layer QAOA circuit for $j\leq p$. Then, optimization is performed by starting with each initial point and the angles giving the minimum cost are chosen as the optimum angles for $p=1$. For $p=2$ and $p=3$, the initial points are chosen such that $\widetilde{\gamma}_{p,r}=\widetilde{\gamma}_{p-1,r}$ and $\beta_{p,r}=\beta_{p-1,r}$ for $r\in[1,p-1]$ while $(\widetilde{\gamma}_{p,p},\beta_{p,p})$ region covering an area of $(2\pi\times2\pi)$ is discretized. It is observed that optimized $\widetilde{\gamma}_{p,p}$ and $\beta_{p,p}$ are in the neighborhood of $\widetilde{\gamma}_{p-1,p-1}$ and $\beta_{p-1,p-1}$, respectively, for $p\in[1,3]$. This observation is utilized while determining the grid sizes and boundaries for the initial point regions for $p\in[4,6]$ while with decreased number of points due to complexity as shown in Table II. For $p=7$, initial point is chosen with a heuristic extrapolation method described next and the same nonlinear optimization method is applied only for a single point. Extrapolation methods are also applied for Max-Cut problems in QAOA for high depths [5], [29]. Optimized angles for $p\in[1,5]$ and $SNR=15$ are presented in Table III. The values of $\widetilde{\boldsymbol{\gamma}}$ for $SNR=15$ are shown in Fig. 1(a) by including the extrapolated curves for $p\in[8,14]$ as discussed in the next section. All optimized parameters are available in data repository [21].

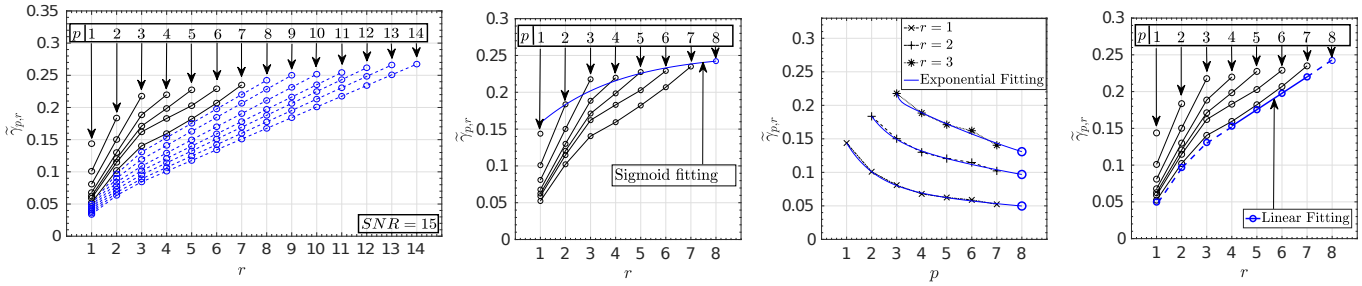
B. Extrapolation of $\boldsymbol{\gamma}$ and $\boldsymbol{\beta}$ for higher depths

We design a heuristic method to calculate angles for $p\in[8,14]$ by observing the pattern of the optimized angles for $p\in[1,7]$. The angles of $\widetilde{\gamma}_{p,r}$ for $p\in[1,7]$ and $r\in[1,p]$ are presented in Fig. 1(b) for SNR $=15$. It is observed that $\widetilde{\gamma}_{p,r}$ values increase as r increases while the curves shift to the right as p increases. Similar pattern is observed also for β values. Then, an initial estimation is performed manually for $p=7$ and an optimization is performed with this initial point to calculate the angles for $p=7$. For the values $p\in[8,14]$, a numerical algorithm is designed which is composed of three basic steps as shown in Algorithm 2. Firstly, we realize extrapolation for $\boldsymbol{\gamma}_{p,p}$ by using the following interpolation function based on the observation of a saturation behavior for $\widetilde{\gamma}_{r,r}$ for $r\in[1,p-1]$:

$$f_{\sigma,\vec{b}}(r)\equiv b_1+\frac{b_2}{1+\exp\left(-(r-b_3)\,/\,b_4\right)} \tag{48}$$

TABLE III
OPTIMIZED ANGLE PARAMETERS OF QAOA FOR SNR $= 15$ AND $p \in [1,5]$

	$p=1$	$p=2$	$p=3$	$p=4$	$p=5$
$\widetilde{\boldsymbol{\gamma}} = n\,\boldsymbol{\gamma}$	0.1438	0.1009, 0.1836	0.0809, 0.1502, 0.2177	0.0678, 0.1300, 0.1885, 0.2198	0.0625, 0.1206, 0.1711, 0.1985, 0.2275
$\boldsymbol{\beta}$	2.5422	2.3830, 2.7575	2.3439, 2.6162, 2.8963	2.3426, 2.5491, 2.7937, 2.9631	2.3078, 2.5215, 2.7408, 2.8822, 3.0037



(a) (b) (c) (d)

Fig. 1. Optimized and extrapolated (a) $\widetilde{\boldsymbol{\gamma}}$ for SNR $= 15$ and $p \in [1,14]$. The values for $p \in [8,14]$ are extrapolated from $p \in [1,7]$. Heuristic extrapolation steps to calculate angles for $p = 8$ with (b) sigmoid, (c) exponential and (d) linear fitting.

Algorithm 2 Extrapolation algorithm for $\widetilde{\gamma}_{p,r}$ and $\beta_{p,r}$ for $p \in [8,14]$ and $r \in [1,p]$

1: Choose saturation fitting curve defined in (48) and optimize coefficients b_j for $j \in [1,4]$ of $f_{\sigma,\vec{b}}(r)$ by using NLS and set $\chi_{p,p}$ equal to $f_{\sigma,\vec{b}}(p)$ where χ is either $\widetilde{\gamma}$ and β.
2: Choose exponential curve fitting model $f_{e,\vec{d}_r}(p')$ in (49) for $r \in [1,3]$ and optimize its coefficients $d_{r,j}$ for $j \in [1,4]$ with NLS and set $\chi_{p,r} = f_{e,\vec{d}_r}(p)$ for $r \in [1,3]$.
3: Calculate $\chi_{p,r}$ for $r \in [4, p-1]$ with linear fitting between $\chi_{p,3}$ and $\chi_{p,p}$.

where $r \in [1,p]$ and the optimized value of $\vec{b} \equiv [b_1\ b_2\ b_3\ b_4]$ is calculated by using nonlinear curve-fitting in least-squares (LS) sense by minimizing $\sum_{r=1}^{p-1} \left(f_{\sigma,\vec{b}}(r) - \widetilde{\gamma}_{r,r}\right)^2$ with respect to $\vec{b}$. Then, $\widetilde{\gamma}_{p,p}$ is set equal to $f_{\sigma,\vec{b}}(p)$. The values of $\widetilde{\gamma}_{p,r}$ for $p \in [1,7]$ and $r \in [1,p]$, interpolation curve and the value of $\widetilde{\gamma}_{8,8}$ are shown in Fig. 1(b). Secondly, we calculate the first three elements of $\widetilde{\gamma}_{8,r}$ for $r \in [1,3]$ by using an exponential fitting based on the observation that the values $\widetilde{\gamma}_{p',r}$ are decreasing exponentially while the values of p' increasing from r to $p-1$ for $r \in [1,3]$ as shown in Fig. 1(c). Exponential fitting curve is defined as follows:

$$f_{e,\vec{d}_r}(p') \equiv d_{r,1}\,\exp\left(d_{r,2}\,p'\right) + d_{r,3}\,\exp\left(d_{r,4}\,p'\right) \qquad (49)$$

where $\vec{d}_r \equiv [d_{r,1}\ d_{r,2}\ d_{r,3}\ d_{r,4}]$ is calculated with nonlinear LS (NLS) algorithm by minimizing $\sum_{p'=r}^{p-1} \left(f_{e,\vec{d}_r}(p') - \widetilde{\gamma}_{p',r}\right)^2$ with respect to $\vec{d}_r$. Then, $\widetilde{\gamma}_{p,r}$ for $r \in [1,3]$ is calculated as $\widetilde{\gamma}_{p,r} = f_{e,\vec{d}_r}(p)$. Finally, the values $\widetilde{\gamma}_{p,r}$ for $r \in [4,p-1]$ are calculated by realizing linear fitting between $\widetilde{\gamma}_{p,3}$ and $\widetilde{\gamma}_{p,p}$ as shown in Fig. 1(d). The same algorithm is utilized to calculate the extrapolation for $\boldsymbol{\beta}$ values. The extrapolated values of $\widetilde{\gamma}_{p,r}$ for $SNR = 15$ are shown in Fig. 1(a). It is observed that $\widetilde{\gamma}_{p,r}$ values increase almost linearly with respect to r as p increases.

C. QAOA performances for $n = 25$

P_e values obtained with QAOA, ML and MMSE solutions for varying SNR and depth $p \in [1,7]$ are shown in Fig. 2(a). It is observed that $p \geq 4$ reaches near-optimum ML performance for all SNR values while MMSE performance is significantly worse compared with QAOA and ML. In Fig. 2(b), the performance is shown for $SNR = 15$ and $p \in [1,14]$ showing the saturation at the optimum performance for $p \geq 5$ with small oscillations due to finite number of experiments M and potentially finite size effects for various p.

The cost performances are shown in Fig. 2(c) for varying SNR for QAOA, ML and MMSE where expected costs are normalized by multiplying with $\left(-2\,n^2\,\sigma_H^2\right) / E_{\boldsymbol{J},\boldsymbol{h}}\left[C(\boldsymbol{s})\right]$ to obtain smoother and accurate comparison between the results due to different number of simulation instances for varying SNR and p. In Fig. 2(c), it is observed that as p increases from 1 to 7, expectation $E_{\boldsymbol{J},\boldsymbol{h}}\left[\langle C_{QAOA}\rangle\right]$ curves for varying SNR get closer to the ML cost curve. Expected solution cost $E_{\boldsymbol{J},\boldsymbol{h}}\left[C(\boldsymbol{s})\right] = -2\,n^2\,\sigma_H^2 = -2\,n^2$ is also shown for reference. In Fig. 2(d), performance is shown for high depths with $p \in [1,14]$ and $SNR = 15$. It is observed that at high SNR, ML cost is almost equal to solution cost while $E_{\boldsymbol{J},\boldsymbol{h}}\left[\langle C_{QAOA}\rangle\right]$ is getting closer to ML cost compared with the performances for $p \in [1,7]$. Therefore, our heuristic extrapolation method surprisingly provides significant performance even though they do not include any optimization. It is observed in Fig. 2(e) that observed minimum values of the cost among QAOA circuit output measurement results are significantly close to ML cost. For example, at $p = 3$, $E_{\boldsymbol{J},\boldsymbol{h}}\left[\min_{\boldsymbol{z}}\{C_{QAOA}(\boldsymbol{z})\}\right]$ is close to ML cost $E_{\boldsymbol{J},\boldsymbol{h}}\left[C_{ML}\right]$.

Approximation ratios r, r^*, r_{ML} and r^*_{ML} are shown for varying $p \in [1,7]$ and SNR values of 1, 5 and 10 in Figs. 3(a), (b) and (c), respectively, while for varying $p \in [1,14]$ and $SNR = 15$ in Fig. 3(d). It is observed in Figs. 3(a), (b) and (c) that $r^*_{ML} \approx 1$ for $p \geq 5$ and $SNR \in \{1,5,10\}$, respectively. In Fig. 3(d), it is observed that both r^* and r^*_{ML} values ≈ 1

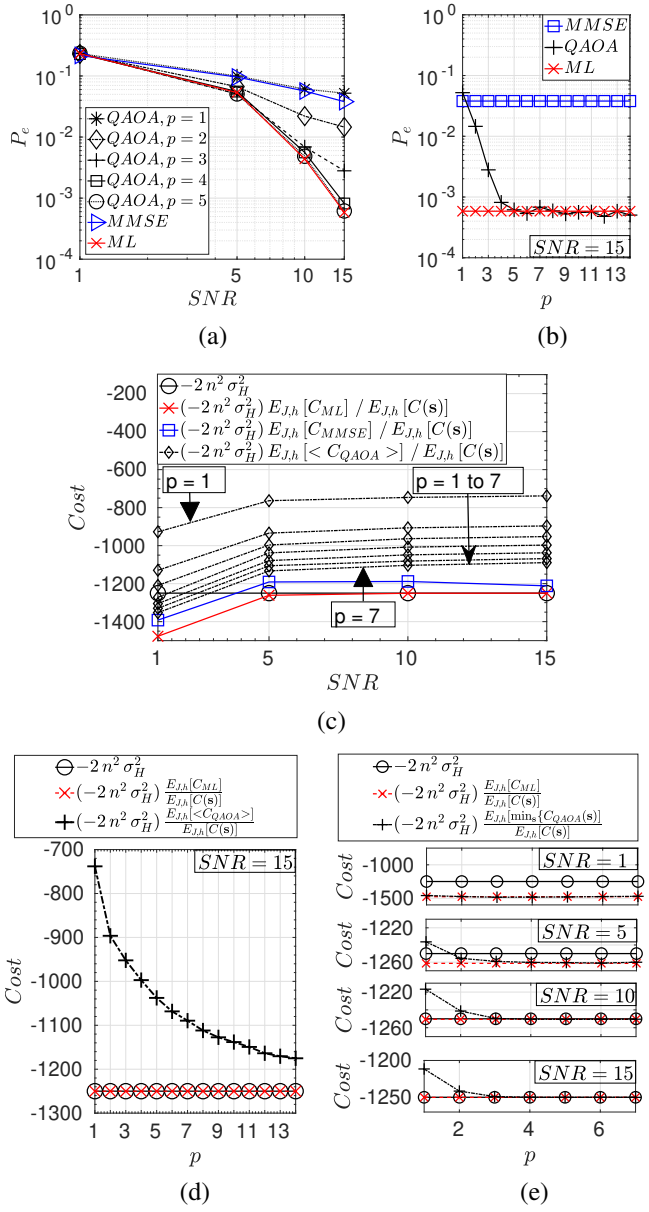


Fig. 2. P_e obtained with ML, MMSE and QAOA for varying (a) SNR $\in \{1, 5, 10, 15\}$ and layer depths $p \in [1, 5]$, and (b) SNR $= 15$ and $p \in [1, 14]$. $E_{\boldsymbol{J},\boldsymbol{h}}[C_{ML}]$ and $E_{\boldsymbol{J},\boldsymbol{h}}\left[\langle C_{QAOA}\rangle\right]$ normalized with multiplication by $(-2\,n^2\,\sigma_H^2)\,/\,E_{\boldsymbol{J},\boldsymbol{h}}[C(\boldsymbol{s})]$ for (c) varying SNR $\in \{1, 5, 10, 15\}$ and $p \in [1, 7]$ and (d) SNR $= 15$ and $p \in [1, 14]$ compared with $E_{\boldsymbol{J},\boldsymbol{h}}[C_{MMSE}]$ and $E_{\boldsymbol{J},\boldsymbol{h}}[C(\boldsymbol{s})] = -2\,n^2\,\sigma_H^2$. (e) Normalized $E_{\boldsymbol{J},\boldsymbol{h}}[C_{ML}]$ and $E_{\boldsymbol{J},\boldsymbol{h}}\left[\min_{\boldsymbol{z}}\{C_{QAOA}(\boldsymbol{z})\}\right]$ for varying SNR $\in \{1, 5, 10, 15\}$ and $p \in [1, 7]$ compared with $-2\,n^2\,\sigma_H^2$.

for $p \geq 5$ and SNR $= 15$. Both ML and QAOA solutions are more separated from the solution for low SNR value of 1 as shown in Fig. 3(a). r is not close to 1 for $SNR = 1$ for all p values while still $r^*_{ML} \approx 1$ with larger difference compared with $SNR \geq 5$. Similar observation is obtained for r_{ML} for $SNR = 1$ with an increase towards unity as p increases.

D. Statistical properties of QAOA cost metrics for varying n

We simulate statistical properties formulated in Theorem 1 and the corresponding results in (31-35) allowing us to present Conjecture 2. In Figs. 4(a) and (b), (31) and (32) are verified with simulations for $p \in [1, 4]$ and $n \in [10, 28]$ showing that $\mathbb{E}_{\boldsymbol{J},\boldsymbol{h}}\left[\langle C_{QAOA}\rangle\right]$ and $\mathbb{E}_{\boldsymbol{J},\boldsymbol{h}}\left[\langle C^2_{QAOA}\rangle\right]$ increase as $O(n^2)$ and $O(n^4)$, respectively, with respect to n. On the other hand, in Figs. 4(c) and (d), it is observed that variances $\mathbb{E}_{\boldsymbol{J},\boldsymbol{h}}\left[\langle C^2_{QAOA}\rangle\right] - \mathbb{E}^2_{\boldsymbol{J},\boldsymbol{h}}\left[\langle C_{QAOA}\rangle\right]$ and $\mathbb{E}_{\boldsymbol{J},\boldsymbol{h}}\left[\langle C^2_{QAOA}\rangle - \langle C_{QAOA}\rangle^2\right]$ increase as $O(n^3)$ with respect to n supporting concentration property defined in (33). It also shown that measurement result of each QAOA instance is also showing an increase of variance as $O(n^3)$ on average over many individual instances with respect to the symbol length n. In Fig. 4(e), variances are shown for $n = 25$ and varying $p \in [1, 14]$ observing that measurement variance $\mathbb{E}_{\boldsymbol{J},\boldsymbol{h}}\left[\langle C^2_{QAOA}\rangle - \langle C_{QAOA}\rangle^2\right]$ decreases with respect to p resembling a saturation while instance variance of the mean $\langle C_{QAOA}\rangle$ has a much smaller oscillation with respect to p while saturating for large p. On the other hand, in Fig. 4(f), it is observed that distance between the measured minimum cost and mean normalized by n^2, i.e., $E_{J,h}\left[< C_{QAOA} > - \min_{\boldsymbol{z}}\{C_{QAOA}(\boldsymbol{z})\}\right] / n^2$, decreases slowly for $p = 1$ and $p = 2$ as n increases supporting (34) in Conjecture 1. The same decrease is not observed for $p = 3$ and $p = 4$ probably due to their much slower decay function $g_p(n)$ compared with $g_2(n)$ and $g_1(n)$. In other words, it is conjectured that $E_{J,h}\left[< C_{QAOA} > - \min_{\boldsymbol{z}}\{C_{QAOA}(\boldsymbol{z})\}\right] / n^2$ for $p \geq 3$ starts to slowly decrease with increasing n if $n > 28$. Similarly, variance of $\min_{\boldsymbol{z}}\{C_{QAOA}(\boldsymbol{z})\}$ increases with n as $O(n^3)$ as shown in Fig. 4(g) supporting (35) in Conjecture 1.

E. Simulations for Conjecture-2

P_e for QAOA circuits optimized for $p \in [1, 4]$ is calculated for varying $n \in [10, 28]$ and compared with P_e of ML solution to estimate the maximum n value denoted by $\widehat{n}_{max}$ for QAOA circuits achieving ML performance. In Fig. 5(a), P_e of QAOA circuits for $p \in [1, 4]$ are compared with P_e of ML detection. It is observed that P_e with respect to n shows a local minimum approximately at the point $\widehat{n}_{max}$ where its performance starts to degrade compared with ML for increasing n. Approximation ratio r^* performance for varying n and $p \in [1, 4]$ is shown in Fig. 5(b). We obtain $\widehat{n}_{max} \in \{16, 18, 23, 28\}$ with the corresponding depth values of $p \in \{1, 2, 3, 4\}$, respectively, for $r^* = 0.9995$ and $\widehat{n}_{max} \in \{18, 22, 28\}$ with $p \in \{1, 2, 3\}$, respectively, for $r^* = 0.9967$. In Fig. 5(c), extrapolation based on (41) in Conjecture 2 is shown where $\widehat{n}_{max}$ for varying $\mathbb{E}_{\boldsymbol{J},\boldsymbol{h}}\left[\langle C_{QAOA}\rangle - C(\boldsymbol{s})\right] / n^2$ is shown. The first four points for $p \in [1, 4]$ are extracted from simulations shown in Fig. 5(b), $\left(\mathbb{E}_{\boldsymbol{J},\boldsymbol{h}}\left[\langle C^2_{QAOA}\rangle\right] - \mathbb{E}^2_{\boldsymbol{J},\boldsymbol{h}}\left[\langle C_{QAOA}\rangle\right]\right) / n^3$ in the numerator of (41) is calculated based on the observations in Fig. 4(e) and $\mathbb{E}_{\boldsymbol{J},\boldsymbol{h}}\left[\langle C_{QAOA}\rangle - C(\boldsymbol{s})\right] / n^2$ is calculated based on experiments for calculating the values in Fig. 2(d). Then, $\widehat{n}_{max}$ is extrapolated for target r^* values. We estimate that approximately $\widehat{n}_{max} < 200$ and $\widehat{n}_{max} < 300$ for $r^* = 0.9995$ and $r^* = 0.9967$ with $p \in [5, 14]$ and $p \in [4, 14]$, respectively. It is an open issue to verify the extrapolated performance with high number of qubits in NISQ devices and near future quantum computer simulators.

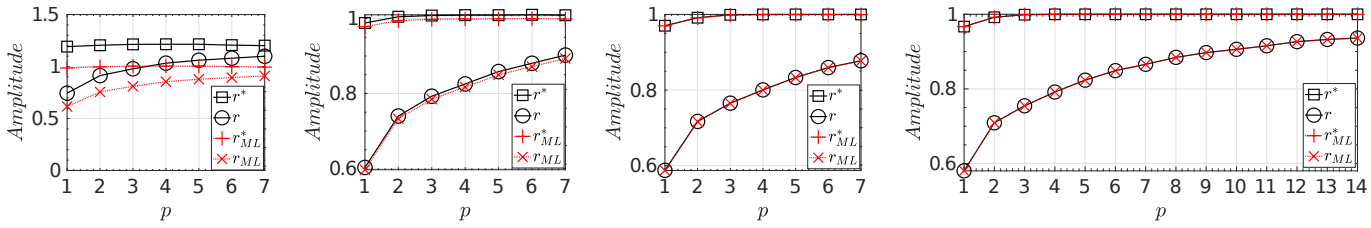


Fig. 3. Ratios r, r^*, r_{ML} and r^*_{ML} are shown for varying $p \in [1, 7]$ and SNR of (a) 1, (b) 5 and (c) 10 while (d) for varying $p \in [1, 14]$ and $SNR = 15$.

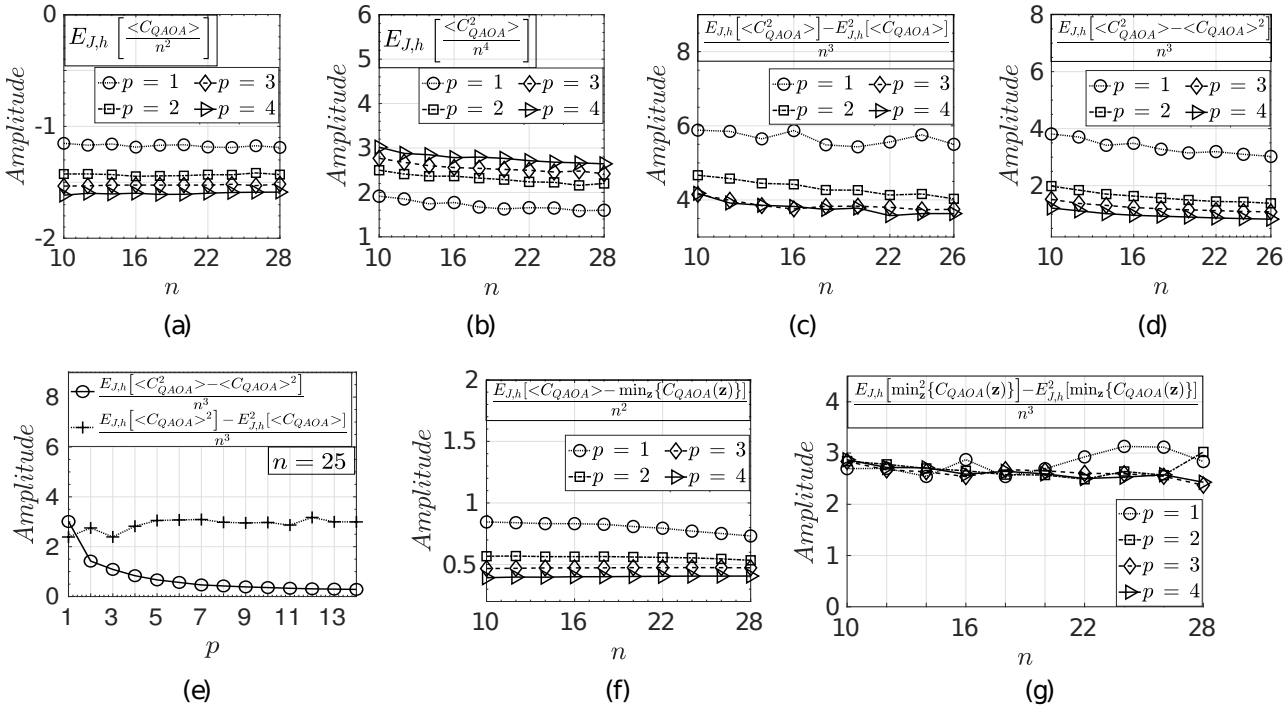


Fig. 4. Statistical properties of QAOA measurements with respect to problem instance based expectations for SNR $= 15$. (a) $\mathbb{E}_{\boldsymbol{J},\boldsymbol{h}}\left[\langle C_{QAOA}\rangle / n^2\right]$ and (b) $\mathbb{E}_{\boldsymbol{J},\boldsymbol{h}}\left[\langle C^2_{QAOA} / n^4\rangle\right]$ for varying $n \in [10, 28]$ and $p \in [1, 4]$. (c) $\left(\mathbb{E}_{\boldsymbol{J},\boldsymbol{h}}\left[\langle C^2_{QAOA}\rangle\right] - \mathbb{E}^2_{\boldsymbol{J},\boldsymbol{h}}\left[\langle C_{QAOA}\rangle\right]\right) / n^3$ and (d) $\left(\mathbb{E}_{\boldsymbol{J},\boldsymbol{h}}\left[\langle C^2_{QAOA}\rangle - \langle C_{QAOA}\rangle^2\right]\right) / n^3$ for varying $n \in [10, 26]$ and $p \in [1, 4]$. (e) $\left(\mathbb{E}_{\boldsymbol{J},\boldsymbol{h}}\left[\langle C^2_{QAOA}\rangle - \langle C_{QAOA}\rangle^2\right]\right) / n^3$ and $\left(\mathbb{E}_{\boldsymbol{J},\boldsymbol{h}}\left[\langle C_{QAOA}\rangle^2\right] - \mathbb{E}^2_{\boldsymbol{J},\boldsymbol{h}}\left[\langle C_{QAOA}\rangle\right]\right) / n^3$ for $n = 25$ and $p \in [1, 14]$. (f) The distance $E_{J,h}\left[< C_{QAOA} > - \min_{\boldsymbol{z}}\{C_{QAOA}(\boldsymbol{z})\}\right] / n^2$ and (g) variance $\left(\mathbb{E}_{\boldsymbol{J},\boldsymbol{h}}\left[\left(\min_{\boldsymbol{z}}\{C_{QAOA}(\boldsymbol{z})\}\right)^2\right] - \mathbb{E}^2_{\boldsymbol{J},\boldsymbol{h}}\left[\min_{\boldsymbol{z}}\{C_{QAOA}(\boldsymbol{z})\}\right]\right) / n^3$ for varying $n \in [10, 28]$ and $p \in [1, 4]$.

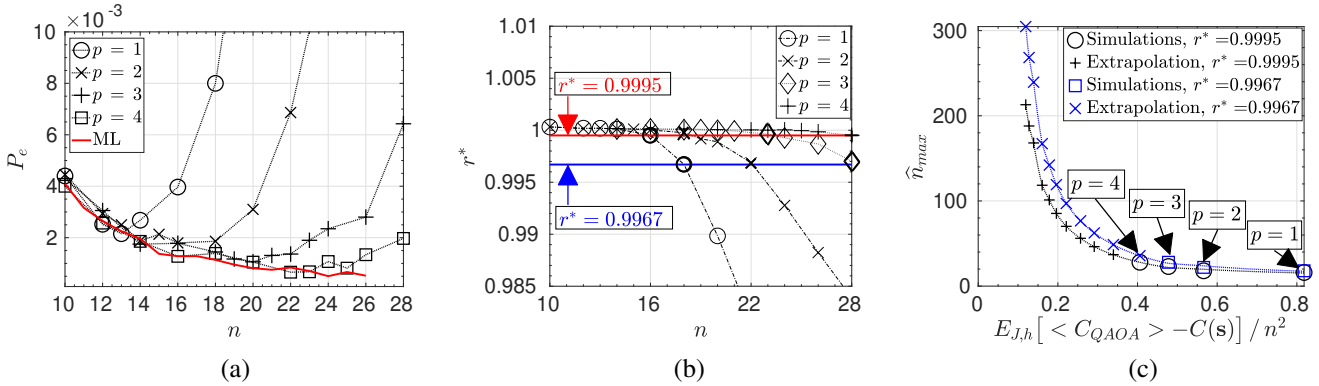


Fig. 5. (a) P_e for QAOA is compared with ML detection and (b) approximation ratio r^* for varying $n \in [10, 28]$ and $p \in [1, 4]$. (c) Extrapolation of maximum n denoted by $\widehat{n}_{max}$ based on (41) in Conjecture 2 for target $r^* = 0.9995$ and $r^* = 0.9967$ with $p \in [5, 14]$ and $p \in [4, 14]$, respectively.

VIII. Conclusions

We apply QAOA on ML detection of massive MIMO systems with an optimization algorithm of angle parameters having computational complexity of $O(16^p)$ for p-layer QAOA circuits based on SK modeling. Proposed algorithm optimizes angles independent from n and specific problem instance with significant practicality. We provide extensive simulation studies for QAOA by analyzing statistical properties of QAOA measurements in IBM Quantum Lab. We present a conjecture about concentration property of QAOA and share corresponding measurement statistics of $\langle C_{QAOA}\rangle$, $\langle C^2_{QAOA}\rangle$, $\min_{\boldsymbol{z}}\{C_{QAOA}(\boldsymbol{z})\}$, $C(\boldsymbol{s})$ and total number of bit errors for a total of 236500 individual QAOA circuits for varying n, p and SNR as a data repository in [21]. We observe that QAOA for $p \geq 4$ achieves near-optimum ML performance in terms of P_e for large scale 25×25 MIMO systems and SNR < 12 dB. In addition, we present extrapolated angles for $p \in [8, 14]$ for $SNR \approx 12$ dB with a conjecture claiming that presented angles achieve near-optimum performance for $n \leq 300$ with applicability for next generation massive MIMO systems.

Acknowledgments

This work was supported by TUBITAK (The Scientific and Technical Research Council of Turkey) under Grant #119E584.

Data Availability Statement

Problem instance specific values of $\langle C_{QAOA}\rangle$, $\langle C^2_{QAOA}\rangle$, $\min_{\boldsymbol{z}}\{C_{QAOA}(\boldsymbol{z})\}$, $C(\boldsymbol{s})$ and total number of bit errors obtained with 236500 randomly generated individual QAOA circuits in IBM Quantum Lab are available in the publicly accessible data repository in IEEE DataPort in [21]. Researchers can access the dataset and its associated documentation for further analysis and verification. Digital object identifier (DOI) of the repository is 10.21227/x0g7-n411.

Appendix A
Statistical Properties of Sherrington-Kirkpatrick Model of ML Problem

The law of large numbers allows to approximate the distributions of $J_{j,k}$ and h_j as Gaussian for $n \gg 1$. We also assume typical input symbols $\boldsymbol{s}$ of length n with s_i being equal to ± 1 with equal probability of $1/2$ leading to $\mu_{\boldsymbol{s}} = 0$. Then, the approximate values of mean and variance are calculated for typical instances of $\boldsymbol{H}$, $\boldsymbol{n}$, $\boldsymbol{J}$, $\boldsymbol{h}$ and $\boldsymbol{s}$ by using $\mu_{\boldsymbol{J}} \approx \frac{1}{n\,(n-1)\,/\,2}\sum_{j<k} J_{j,k}$, $\mu_{\boldsymbol{h}} \approx \frac{1}{n}\sum_j h_j$, $\sigma^2_{\boldsymbol{J}} \approx \frac{1}{n\,(n-1)\,/\,2}\sum_{j<k}(J_{j,k}-\mu_{\boldsymbol{J}})^2$ and $\sigma^2_{\boldsymbol{h}} \approx \frac{1}{n}\sum_j (h_j-\mu_{\boldsymbol{h}})^2$. Furthermore, we utilize $\frac{1}{n}\sum_\ell H_{\ell,k} = \frac{1}{n}\sum_k H_{\ell,k} \approx \mu_H = 0$, $\frac{1}{n}\sum_\ell H^2_{\ell,k} = \frac{1}{n}\sum_k H^2_{\ell,k} \approx \sigma^2_H$, $\frac{1}{n}\sum_\ell n_\ell \approx \mu_{\boldsymbol{n}} = 0$, $\frac{1}{n}\sum_\ell n^2_\ell \approx \sigma^2_{\boldsymbol{n}} = n\,\sigma^2_H\,/\,SNR$, $\frac{1}{n}\sum_k s_k \approx \mu_s = 0$. Then, by substituting (9) and (10), we obtain $\mu_{\boldsymbol{J}} \approx 2\,n\,\mu^2_H = 0$ and $\mu_{\boldsymbol{h}} \approx 0$. σ^2_J is approximated where the terms proportional to μ_H and $O(1\,/\,n^2)$ are excluded in the steps of the derivation by setting them equal to zero resulting in $\sigma^2_{\boldsymbol{J}} \approx n\,(4\,\sigma^4_H)$. In a similar manner, σ^2_h is approximated as $\sigma^2_{\boldsymbol{h}} \approx n^2\,(4\,\sigma^4_H)\,\left(2 - \frac{1}{n} + \frac{n-1}{n}\frac{1}{SNR}\right)$ where we use the fact that the expectation of $H^4(\ell, j)$ is equal to $3\,\sigma^4_H$ and we excluded the terms proportional to $O(1/n^2)$.

On the other hand, the mean and variance of the cost for the target solution are calculated by using (1), (9) and (10) as $\mu_{C_s} \equiv \mathbb{E}_{\boldsymbol{H},\boldsymbol{n}}[C(\boldsymbol{s})] = -2\,n^2\,\sigma^2_H$. The variance denoted by $\sigma^2_{C_s}$ is trivially calculated after a set of conversions as $\sigma^2_{C_s} \equiv \mathbb{E}_{\boldsymbol{H},\boldsymbol{n}}\left[C^2(\boldsymbol{s}) - \mu^2_{C_s}\right] = \sigma^4_H\,(2 + 6\,/\,n + 4\,/\,SNR)\,n^3 = O(n^3)$. This results in concentration property as follows:

$$\mathbb{P}_{\boldsymbol{H},\boldsymbol{n}}\left(\left|C(\boldsymbol{s})/n^2 + 2\,\sigma^2_H\right| > \epsilon\right) \to 0 \quad \text{as} \quad n \to \infty \tag{50}$$

Appendix B
Proof of Theorem 1

Following the approach in [5] as defined between equations (73)-(76), $\langle e^{\imath\,\frac{\lambda C}{n^2}}\rangle$ is calculated as follows by also including the effect of the local field of SK model:

$$\begin{aligned}\langle e^{\imath\,\frac{\lambda C}{n^2}}\rangle = \frac{1}{2^n}\sum_{\boldsymbol{z}^m}\sum_{\boldsymbol{z}^{[\pm 1]},\ldots,\boldsymbol{z}^{[\pm p]}}\left(\prod_{r=1}^{p} f_r(\boldsymbol{z}^{[r]})\,f^*_r(\boldsymbol{z}^{[-r]})\right)\\ \times\,\exp\left(\frac{\imath}{n}\sum_{j<k} J_{j,k}\,z^m_j\,z^m_k\,\left(\widetilde{\phi}_{j,k}+\frac{\lambda}{n}\right)\right)\\ \times\exp\left(\frac{\imath}{n}\sum_j h_j\,z^m_j\,\left(\widetilde{\phi}_j+\frac{\lambda}{n}\right)\right)\end{aligned} \tag{51}$$

where $z^{[r:p]}_j \equiv \prod_{l=r}^{p} z^{[l]}_j$ and $z^{[-r:-p]}_j \equiv \prod_{l=-p}^{-r} z^{[l]}_j$ are consecutive multiplications for the indices, $\boldsymbol{z}^{[\pm r]} = \left(z^{[\pm r]}_1, z^{[\pm r]}_2, \ldots, z^{[\pm r]}_n\right) \in \{+1,-1\}^n$ denote n-bit strings for $r \in [1,p]$, $\boldsymbol{z}^m = (z^m_1, z^m_2, \ldots, z^m_n) \in \{+1,-1\}^n$, $f_r(\boldsymbol{z}.\boldsymbol{z}') \equiv \langle \boldsymbol{z}\,|e^{\imath\,\beta_r\,B}|\,\boldsymbol{z}'\rangle = \langle \boldsymbol{z}.\boldsymbol{z}'|e^{\imath\,\beta_r\,B}|\boldsymbol{1}\rangle = (\cos\beta_r)^{g^+(\boldsymbol{z})}\,(\imath\,\sin\beta_r)^{g^-(\boldsymbol{z})}$, $\boldsymbol{z}.\boldsymbol{z}'$ is the bit-wise product of $\boldsymbol{z}$ and $\boldsymbol{z}'$, $g^+(\boldsymbol{z})$ and $g^-(\boldsymbol{z})$ count the number of ones and minus ones in $\boldsymbol{z}$, respectively, and $\widetilde{\phi}_{j,k}$ and $\widetilde{\phi}_j$ defined as follows:

$$\widetilde{\phi}_{j,k} = \sum_{r=1}^{p}\widetilde{\gamma}_{p,r}\left(z^{[r:p]}_j\,z^{[r:p]}_k - z^{[-r:-p]}_j\,z^{[-r:-p]}_k\right) \tag{52}$$

$$\widetilde{\phi}_j = \sum_{r=1}^{p}\widetilde{\gamma}_{p,r}\left(z^{[r:p]}_j - z^{[-r:-p]}_j\right) \tag{53}$$

where $\gamma_{p,r} \equiv \widetilde{\gamma}_{p,r}\,/\,n$ for $r \in [1,p]$ based on our assumption.

The characteristic function of a Gaussian random variable X with mean μ_X and variance σ^2_X or the expectation $\mathbb{E}_X[\exp(\imath\,t\,X)]$ equals to $\exp(\imath\,\mu_X\,t)\,\exp(-\,\sigma^2_X\,t^2\,/\,2)$. Furthermore, taking expectation with respect to $J_{j,k}\,z^m_j\,z^m_k$ and $h_j\,z^m_j$ are the same with taking the expectation with respect to $J_{j,k}$ and h_j, respectively, due to symmetry. After replacing $J_{j,k}$ with $J_{j,k}\,z^m_j\,z^m_k$ and h_j with $h_j\,z^m_j$, the expectation of $\langle e^{\imath\,\lambda\,C\,/\,n^2}\rangle$ with respect to $\boldsymbol{J}$ and $\boldsymbol{h}$ is calculated as follows:

$$\begin{aligned}\mathbb{E}_{\boldsymbol{J},\boldsymbol{h}}\left[\langle e^{\imath\,\lambda\,C\,/\,n^2}\rangle\right] = \sum_{\boldsymbol{z}^{[\pm 1]},\ldots,\boldsymbol{z}^{[\pm p]}}\prod_{r=1}^{p}\left(f_r(\boldsymbol{z}^{[r]})\,f^*_r(\boldsymbol{z}^{[-r]})\right)\\ \exp\left(-\frac{\widetilde{\sigma_J}^2}{2\,n}\sum_{j<k}\left(\widetilde{\phi}_{j,k}+\frac{\lambda}{n}\right)^2 - \frac{\widetilde{\sigma_h}^2}{2}\sum_j\left(\widetilde{\phi}_j+\frac{\lambda}{n}\right)^2\right)\end{aligned} \tag{54}$$

where $\widetilde{\sigma_J}^2 = \sigma_J^2 / n$, $\widetilde{\sigma_h}^2 = \sigma_h^2 / n^2$. In Section 6 of [5], a simplifying approach is proposed to calculate a similar form by summing in a configuration basis instead of all $(2^n)^{2p}$ possible strings $\left(\boldsymbol{z}^{[1]}, \ldots, \boldsymbol{z}^{[p]}, \boldsymbol{z}^{[-p]}, \ldots, \boldsymbol{z}^{[-1]}\right)$ in (54). Summation in (54) is converted into the following by expanding the square terms and then taking the first derivative of both sides with respect to λ at $\lambda = 0$ (similar to the formulation between the equations (86)-(95) and (117)-(118) in [5]):

$$\begin{aligned}\mathbb{E}_{\boldsymbol{J},\boldsymbol{h}}\left[\langle C / n^2\rangle\right] &= \sum_{\{n_a\}} \binom{n}{\{n_{\boldsymbol{a}}\}} \prod_{\boldsymbol{a}\in A} Q_a^{n_a} \\ &\left(\imath \frac{\widetilde{\sigma_J}^2}{2\,n^2} \sum_{\boldsymbol{u},\boldsymbol{v}\in A} \Phi_{\boldsymbol{u}.\boldsymbol{v}} n_{\boldsymbol{u}} n_{\boldsymbol{v}} + \imath \frac{\widetilde{\sigma_h}^2}{n} \sum_{\boldsymbol{u}\in A} \Phi_{\boldsymbol{u}} n_{\boldsymbol{u}}\right) \\ &\exp\left(-\frac{\widetilde{\sigma_J}^2}{4\,n} \sum_{\boldsymbol{a},\boldsymbol{b}\in A} \Phi_{\boldsymbol{a}.\boldsymbol{b}}^2 n_{\boldsymbol{a}} n_{\boldsymbol{b}} - \frac{\widetilde{\sigma_h}^2}{2} \sum_{\boldsymbol{a}\in A} \Phi_{\boldsymbol{a}}^2 n_{\boldsymbol{a}}\right) \quad (55)\end{aligned}$$

where $\boldsymbol{a}.\boldsymbol{b}$ denotes bitwise product of $\boldsymbol{a}$ and $\boldsymbol{b}$, multinomial coefficient $\binom{n}{\{n_{\boldsymbol{a}}\}}$ is equal to the value of $n!\,/\,(n_1!\,n_2!\,\ldots\,n_{2^{2p}}!)$ such that $\sum_{i=1}^{2^{2p}} n_i = n$, $\Phi_{\boldsymbol{a}} = \sum_{r=1}^{p} \widetilde{\gamma}_{p,r}\,(a_{r:p} - a_{-r:-p})$, $a_{r:p} = \prod_{l=r}^{p} a_l$, $a_{-r:-p} = \prod_{l=-p}^{-r} a_l$ for $r \in [1,p]$ and Q_a is defined as follows:

$$Q_a = \prod_{j=1}^{p} (\cos\beta_{p,j})^{1+\frac{a_j+a_{-j}}{2}}\,(\sin\beta_{p,j})^{1-\frac{a_j+a_{-j}}{2}}\,(i)^{\frac{a_{-j}-a_j}{2}} \quad (56)$$

Similarly, the following is obtained for the second moment:

$$\begin{aligned}\mathbb{E}_{\boldsymbol{J},\boldsymbol{h}}\left[\langle C^2 / n^4\rangle\right] &= \frac{\widetilde{\sigma_J}^2(n-1)}{2\,n^2} + \frac{\widetilde{\sigma_h}^2}{n} \\ &- \sum_{\{n_a\}} \binom{n}{\{n_{\boldsymbol{a}}\}} \prod_{\boldsymbol{a}\in A} Q_a^{n_a} \\ &\left[\frac{\widetilde{\sigma_J}^4}{4\,n^4}\left(\sum_{\boldsymbol{u},\boldsymbol{v}\in A} \Phi_{\boldsymbol{u}.\boldsymbol{v}} n_{\boldsymbol{u}} n_{\boldsymbol{v}}\right)^2 + \frac{\widetilde{\sigma_h}^4}{n^2}\left(\sum_{\boldsymbol{u}\in A} \Phi_{\boldsymbol{u}} n_{\boldsymbol{u}}\right)^2\right. \\ &\left. + \frac{\widetilde{\sigma_J}^2\widetilde{\sigma_h}^2}{n^3}\left(\sum_{\boldsymbol{u},\boldsymbol{v}\in A} \Phi_{\boldsymbol{u}.\boldsymbol{v}} n_{\boldsymbol{u}} n_{\boldsymbol{v}}\right)\left(\sum_{\boldsymbol{u}\in A} \Phi_{\boldsymbol{u}} n_{\boldsymbol{u}}\right)\right] \\ &\exp\left(-\frac{\widetilde{\sigma_J}^2}{4\,n} \sum_{\boldsymbol{a},\boldsymbol{b}\in A} \Phi_{\boldsymbol{a}.\boldsymbol{b}}^2 n_{\boldsymbol{a}} n_{\boldsymbol{b}} - \frac{\widetilde{\sigma_h}^2}{2} \sum_{\boldsymbol{a}\in A} \Phi_{\boldsymbol{a}}^2 n_{\boldsymbol{a}}\right) \quad (57)\end{aligned}$$

Assume that the set $\boldsymbol{R}$ denotes one of sets $\{\boldsymbol{u},\boldsymbol{v},\boldsymbol{x},\boldsymbol{y}\}$, $\{\boldsymbol{u},\boldsymbol{v},\boldsymbol{x}\}$, $\{\boldsymbol{u},\boldsymbol{v}\}$ or $\{\boldsymbol{u}\}$ for given strings $\boldsymbol{u},\boldsymbol{v},\boldsymbol{x}$ and $\boldsymbol{y}$. Following the similar methodology in [5], the following functions corresponding to $S_{\boldsymbol{u},\boldsymbol{v},\boldsymbol{x},\boldsymbol{y}}$, $S_{\boldsymbol{u},\boldsymbol{v},\boldsymbol{x}}$, $S_{\boldsymbol{u},\boldsymbol{v}}$ and $S_{\boldsymbol{u}}$ are defined:

$$\begin{aligned}S_{\boldsymbol{R}} &= \sum_{\{n_a\}} \binom{n}{\{n_{\boldsymbol{a}}\}} \prod_{\boldsymbol{a},\,\boldsymbol{b}\,\in A} g_{\boldsymbol{a}.\boldsymbol{b}}^{n_{\boldsymbol{a}}\,n_{\boldsymbol{b}}\,/\,2} \prod_{\boldsymbol{a}\,\in A} g_{\boldsymbol{a}}^{n_{\boldsymbol{a}}\,/\,2} \\ &\prod_{\boldsymbol{a}\in A} Q_a^{n_a}\,\widetilde{f}(R,n) \quad (58)\end{aligned}$$

where $g_{\boldsymbol{a}.\boldsymbol{b}}$ and $g_{\boldsymbol{a}}$ are defined as follows:

$$\begin{aligned}g_{\boldsymbol{a}.\boldsymbol{b}} &\equiv \exp\left(\left(-\widetilde{\sigma_J}^2\,/\,(2\,n)\right)\Phi_{\boldsymbol{a}.\boldsymbol{b}}^2\right) \quad (59)\\ g_{\boldsymbol{a}} &\equiv \exp\left(-\widetilde{\sigma_h}^2\,\Phi_{\boldsymbol{a}}^2\right) \quad (60)\end{aligned}$$

and the function $\widetilde{f}(R,n)$ is defined as follows:

$$\widetilde{f}(R,n) = \begin{cases}\binom{n_{\boldsymbol{u}}}{n}\binom{n_{\boldsymbol{v}}}{n}\binom{n_{\boldsymbol{x}}}{n}\binom{n_{\boldsymbol{y}}}{n}, & \text{if } R = \{\boldsymbol{u},\boldsymbol{v},\boldsymbol{x},\boldsymbol{y}\}\\ \binom{n_{\boldsymbol{u}}}{n}\binom{n_{\boldsymbol{v}}}{n}\binom{n_{\boldsymbol{x}}}{n}, & \text{if } R = \{\boldsymbol{u},\boldsymbol{v},\boldsymbol{x}\}\\ \binom{n_{\boldsymbol{u}}}{n}\binom{n_{\boldsymbol{v}}}{n}, & \text{if } R = \{\boldsymbol{u},\boldsymbol{v}\}\\ \binom{n_{\boldsymbol{u}}}{n}, & \text{if } R = \{\boldsymbol{u}\}\end{cases} \quad (61)$$

Then, (55) and (57) are transformed into the following:

$$\begin{aligned}\mathbb{E}_{\boldsymbol{J},\boldsymbol{h}}\left[\left\langle\frac{C}{n^2}\right\rangle\right] &= \frac{\imath\,\widetilde{\sigma_J}^2}{2} \sum_{\boldsymbol{u},\boldsymbol{v}\in A} \Phi_{\boldsymbol{u}.\boldsymbol{v}} S_{\boldsymbol{u},\boldsymbol{v}} \\ &\quad + \imath\,\widetilde{\sigma_h}^2 \sum_{\boldsymbol{u}\in A} \Phi_{\boldsymbol{u}} S_{\boldsymbol{u}} \quad (62)\\ \mathbb{E}_{\boldsymbol{J},\boldsymbol{h}}\left[\left\langle\frac{C^2}{n^4}\right\rangle\right] &= \frac{\widetilde{\sigma_J}^2(n-1)}{2\,n^2} + \frac{\widetilde{\sigma_h}^2}{n} \\ &\quad - \frac{\widetilde{\sigma_J}^4}{4} \sum_{\boldsymbol{u},\boldsymbol{v},\boldsymbol{x},\boldsymbol{y}\in A} \Phi_{\boldsymbol{u}.\boldsymbol{v}}\,\Phi_{\boldsymbol{x}.\boldsymbol{y}}\,S_{\boldsymbol{u},\boldsymbol{v},\boldsymbol{x},\boldsymbol{y}} \\ &\quad - \widetilde{\sigma_h}^4 \sum_{\boldsymbol{u},\boldsymbol{v}\in A} \Phi_{\boldsymbol{u}}\,\Phi_{\boldsymbol{v}}\,S_{\boldsymbol{u},\boldsymbol{v}} \\ &\quad - \widetilde{\sigma_J}^2\widetilde{\sigma_h}^2 \sum_{\boldsymbol{u},\boldsymbol{v},\boldsymbol{x}\in A} \Phi_{\boldsymbol{u}.\boldsymbol{v}}\Phi_{\boldsymbol{x}}\,S_{\boldsymbol{u},\boldsymbol{v},\boldsymbol{x}} \quad (63)\end{aligned}$$

Therefore, the problem is converted to the calculations of $S_{\boldsymbol{u},\boldsymbol{v},\boldsymbol{x},\boldsymbol{y}}$, $S_{\boldsymbol{u},\boldsymbol{v},\boldsymbol{x}}$, $S_{\boldsymbol{u},\boldsymbol{v}}$ and $S_{\boldsymbol{u}}$ for $\boldsymbol{u},\boldsymbol{v},\boldsymbol{x},\boldsymbol{y} \in A$. In a similar manner to the proof of Lemma-2 in Section 6.2 in [5], these functions can be expressed by separating the summation with respect to $\{n_a\}$ into two subsets of the set A. The partition of A as $A = B \cup A_{p+1} = D \cup D^c \cup A_{p+1}$ is described in Section V. Then, $S_{\boldsymbol{R}}$ is expressed as follows:

$$\begin{aligned}&= \sum_{t=0}^{n} \binom{n}{t} \sum_{\{n_{\boldsymbol{a}}:\boldsymbol{a}\in A_{p+1}\}} \sum_{\{n_{\boldsymbol{b}}:\boldsymbol{b}\in B\}} \binom{t}{\{n_{\boldsymbol{b}}\}} \binom{n-t}{\{n_{\boldsymbol{a}}\}} \\ &\prod_{\boldsymbol{b},\boldsymbol{b}'\in B} g_{\boldsymbol{b}.\boldsymbol{b}'}^{n_b n_{b'}/2} \prod_{\boldsymbol{a},\boldsymbol{a}'\in A_{p+1}} g_{\boldsymbol{a}.\boldsymbol{a}'}^{n_a n_{a'}/2} \prod_{\boldsymbol{a}\in A_{p+1},\,\boldsymbol{b}\in B} g_{\boldsymbol{a}.\boldsymbol{b}}^{n_a n_b/2} \\ &\prod_{\boldsymbol{a}\in B,\,\boldsymbol{b}\in A_{p+1}} g_{\boldsymbol{a}.\boldsymbol{b}}^{n_a n_b/2} \prod_{\boldsymbol{a}\in A_{p+1}} g_{\boldsymbol{a}}^{n_a\,/\,2} \prod_{\boldsymbol{b}\in B} g_{\boldsymbol{b}}^{n_b\,/\,2} \\ &\prod_{\boldsymbol{a}\in A_{p+1}} Q_{\boldsymbol{a}}^{n_a} \prod_{\boldsymbol{b}\in B} Q_{\boldsymbol{b}}^{n_b}\,\widetilde{f}(R,n) \quad (64)\end{aligned}$$

The following observations as described in Sections 6.1 and 6.2 in [5] are used to simplify the equation. It is observed that $g_{\boldsymbol{a}.\boldsymbol{b}} = g_{\boldsymbol{b}.\boldsymbol{a}}$, $g_{\boldsymbol{a}.\boldsymbol{a}'} = 1$ since $\Phi_{\boldsymbol{a}.\boldsymbol{a}'} = 0$ for $\boldsymbol{a},\,\boldsymbol{a}' \in A_{p+1}$ and $g_{\boldsymbol{a}} = 1$ since $\Phi_{\boldsymbol{a}} = \Phi_{\boldsymbol{a}.\boldsymbol{1}} = 0$ where both all ones string $\boldsymbol{1}$ and $\boldsymbol{a} \in A_{p+1}$. Then, (64) is converted into $S_{\boldsymbol{R}} = \sum_{t=0}^{n} s_{\boldsymbol{R}}(t,n)$ where $s_{\boldsymbol{R}}(t,n)$ is simplified as follows:

$$\begin{aligned}s_{\boldsymbol{R}}(t,n) &= \binom{n}{t} \sum_{\{n_{\boldsymbol{b}}:\boldsymbol{b}\in B\}} \binom{t}{\{n_{\boldsymbol{b}}\}} \prod_{\boldsymbol{b},\boldsymbol{b}'\in B} g_{\boldsymbol{b}.\boldsymbol{b}'}^{n_b n_{b'}/2} \\ &\prod_{\boldsymbol{b}\in B} \widetilde{Q}_{\boldsymbol{b}}^{n_b} \sum_{\{n_{\boldsymbol{a}}:\boldsymbol{a}\in A_{p+1}\}} \binom{n-t}{\{n_{\boldsymbol{a}}\}} \\ &\prod_{\boldsymbol{a}\in A_{p+1},\,\boldsymbol{b}\in B} g_{\boldsymbol{a}.\boldsymbol{b}}^{n_a n_b} \prod_{\boldsymbol{a}\in A_{p+1}} \widetilde{Q}_{\boldsymbol{a}}^{n_a}\,\widetilde{f}(R,n) \quad (65)\end{aligned}$$

where $\widetilde{Q}_b = Q_b\,g_{\boldsymbol{b}}^{1/2}$. The only difference between (65) and the equation (132) in Proof of Lemma-2 in Section 6.2 in [5] is

that Q_b is replaced with $\widetilde{Q}_b = Q_b\, g_{\boldsymbol{b}}^{1\,/\,2}$ where $g_{\boldsymbol{b}}^{1\,/\,2}$ does not depend on n in a similar manner with Q_b. Therefore, all the following proofs and calculations in [5] are directly applicable for the current model in this article without requiring the repetition of the proof in [5]. It is shown that $s_{\boldsymbol{R}}(t,n)$ is expressed as $s_{\boldsymbol{R}}(t,n) = \ell_{\boldsymbol{R}}(t)\,(1+O(1\,/\,n))$ with some calculated $\ell_{\boldsymbol{R}}(t)$ not depending on n. As a result, based on the proofs described in Lemma-2 in [5], the following equalities are obtained:

$$\lim_{n\to\infty} S_{\boldsymbol{R}} == \begin{cases} \widetilde{W}_{\boldsymbol{u}}\,\widetilde{W}_{\boldsymbol{v}}\,\widetilde{W}_{\boldsymbol{x}}\,\widetilde{W}_{\boldsymbol{y}}, & \text{if } R = \{\boldsymbol{u},\boldsymbol{v},\boldsymbol{x},\boldsymbol{y}\} \\ \widetilde{W}_{\boldsymbol{u}}\,\widetilde{W}_{\boldsymbol{v}}\,\widetilde{W}_{\boldsymbol{x}}\,, & \text{if } R = \{\boldsymbol{u},\boldsymbol{v},\boldsymbol{x}\} \\ \widetilde{W}_{\boldsymbol{u}}\,\widetilde{W}_{\boldsymbol{v}}, & \text{if } R = \{\boldsymbol{u},\boldsymbol{v}\} \\ \widetilde{W}_{\boldsymbol{u}}, & \text{if } R = \{\boldsymbol{u}\} \end{cases} \tag{66}$$

where the algorithm to obtain $\widetilde{W}_{\boldsymbol{u}}$ for $\boldsymbol{u}\in A$ is presented in Algorithm 1 based on the iterative formulation in page 26 in Section 6.2 in [5] with the main difference of replacing Q_b with $\widetilde{Q}_b = Q_b\, g_{\boldsymbol{b}}^{1\,/\,2}$. Besides that, after the observation in the equations (119) and (120) in [5] and inserting (66) into (62), the following is obtained for $\lim_{n\to\infty}\mathbb{E}_{\boldsymbol{J},\boldsymbol{h}}\left[\langle C\,/\,n^2\rangle\right]$:

$$= \quad \frac{\imath\,\widetilde{\sigma_J}^2}{2}\sum_{r=1}^{p}\widetilde{\gamma}_{p,r}\Gamma_r^+\,\Gamma_r^- \;+\; \imath\,\widetilde{\sigma_h}^2\sum_{r=1}^{p}\widetilde{\gamma}_{p,r}\,\Gamma_r^- \tag{67}$$

while the following is obtained for $\lim_{n\to\infty}\mathbb{E}_{\boldsymbol{J},\boldsymbol{h}}\left[\langle C^2\,/\,n^4\rangle\right]$:

$$\begin{aligned} = \quad & \lim_{n\to\infty}\frac{\widetilde{\sigma_J}^2(n-1)}{2\,n^2} + \lim_{n\to\infty}\frac{\widetilde{\sigma_h}^2}{n} \\ & - \left(\frac{\widetilde{\sigma_J}^2}{2}\sum_{r=1}^{p}\widetilde{\gamma}_{p,r}\,\Gamma_r^+\,\Gamma_r^-\right)^2 - \left(\widetilde{\sigma_h}^2\sum_{r=1}^{p}\widetilde{\gamma}_{p,r}\Gamma_r^-\right)^2 \\ & - \left(\widetilde{\sigma_J}^2\sum_{r=1}^{p}\widetilde{\gamma}_{p,r}\Gamma_r^+\,\Gamma_r^-\right)\left(\widetilde{\sigma_h}^2\sum_{r=1}^{p}\widetilde{\gamma}_{p,r}\Gamma_r^-\right) \end{aligned} \tag{68}$$

It is easily shown that the concentration property holds by excluding the first two terms decreasing with $O(1\,/\,n)$ and observing the remaining terms being equal to (68) as follows:

$$\left(\lim_{n\to\infty}\mathbb{E}_{\boldsymbol{J},\boldsymbol{h}}\left[\langle C\,/\,n^2\rangle\right]\right)^2 \;=\; \lim_{n\to\infty}\mathbb{E}_{\boldsymbol{J},\boldsymbol{h}}\left[\langle C^2\,/\,n^4\rangle\right] \tag{69}$$

Besides that, observe that both $\mathbb{E}_{\boldsymbol{J},\boldsymbol{h}}\left[\langle C\,/\,n^2\,\rangle\right]$ and $\mathbb{E}_{\boldsymbol{J},\boldsymbol{h}}\left[\langle C^2\,/\,n^4\rangle\right]$ include additive terms decreasing with $O(1\,/\,n)$. This completes the proof.

REFERENCES

[1] J. Preskill, "Quantum computing in the NISQ era and beyond," *Quantum*, vol. 2, p. 79, 2018.

[2] E. Farhi, J. Goldstone, and S. Gutmann, "A quantum approximate optimization algorithm," *arXiv:1411.4028*, 2014.

[3] E. Farhi and A. W. Harrow, "Quantum supremacy through the quantum approximate optimization algorithm," *arXiv:1602.07674*, 2016.

[4] A. Ozaeta, W. van Dam, and P. L. McMahon, "Expectation values from the single-layer quantum approximate optimization algorithm on Ising problems," *Quantum Science and Technology*, vol. 7, no. 4, p. 045036, 2022.

[5] E. Farhi, J. Goldstone, S. Gutmann, and L. Zhou, "The quantum approximate optimization algorithm and the Sherrington-Kirkpatrick model at infinite size," *Quantum*, vol. 6, p. 759, 2022.

[6] A. Montanari, "Optimization of the Sherrington–Kirkpatrick Hamiltonian," *SIAM Journal on Computing*, no. 0, pp. FOCS19–1, 2021.

[7] 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," in *IEEE 63rd Annual Symposium on Foundations of Computer Science (FOCS)*, pp. 335–343, 2022.

[8] J. Cui, Y. Xiong, S. X. Ng, and L. Hanzo, "Quantum approximate optimization algorithm based maximum likelihood detection," *IEEE Transactions on Communications*, vol. 70, no. 8, pp. 5386–5400, 2022.

[9] A. K. Singh, K. Jamieson, P. L. McMahon, and D. Venturelli, "Ising machines' dynamics and regularization for near-optimal MIMO detection," *IEEE Transactions on Wireless Communications*, vol. 21, no. 12, pp. 11 080–11 094, 2022.

[10] J. Cui, G. L. Long, and L. Hanzo, "General Hamiltonian representation of ML detection relying on the quantum approximate optimization algorithm," *arXiv:2204.05126*, 2022.

[11] T. Matsumine, T. Koike-Akino, and Y. Wang, "Channel decoding with quantum approximate optimization algorithm," in *IEEE Int. Symp. on Inf. Theory (ISIT)*, pp. 2574–2578, 2019.

[12] B. Liu, T. Koike-Akino, Y. Wang, and K. Parsons, "Learning to learn quantum turbo detection," *arXiv:2205.08611*, 2022.

[13] A. K. Singh, A. Kapelyan, D. Venturelli, and K. Jamieson, "Uplink MIMO detection using Ising machines: A multi-stage Ising approach," *arXiv:2304.12830*, 2023.

[14] S. Kasi, A. K. Singh, D. Venturelli, and K. Jamieson, "Quantum annealing for large mimo downlink vector perturbation precoding," in *IEEE Int. Conf. on Comm. (ICC)*, pp. 1–6, 2021.

[15] Z. I. Tabi, Á. Marositz, Z. Kallus, P. Vaderna, I. Gódor, and Z. Zimborás, "Evaluation of quantum annealer performance via the massive MIMO problem," *IEEE Access*, vol. 9, pp. 131 658–131 671, 2021.

[16] M. Kim and K. Jamieson, "Finer-grained decomposition for parallel quantum MIMO processing," in *IEEE Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 1–5, 2023.

[17] M. Kim, D. Venturelli, and K. Jamieson, "Leveraging quantum annealing for large MIMO processing in centralized radio access networks," in *Proc. of the ACM Special Interest Group on Data Comm.*, pp. 241–255, 2019.

[18] M. Norimoto, R. Mori, and N. Ishikawa, "Quantum algorithm for higher-order unconstrained binary optimization and MIMO maximum likelihood detection," *IEEE Transactions on Communications*, vol. 71, no. 4, pp. 1926–1939, 2023.

[19] L. Zhou et al., "Quantum approximate optimization algorithm: Performance, mechanism, and implementation on near-term devices,"*Physical Review X*, vol. 10, no. 2, p. 021067, 2020.

[20] "IBM Quantum Lab," [Online], 2023, available: https://quantum-computing.ibm.com.

[21] B. Gulbahar, September 20, 2023, "Maximum-likelihood detection with QAOA for massive MIMO and Sherrington-Kirkpatrick model with local field at infinite size", *IEEE Dataport*, doi: https://dx.doi.org/10.21227/x0g7-n411.

[22] G. S. Barron and C. J. Wood, "Measurement error mitigation for variational quantum algorithms," *arXiv:2010.08520*, 2020.

[23] A. Weidinger, G. B. Mbeng, and W. Lechner, "Error mitigation for quantum approximate optimization," *arXiv:2301.05042*, 2023.

[24] L. Cheng, Y.-Q. Chen, S.-X. Zhang, and S. Zhang, "Error-mitigated quantum approximate optimization via learning-based adaptive optimization," *arXiv:2303.14877*, 2023.

[25] S. Hadfield et al., "From the quantum approximate optimization algorithm to a quantum alternating operator ansatz," *Algorithms*, vol. 12, no. 2, p. 34, 2019.

[26] P. Díez-Valle, D. Porras, and J. J. García-Ripoll, "Quantum approximate optimization algorithm pseudo-Boltzmann states," *Physical Review Letters*, vol. 130, no. 5, p. 050601, 2023.

[27] P. C. Lotshaw, G. Siopsis, J. Ostrowski, R. Herrman, R. Alam, S. Powers, and T. S. Humble, "Approximate Boltzmann distributions in quantum approximate optimization," *arXiv:2212.01857*, 2022.

[28] E. Biglieri, R. Calderbank, A. Constantinides, A. Goldsmith, A. Paulraj, and H. V. Poor, *MIMO Wireless Communications*. Cambridge University Press, 2007.

[29] J. Basso, E. Farhi, K. Marwaha, B. Villalonga, and L. Zhou, "The quantum approximate optimization algorithm at high depth for maxcut on large-girth regular graphs and the Sherrington-Kirkpatrick model," *arXiv:2110.14206*, 2021.