Exploring the neighborhood of 1-layer QAOA with instantaneous quantum polynomial circuits

Sebastian Leontica [1,2,*] and David Amaro [1,†]

[1] *Quantinuum, Partnership House, Carlisle Place, London SW1P 1BX, United Kingdom*

[2] *CMMP Research Group, University College London, Department of Physics and Astronomy, Gower Street, London WC1E 6BT, United Kingdom*

(Received 14 October 2022; revised 6 November 2023; accepted 21 December 2023; published 19 January 2024)

We embed 1-layer QAOA circuits into the larger class of parametrized instantaneous quantum polynomial circuits to produce an improved variational quantum algorithm for solving combinatorial optimization problems. The use of analytic expressions to find optimal parameters classically makes our protocol robust against barren plateaus and hardware noise. The average overlap with the ground state scales as $2^{-0.31(2)N}$ with the number of qubits N for random Sherrington-Kirkpatrick (SK) Hamiltonians of up to 29 qubits, a polynomial improvement over 1-layer QAOA. Additionally, we observe that performing variational imaginary time evolution on the manifold approximates low-temperature pseudo-Boltzmann states. Our protocol outperforms 1-layer QAOA on the recently released Quantinuum H2 trapped-ion quantum hardware and emulator, where we obtain an average approximation ratio of 0.985 across 312 random SK instances of 7 to 32 qubits, from which almost 44% are solved optimally using only 4 to 1208 shots per instance.

DOI: 10.1103/PhysRevResearch.6.013071

I. INTRODUCTION

Since its introduction by Farhi *et al.* [1] in 2014, the quantum approximate optimization algorithm (QAOA) has been explored in the quantum computing literature as one of the most promising heuristics for achieving quantum advantage on near-term devices [2,3]. This is only one example of a larger class of variational quantum optimization algorithms, which attempt to produce good solutions to combinatorial optimization problems by sampling a parametrized quantum circuit [4–8]. In the absence of full quantum error correction [9], the required circuits must be sufficiently shallow to withstand noise, yet expressive enough to find states with high overlap onto the ground state. QAOA is a particularly good choice for satisfying these criteria, as it has an adjustable number of layers p. It can be understood as a Trotterized version of the quantum adiabatic algorithm (QAA), for which compelling theoretical evidence of performance exists [10]. Additionally, it was shown that even for small numbers of layers, sampling from the QAOA ansatz is a hard task for classical computers [11].

In this regime of a small number of layers, the form of the Trotterized QAOA operators may not be the best choice. This has motivated [12–15] the addition of extra parameters to the QAOA ansatz so that, instead of evolving the state according to the problem Hamiltonian, each parameter in the ansatz has the freedom to evolve independently. By doing this, an ansatz of the same depth may incorporate corrections that would otherwise require multiple layers.

In particular, 1-layer QAOA circuits—with and without the additional parametrization—belong to the class of parametrized quantum circuits known as weighted graph states (WGS) used to simulate condensed matter systems [16–22]. For these states, the reduced density matrix in a subsystem of fixed size can be computed classically, allowing the efficient evaluation of local observables on a classical computer. This property permits the derivation of analytic and exact expressions for 1-layer QAOA on arbitrary local Hamiltonians [23] and for extra-parametrized circuits on some restricted local Hamiltonians [12,13]. Such expressions are used to train the model classically, bypassing typical limitations such as the appearance of barren plateaus [24].

In this manuscript, we explore the embedding of 1-layer QAOA into the broader class of parametrized instantaneous quantum polynomial (IQP) circuits, for which similar hardness of sampling theorems exist [25,26], even in the presence of moderate noise [27]. IQP circuits also belong to the class of WGS, but compared to QAOA and existing extra-parametrized variants our ansatz uses all-to-all two-qubit interactions, making its implementation problem independent and most natural for trapped-ion quantum computers. We additionally show that analytic and exact expressions can be obtained for arbitrary local Hamiltonians, and use them to train the model via robust classical techniques like the Runge-Kutta method [28]. We emphasize the role of starting the training from the optimal QAOA and finding a nearby local minimum rather than aiming for a global optimum, which avoids the challenging exploration of nontrivial landscapes [29].

*sebastian.leontica.22@ucl.ac.uk

†david.amaro@quantinuum.com

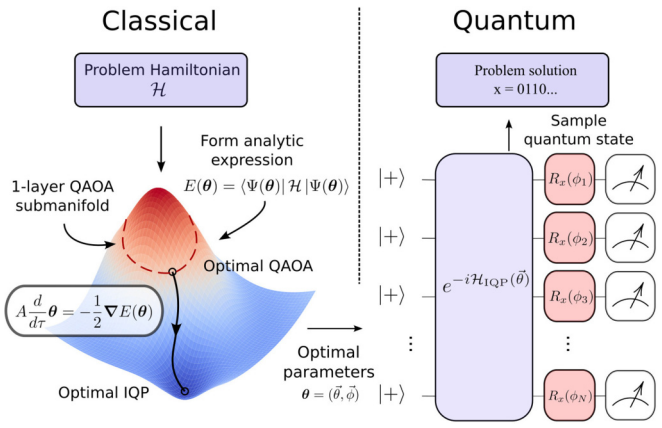

FIG. 1. Diagrammatic representation of the algorithm. The 1-layer QAOA ansatz is a submanifold of the IQP ansatz and provides a warm start in the optimization protocol. The trajectory between the QAOA optimum and the IQP optimum is defined via the McLachlan variational principle and is computed classically. Color coding the optimization landscape represents the effective temperature of the associated state, with lower temperature states (blue) having a higher chance of sampling the ground state. The quantum computer is only used during the sampling step, which is known to be difficult classically.

This leaves only the key ingredient of sampling from the final quantum state to be performed on the quantum device, as illustrated in Fig. 1. A recent investigation of the states produced by 1-layer QAOA [30] shows that sampling produces a distribution close to a Boltzmann distribution, at temperatures beyond the reach of classical sampling techniques such as Markov Chain Monte Carlo (MCMC) [31]. We improve on this result by lowering the temperature further, using variational quantum imaginary time evolution (VarQITE) [32,33]. However, the constraint of keeping the state in the variational manifold limits our ability to follow exact imaginary time evolution, distorting the distribution.

The manuscript is structured as follows. Section II provides a brief review of QAOA. Our IQP ansatz is introduced in Sec. III, where we make the connection to 1-layer QAOA, describe the derivation of analytical expressions and how to use them for classical training, and discuss a previous work [34] that challenges the possibility of quantum advantage with IQP circuits. In Sec. IV we describe our protocol for approximating thermal distributions and solving combinatorial optimization problems, while Sec. V presents numerical performance results. First, the average overlap with the ground state obtained with an exact state-vector simulator is polynomially better than for 1-layer QAOA on random Sherrington-Kirkpatrick (SK) Hamiltonians of up to 29 qubits. Second, when approximating thermal distributions we can reach lower temperatures than 1-layer QAOA but the approximation quality reduces. Third, we demonstrate a better performance than 1-layer QAOA at solving random SK Hamiltonians of up to 32 qubits in the recently released Quantinuum's trapped-ion H2 quantum hardware and emulator. Using a reduced number of shots, the best solution per instance presents a large approximation ratio and is optimal for a large fraction of instances. Finally, Sec. VI discusses the methods, results, and future research directions.

II. THE QUANTUM APPROXIMATE OPTIMIZATION ALGORITHM

The standard implementation of the QAOA [1] attempts to create states with large overlap onto the ground eigenspace of some optimization problem, typically defined through an Ising Hamiltonian,

$$\mathcal{H} = \sum_i h_i Z_i + \sum_{i<j} J_{ij} Z_i Z_j, \tag{1}$$

where the Z_i variables can be interpreted as the projections onto the z axis of a classical or quantum mechanical ensemble of N spin-$\frac{1}{2}$ particles, and (h_i, J_{ij}) are real coefficients. This is achieved by starting with the ground state $|+\rangle^{\otimes N}$ of the trivial transverse field mixing Hamiltonian $\mathcal{H}_x = -\sum_i X_i$ and evolving the state under the alternating application of the propagators of $\mathcal{H}$ and $\mathcal{H}_x$. The final trial state is of the form

$$|\Psi(\boldsymbol{\gamma}, \boldsymbol{\beta})\rangle = \prod_{k=1}^{p} \exp\left(i\beta_k \mathcal{H}_x\right) \exp\left(-i\gamma_k \mathcal{H}\right) |+\rangle^{\otimes N}, \tag{2}$$

where p is called the level of the QAOA and the sets $\boldsymbol{\beta}$, $\boldsymbol{\gamma}$ of real coefficients β_k, γ_k are used as variational parameters. The most commonly used cost function in the optimization of the ansatz is the expectation value of the problem Hamiltonian

$$E = \langle\Psi(\boldsymbol{\gamma}, \boldsymbol{\beta})|\,\mathcal{H}\,|\Psi(\boldsymbol{\gamma}, \boldsymbol{\beta})\rangle, \tag{3}$$

although alternative objective functions have been proposed [35,36]. For the rest of this work, we will only consider the 1-layer QAOA, which is sufficiently shallow to withstand the effects of moderate noise and obtains an enhanced average probability of sampling the ground state quadratically larger than random guessing [30], i.e., scaling as $2^{-0.5N}$.

III. THE INSTANTANEOUS QUANTUM POLYNOMIAL CIRCUIT

The IQP is a nonuniversal model of quantum computation with similar roots to the boson sampling problem, whose aim is to strengthen the general belief that quantum computers are more powerful than classical machines [25,26]. Under certain widely believed complexity-theoretic assumptions, sampling from the IQP state $H^{\otimes N}\exp(-i\mathcal{H}_{\rm IQP}(\vec{\theta}))\,|+\rangle^{\otimes N}$ in the computational basis of all qubits is a hard task for a classical computer [26]. Here the IQP Hamiltonian is defined as $\mathcal{H}_{\rm IQP}(\vec{\theta}) = \frac{1}{2}\sum_i \theta_i Z_i + \frac{1}{2}\sum_{i<j} \theta_{ij} Z_i Z_j$ and H is the Hadamard gate.

The IQP ansatz employed in this work is a generalization where Hadamard gates are replaced with independent parametrized single-qubit rotations $R_x(\phi) = \exp(-i\phi X/2)$, leading to the quantum circuit

$$|\Psi(\boldsymbol{\theta})\rangle = \bigotimes_{i\in\mathcal{N}} R_x(\phi_i) \cdot \exp(-i\mathcal{H}_{\rm IQP}(\vec{\theta}))|+\rangle^{\otimes N}, \tag{4}$$

where $\boldsymbol{\theta} = (\vec{\phi}, \vec{\theta})$ are free, real parameters. The IQP state is recovered by setting $\phi_i = \pi/2$ and making the transformation $\theta_i \longrightarrow \theta_i - \pi/2$. Since the IQP state can be brought to this form by modifying the final layer of single qubit rotations, we expect generic states of this form to be difficult to sample

classically as well. We also make the important observation that, up to single qubit rotations and energy rescaling, the IQP state in Eq. (4) is the same as that produced by a 1-layer QAOA designed to solve for the ground state of $\mathcal{H}_{\mathrm{IQP}}$.

This ansatz generalizes the optimization cost function of Eq. (3) to

$$E(\boldsymbol{\theta}) = \langle \Psi(\boldsymbol{\theta})| \mathcal{H} |\Psi(\boldsymbol{\theta})\rangle = \langle \mathcal{H} \rangle_{\boldsymbol{\theta}}, \tag{5}$$

which we refer to as the optimization landscape. The task of computing the cost function defined in Eq. (5) is then reduced to estimating the expectation values of the spins $\langle Z_i \rangle_{\boldsymbol{\theta}}$ and correlators $\langle Z_i Z_j \rangle_{\boldsymbol{\theta}}$ in an arbitrary state $|\Psi(\boldsymbol{\theta})\rangle$. In the Supplemental Material [37] we show that the latter expression can be reduced to calculating partition functions of reduced Ising Hamiltonians of the form

$$\mathcal{Z}_e = \frac{1}{2^N} \sum_{\{x\}} e^{-i\mathcal{H}_e(x)}, \tag{6}$$

where e's are single or two qubit subsets. The reduced generator $\mathcal{H}_e$ retains only the terms in $\mathcal{H}_{\mathrm{IQP}}$ that anticommute with the operator $X_e = \bigotimes_{i \in e} X_i$. This leads to a highly restricted graph topology, for which partition functions can be evaluated exactly. We generalize this method to show that IQPs have simple analytic expressions for all expectation values of the form $\langle Z_e \rangle_{\boldsymbol{\theta}}$, with a number of terms that scales like $\mathcal{O}(2^{|e|})$.

These properties of the IQP ansatz make it a good candidate for solving optimization problems, as it is guaranteed to be at least as powerful as 1-layer QAOA and the training can be performed efficiently using only classical resources. The access to exact, analytic expressions for the cost function also means we do not need to worry about finite sampling or device errors during training. Barren plateau issues can also be ruled out, as we can evaluate gradients to arbitrary precision and use adaptive step sizes. Access to a quantum computer is only necessary during the final sampling step, so we expect our protocol to perform well under moderate hardware noise.

As opposed to the standard QAOA ansatz, the IQP is sufficiently flexible to produce all computational states. In particular, this means that if a classical algorithm were able to find the global optimum of Eq. (5), it would also find the exact ground state of $\mathcal{H}$. In [34], it is shown that the optimization landscapes of IQP ansatz with only polynomially many terms (like our ansatz) are generally nonconvex, and computational states other than the solution may form local minima, which we call trivial minima. Consequently, converging to such local minima would imply the algorithm does not need access to a quantum computer, as the bits x_i of the solution corresponding to the optimal parameters are given by $\langle Z_i \rangle_{\boldsymbol{\theta}}$, which can be efficiently computed classically.

We prove that the optimization landscapes can contain nontrivial minima, and give a minimal example of this in the Supplemental Material [37]. Remarkably, we provide numerical evidence that for the SK model such a local minimum is located in the vicinity of the QAOA parameters, and show that sampling the IQP circuit at this point greatly enhances the chance of finding the ground state compared to QAOA.

IV. METHODS

A remarkable result of [30] is that for a wide range of optimization problems that can be formulated as in Eq. (1), the 1-layer QAOA is capable of approximating pseudo-Boltzmann states proportional to $\exp(-\beta\mathcal{H}/2)|+\rangle^{\otimes N}$, with large inverse temperature β, up to relative phases that do not affect the distribution. This is important because sampling this state produces the same distribution as sampling the mixed thermal state $\rho_\beta = e^{-\beta\mathcal{H}}/\mathcal{Z}$ for classical Hamiltonians, which is useful for a variety of optimization tasks.

In our work, we use this result to justify the QAOA as a good starting point in optimizing the IQP ansatz. Since the 1-layer QAOA ansatz can be recovered by restricting the parameters of the full IQP, we find the optimal QAOA position classically, using the BFGS [38] algorithm on the submanifold. To find a local optimum in the vicinity of this position, it is sufficient to use simple gradient descent. However, we also explore the feasibility of our algorithm for producing low-energy thermal states, which is achieved using a different approach called VarQITE [32,33]. This protocol aims to find the trajectory on the manifold that best approximates the action of $\exp(-\tau\mathcal{H})$ on the state. If the initial state is pseudo-Boltzmann, then applying this operator leads to a decrease in temperature. The parameters in the ansatz are evolved according to the McLachlan variational principle [39]:

$$A\frac{d}{d\tau}\boldsymbol{\theta} = -\frac{1}{2}\nabla E(\boldsymbol{\theta}), \tag{7}$$

where the coupling matrix A describes the geometry of the variational manifold (i.e., it is the Gram matrix of the tangent vectors corresponding to each parameter) and τ is the imaginary time variable. In the Supplemental Material [37] we show that the coefficients of the Gram matrix can be expressed as expectation values of low-weight Pauli operators in the IQP, for which we find simple analytic expressions. However, this calculation is computationally expensive, so when the focus is on finding a local minimum rather than preserving a thermal profile, we set $A = I$ and perform simple gradient descent.

In both cases, this linear system of ODEs defines a flow on the variational manifold, that we solve numerically using the Runge-Kutta method [28]. We stop this procedure when we arrive at a local minimum, or when A becomes noninvertible. This typically happens after a long plateau in the energy profile, which we illustrate in the Supplemental Material [37]. Such an event becomes a rare occurrence when we increase the number of qubits, but for problems that exhibit this behavior, we choose the optimal parameters in the middle of the plateau. After finding the optimal parameters, we sample the circuit and compute the probability of finding the ground state. We share the code used for implementing this protocol in [40].

We characterize our distributions using an effective inverse temperature β. This is obtained by minimizing the Kullbach-Leibler divergence of the IQP distribution to the family of thermal distributions. Here, we compute the KL divergence exactly, but in practice, this would be estimated from samples [41].

V. RESULTS

We test our method on Sherrington-Kirkpatrick (SK) Hamiltonians [42–44] of up to $N = 29$ spins using Qiskit exact state-vector quantum simulators [45]. These Hamiltonians are of the form of Eq. (1) with $h_i = 0$ and J_{ij} indepen-

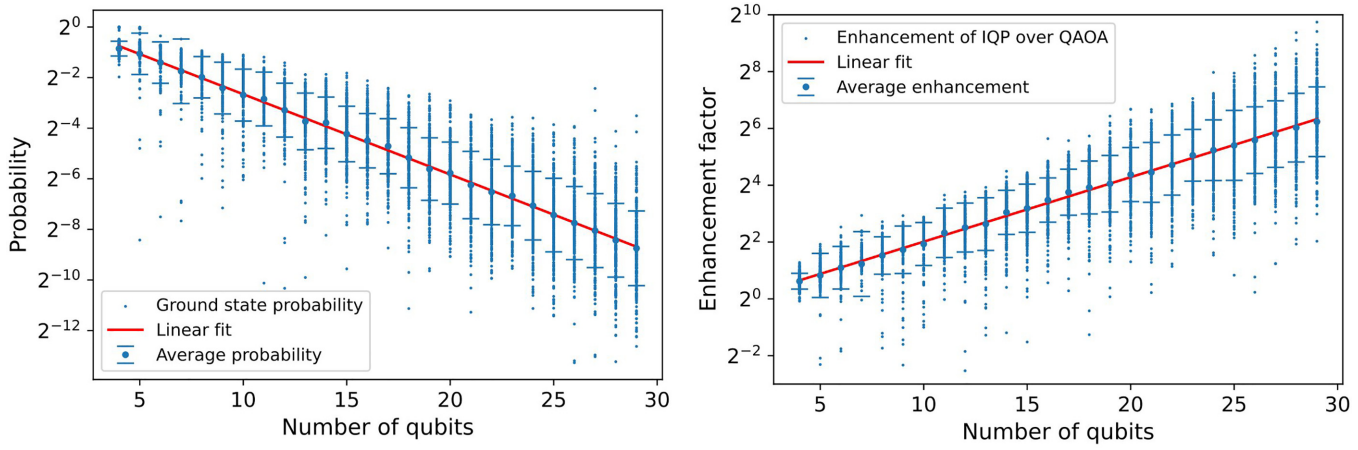

FIG. 2. Optimization results for 300 randomly generated Sherrington-Kirkpatrick Hamiltonians of up to 29 spins. (a) Probability of sampling the ground state configuration in the optimal IQP ansatz. (b) Enhancement factor $p_{\mathrm{IQP}}/p_{\mathrm{QAOA}}$ for finding the ground state in the optimized IQP ansatz compared to the original QAOA. The IQP was optimized until convergence using simple gradient descent. Using a linear fit, we find the average probability of sampling the ground state $p_{\mathrm{IQP}} \sim 2^{-\alpha N}$ with $\alpha = 0.31 \pm 0.02$ and the average enhancement factor $p_{\mathrm{IQP}}/p_{\mathrm{QAOA}} \sim 2^{\delta N}$ with $\delta = 0.23 \pm 0.02$. The errors indicate the variability in gradient at one standard deviation.

dent and identically distributed Gaussian random variables of zero mean and a standard deviation of $1/\sqrt{N}$. The unbiased SK model presents a $\mathbb{Z}_2$ symmetry, so the ground state is unique up to flipping all qubits. This is a well-understood spin model with compelling classical solvers [46]. In quantum optimization it is one of the most studied benchmark problems [30,47–51].

In Fig. 2, we show how the overlap of the optimized IQP state onto the ground eigenspace varies with the problem size, and how it compares to the overlap achieved by the initial QAOA. Both plots show a clear exponential trend with relatively low and slowly increasing variance. This confirms that our algorithm has a significantly better exponential scaling compared to 1-layer QAOA.

We also study how the temperature of the distribution changes as we perform imaginary time evolution on our variational manifold up to time $\tau = 10$, close to convergence. In Fig. 3 we show how the optimal normalized temperatures achieved by the final optimized IQP state are lower than those achieved by the starting QAOA state. However, the KL divergence between the optimized IQP state and the best-fitting thermal state is higher and presents more dispersion than QAOA across Hamiltonian instances. This indicates that IQP states might be beneficial for the task of sampling low-energy eigenstates while QAOA provides a better approximation to thermal distributions.

In Fig. 4 we plot example distributions produced by the QAOA and the optimized IQP ansatz. From the qualitative aspect of the IQP distribution, we see that the performance of our algorithm in increasing the ground state overlap cannot be entirely explained as a consequence of having a lower temperature. The distribution becomes arched, and the probabilities of sampling the low-energy eigenstates rise orders of magnitude above the predictions of the thermal fit. Future theoretical work is necessary to understand how this effect emerges, and whether it is recovered in more general optimization problems.

Our algorithm is also studied in a more realistic setting, where quantum circuits are affected by hardware noise. We use the recently released Quantinuum H2 trapped-ion quantum hardware and emulator [52]. The emulator performs exact state-vector simulations under a noise model that replicates the noisy behavior of the real device. The device presents all-to-all connectivity and high-fidelity parametrized gates of the form $\exp(-i\theta ZZ)$, making it ideal for our protocol and for QAOA on densely connected Hamiltonians.

For this analysis, we study biased SK models with the coefficients h_i independently sampled from the same Gaussian distribution as the coefficients J_{ij}. The presence of the bias breaks the $\mathbb{Z}_2$ symmetry, halving the initial overlap with the ground eigenspace and making the problem slightly more general. The bias adds an additional slope in the vicinity of

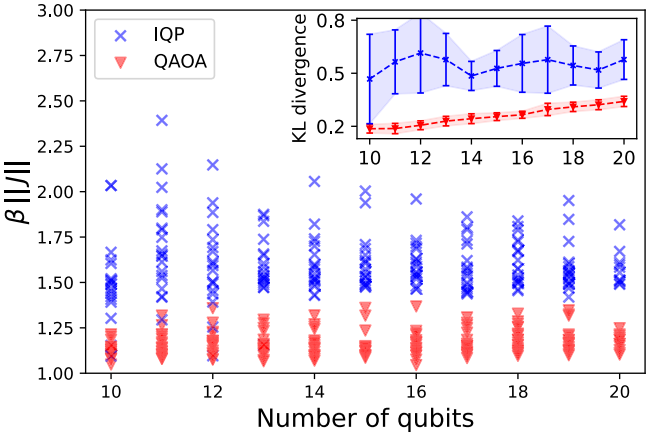

FIG. 3. Normalized effective inverse temperatures $\beta||J||$ in the QAOA state and the IQP state after VarQITE evolution for a time $\tau = 10$, for 20 randomly generated Sherrington-Kirkpatrick Hamiltonians of each size from 10 to 20 qubits. We also show the average and standard deviations for the KL divergences of each problem size.

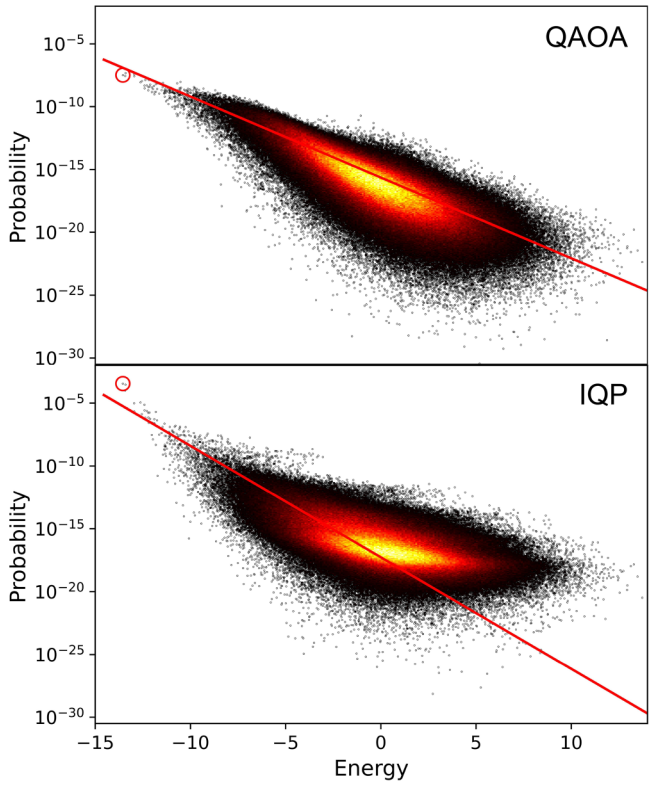

FIG. 4. Overlap of the state produced by our ansatz onto different Hamiltonian eigenvalues as a function of energy for the QAOA parameters (top) and optimized IQP parameters (bottom), for a randomly generated 20 qubit Sherrington-Kirkpatrick Hamiltonian. Brighter color indicates higher coarse-grain point density. The red line illustrates the thermal distribution model that minimizes the KL divergence. A red circle marks the location of the ground state.

QAOA, that sometimes dissolves the local minimum that we exploit in the previous study, leaving no obvious method to pick a point in the gradient-descent trajectory. Our aim for this analysis is, however, to study the performance in the neighborhood of QAOA, rather than providing the most optimized form of our protocol. For this purpose we pick the optimized 1-layer QAOA as the first circuit, and three equally spaced circuits corresponding to three of the first gradient-descent steps. The Supplemental Material [37] describes the criterion we used to pick these circuits.

Figure 5 compares the quality of the best solutions obtained by the corresponding four circuits. From the 312 instances, we optimally solve 5, 21, 59, and 86, respectively, for the four circuits. The best solution sampled for each instance has an average approximation ratio and standard deviation of (0.87, 0.10), (0.935, 0.083), (0.948, 0.083), and (0.970, 0.060), respectively. When considering for each instance only the best solution obtained from the four circuits as the output of our algorithm, 136 instances are solved optimally (almost 44%) and the distribution has an average approximation ratio and standard deviation of (0.985, 0.029).

VI. DISCUSSION

The algorithm we introduce explores the natural connection between the 1-layer QAOA state and IQP circuits. Studying the vicinity of the QAOA in this broader variational manifold leads to a better understanding of its optimality as a shallow-depth quantum heuristic, as well as how it can be improved.

We show that, for the case of SK Hamiltonians, our approach amplifies the probability of sampling the ground state, beyond what can be obtained using classical tools such as MCMC. The hardware implementation is as resource-demanding as it is for 1-layer QAOA, and parameter training can be performed classically in time $\mathcal{O}(N^3)$. Results on the Quantinuum H2 show the reliability of our protocol to solve large instances with scarce quantum resources.

We leave as a future work the development of an optimized strategy to pick points along the gradient-descent trajectory

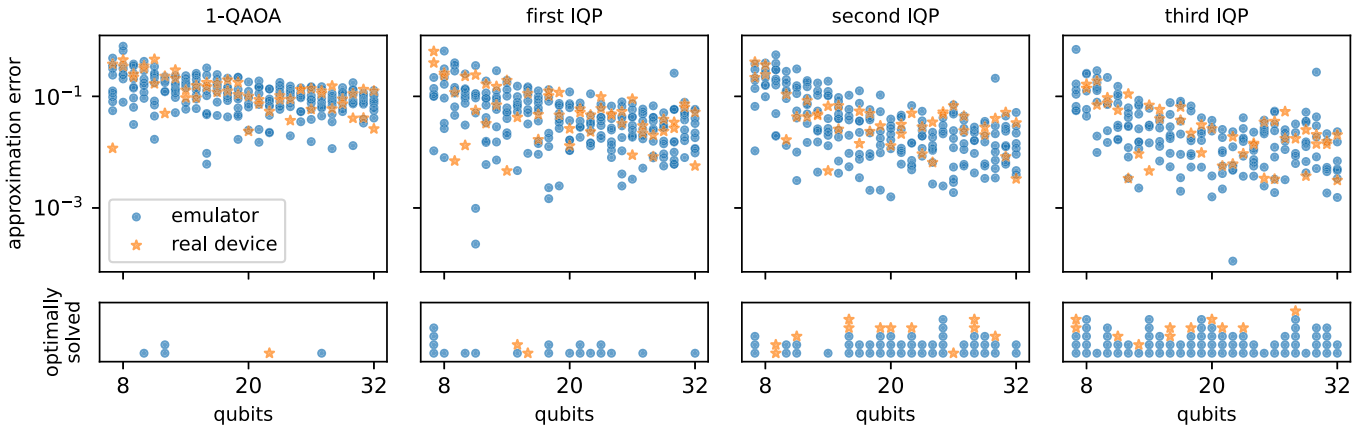

FIG. 5. Optimization results on the Quantinuum H2 trapped-ion quantum hardware and emulator for randomly generated biased Sherrington-Kirkpatrick Hamiltonians of 7 to 32 qubits: two instances per problem size on the device (stars) and ten instances on the emulator (circles). For each instance we pick four steps along the gradient-descent trajectory, corresponding to the standard 1-layer QAOA, and three IQP circuits. Then take $\sim 2^{0.32N} \in [4, 1208]$ shots, equally distributed across the four circuits. Each data point in the figure corresponds to the best solution sampled for each instance. If the best solution is optimal the point is placed in the lower row, while for suboptimal solutions we place the point in the upper row to visualize the approximation error.

where sampling from the quantum computer might yield even better performance.

The results presented motivate the development of strategies to compare the performance of our protocol against state-of-the-art classical algorithms at the scale of real-world combinatorial optimization problems. For example, the access to the analytical expectation value of the problem Hamiltonian and higher powers of it might provide an efficient way to estimate the probability of sampling the low-energy tail for large-scale problems.

ACKNOWLEDGMENTS

We thank M. Lubasch, L. Coopmans, M. Benedetti, M. Rosenkranz, M. DeCross, M. Foss-Feig, and C. Cirstoiu for their feedback. We appreciate the helpful discussions with C. Arenz and thank the authors of [30] for sharing their results with us. Finally, we acknowledge the great work of the Quantinuum H2 developing team: S. Moses, M. Mills, J. Johansen, J. Dreiling, J. Gaebler, P. Siegfried, J. Pino, and C. Figgatt.

[1] E. Farhi, J. Goldstone, and S. Gutmann, A quantum approximate optimization algorithm, arXiv:1411.4028.

[2] E. Farhi, J. Goldstone, and S. Gutmann, A quantum approximate optimization algorithm applied to a bounded occurrence constraint problem, arXiv:1412.6062.

[3] Z. Jiang, E. G. Rieffel, and Z. Wang, Near-optimal quantum circuit for Grover's unstructured search using a transverse field, Phys. Rev. A **95**, 062317 (2017).

[4] P. Díez-Valle, D. Porras, and J. J. García-Ripoll, Quantum variational optimization: The role of entanglement and problem hardness, Phys. Rev. A **104**, 062426 (2021).

[5] X. Liu, A. Angone, R. Shaydulin, I. Safro, Y. Alexeev, and L. Cincio, Layer VQE: A variational approach for combinatorial optimization on noisy quantum computers, IEEE Trans. Quantum Eng. **3**, 1 (2022).

[6] S. Ebadi, A. Keesling, M. Cain, T. T. Wang, H. Levine, D. Bluvstein, G. Semeghini, A. Omran, J.-G. Liu, R. Samajdar, X.-Z. Luo, B. Nash, X. Gao, B. Barak, E. Farhi, S. Sachdev, N. Gemelke, L. Zhou, S. Choi, H. Pichler *et al.*, Quantum optimization of maximum independent set using Rydberg atom arrays, Science **376**, 1209 (2022).

[7] D. Amaro, C. Modica, M. Rosenkranz, M. Fiorentini, M. Benedetti, and M. Lubasch, Filtering variational quantum algorithms for combinatorial optimization, Quantum Sci. Technol. **7**, 015021 (2022).

[8] D. Amaro, M. Rosenkranz, N. Fitzpatrick, K. Hirano, and M. Fiorentini, A case study of variational quantum algorithms for a job shop scheduling problem, EPJ Quantum Technol. **9**, 5 (2022).

[9] D. Litinski, A game of surface codes: Large-scale quantum computing with lattice surgery, Quantum **3**, 128 (2019).

[10] E. Farhi, J. Goldstone, S. Gutmann, J. Lapan, A. Lundgren, and D. Preda, A quantum adiabatic evolution algorithm applied to random instances of an NP-complete problem, Science **292**, 472 (2001).

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

[12] R. Herrman, P. C. Lotshaw, J. Ostrowski, T. S. Humble, and G. Siopsis, Multi-angle quantum approximate optimization algorithm, Sci. Rep. **12**, 6781 (2022).

[13] K. Shi, R. Herrman, R. Shaydulin, S. Chakrabarti, M. Pistoia, and J. Larson, Multiangle qaoa does not always need all its angles, in *2022 IEEE/ACM 7th Symposium on Edge Computing (SEC)* (2022), pp. 414–419.

[14] M. Chalupnik, H. Melo, Y. Alexeev, and A. Galda, Augmenting qaoa ansatz with multiparameter problem-independent layer, in *2022 IEEE International Conference on Quantum Computing and Engineering (QCE)* (2022), pp. 97–103.

[15] P. C. Lotshaw, G. Siopsis, J. Ostrowski, R. Herrman, R. Alam, S. Powers, and T. S. Humble, Approximate Boltzmann distributions in quantum approximate optimization, Phys. Rev. A **108**, 042411 (2023).

[16] S. Anders, M. B. Plenio, W. Dür, F. Verstraete, and H.-J. Briegel, Ground-state approximation for strongly interacting spin systems in arbitrary spatial dimension, Phys. Rev. Lett. **97**, 107206 (2006).

[17] S. Anders, H. J. Briegel, and W. Dür, A variational method based on weighted graph states, New J. Phys. **9**, 361 (2007).

[18] L. Hartmann, J. Calsamiglia, W. Dür, and H. J. Briegel, Weighted graph states and applications to spin chains, lattices and gases, J. Phys. B: At. Mol. Opt. Phys. **40**, S1 (2007).

[19] A. D. K. Plato, O. C. Dahlsten, and M. B. Plenio, Random circuits by measurements on weighted graph states, Phys. Rev. A **78**, 042332 (2008).

[20] R. Hübener, C. Kruszynska, L. Hartmann, W. Dür, F. Verstraete, J. Eisert, and M. B. Plenio, Renormalization algorithm with graph enhancement, Phys. Rev. A **79**, 022317 (2009).

[21] R. Hübener, C. Kruszynska, L. Hartmann, W. Dür, M. B. Plenio, and J. Eisert, Tensor network methods with graph enhancement, Phys. Rev. B **84**, 125103 (2011).

[22] P. M. Schindler, T. Guaita, T. Shi, E. Demler, and J. I. Cirac, Variational ansatz for the ground state of the quantum Sherrington-Kirkpatrick model, Phys. Rev. Lett. **129**, 220401 (2022).

[23] Z. Wang, S. Hadfield, Z. Jiang, and E. G. Rieffel, Quantum approximate optimization algorithm for MaxCut: A fermionic view, Phys. Rev. A **97**, 022304 (2018).

[24] M. Cerezo, A. Sone, T. Volkoff, L. Cincio, and P. J. Coles, Cost function dependent barren plateaus in shallow parametrized quantum circuits, Nat. Commun. **12**, 1791 (2021).

[25] D. Shepherd and M. J. Bremner, Temporally unstructured quantum computation, Proc. Math. Phys. Eng. Sci. **465**, 1413 (2009).

[26] M. J. Bremner, A. Montanaro, and D. J. Shepherd, Average-case complexity versus approximate simulation of commuting quantum computations, Phys. Rev. Lett. **117**, 080501 (2016).

[27] M. J. Bremner, A. Montanaro, and D. J. Shepherd, Achieving quantum supremacy with sparse and noisy commuting quantum computations, Quantum **1**, 8 (2017).

[28] J. Dormand and P. Prince, A family of embedded Runge-Kutta formulae, J. Comput. Appl. Math. **6**, 19 (1980).

[29] L. Bittel and M. Kliesch, Training variational quantum algorithms is NP-hard, Phys. Rev. Lett. **127**, 120502 (2021).
[30] P. Díez-Valle, D. Porras, and J. J. García-Ripoll, Quantum approximate optimization algorithm pseudo-Boltzmann states, Phys. Rev. Lett. **130**, 050601 (2023).
[31] R. Eldan, F. Koehler, and O. Zeitouni, A spectral condition for spectral gap: Fast mixing in high-temperature ising models, arXiv:2007.08200.
[32] X. Yuan, S. Endo, Q. Zhao, Y. Li, and S. C. Benjamin, Theory of variational quantum simulation, Quantum **3**, 191 (2019).
[33] S. McArdle, T. Jones, S. Endo, Y. Li, S. C. Benjamin, and X. Yuan, Variational ansatz-based quantum simulation of imaginary time evolution, npj Quantum Inf. **5**, 75 (2019).
[34] J. Lee, A. B. Magann, H. A. Rabitz, and C. Arenz, Progress toward favorable landscapes in quantum combinatorial optimization, Phys. Rev. A **104**, 032401 (2021).
[35] L. Li, M. Fan, M. Coram, P. Riley, and S. Leichenauer, Quantum optimization with a novel Gibbs objective function and ansatz architecture search, Phys. Rev. Res. **2**, 023074 (2020).
[36] P. K. Barkoutsos, G. Nannicini, A. Robert, I. Tavernelli, and S. Woerner, Improving variational quantum optimization using CVaR, Quantum **4**, 256 (2020).
[37] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevResearch.6.013071 for analytic expression of expectation values of the energy and Gram matrix in the IQP ansatz, a proof that IQP landscapes can have non-trivial local minima and guidelines for selecting the IQP circuit along the evolution to sample from.
[38] R. Fletcher, *Practical Methods of Optimization* (John Wiley & Sons, New York, NY, 1987).
[39] A. McLachlan, A variational solution of the time-dependent Schrodinger equation, Mol. Phys. **8**, 39 (1964).
[40] S. Leontica, github.com/LeonQu4nt/iqp_optimisation (2022).
[41] M. Benedetti, J. Realpe-Gómez, R. Biswas, and A. Perdomo-Ortiz, Estimation of effective temperatures in quantum annealers for sampling applications: A case study with possible applications in deep learning, Phys. Rev. A **94**, 022308 (2016).
[42] D. Sherrington and S. Kirkpatrick, Solvable model of a spin-glass, Phys. Rev. Lett. **35**, 1792 (1975).
[43] G. Parisi, A sequence of approximated solutions to the S-K model for spin glasses, J. Phys. A: Math. Theor. **13**, L115 (1980).
[44] D. Panchenko, The Sherrington-Kirkpatrick model: An overview, J. Stat. Phys. **149**, 362 (2012).
[45] G. Aleksandrowicz, T. Alexander, P. Barkoutsos, L. Bello, Y. Ben-Haim, D. Bucher, F. J. Cabrera-Hernández, J. Carballo-Franquis *et al.*, Qiskit: An open-source framework for quantum computing, doi: 10.5281/zenodo.2562111 (2019).
[46] A. Montanari, Optimization of the Sherrington-Kirkpatrick hamiltonian, in *2019 IEEE 60th Annual Symposium on Foundations of Computer Science (FOCS)* (2019), pp. 1417–1433.
[47] K. J. Sung, J. Yao, M. P. Harrigan, N. C. Rubin, Z. Jiang, L. Lin, R. Babbush, and J. R. McClean, Using models to improve optimizers for variational quantum algorithms, Quantum Sci. Technol. **5**, 044008 (2020).
[48] M. P. Harrigan, K. J. Sung, M. Neeley, K. J. Satzinger, F. Arute, K. Arya, J. Atalaya, J. C. Bardin, R. Barends, S. Boixo, M. Broughton, B. B. Buckley, D. A. Buell, B. Burkett, N. Bushnell, Y. Chen, Z. Chen, B. Chiaro, R. Collins, W. Courtney *et al.*, Quantum approximate optimization of non-planar graph problems on a planar superconducting processor, Nat. Phys. **17**, 332 (2021).
[49] R. Babbush, J. R. McClean, M. Newman, C. Gidney, S. Boixo, and H. Neven, Focus beyond quadratic speedups for error-corrected quantum advantage, PRX Quantum **2**, 010103 (2021).
[50] A. M. Dalzell, N. Pancotti, E. T. Campbell, and F. G. S. L. Brandão, Mind the gap: Achieving a super-Grover quantum speedup by jumping to the end, in *Proceedings of the 55th Annual ACM Symposium on Theory of Computing, Orlando FL, USA* (Association for Computing Machinery, New York, 2023).
[51] E. Farhi, J. Goldstone, S. Gutmann, and L. Zhou, The quantum approximate optimization algorithm and the Sherrington-Kirkpatrick model at infinite size, Quantum **6**, 759 (2022).
[52] Quantinuum H2, https://www.quantinuum.com/, experiments performed between February 14 and March 16 2023.