Polynomial scaling of the quantum approximate optimization algorithm for ground-state preparation of the fully connected *p*-spin ferromagnet in a transverse field

Matteo M. Wauters[⬤],[1] Glen B. Mbeng,[2] and Giuseppe E. Santoro[3,4,5]

[1]*Niels Bohr International Academy and Center for Quantum Devices, University of Copenhagen, DK-2100 Copenhagen, Denmark*
[2]*Institut für Theoretische Physik, Universität Innsbruck, Technikerstrasse 21 a, A-6020 Innsbruck, Austria*
[3]*SISSA, Via Bonomea 265, I-34136 Trieste, Italy*
[4]*International Centre for Theoretical Physics, P.O.Box 586, I-34014 Trieste, Italy*
[5]*CNR-IOM Democritos National Simulation Center, Via Bonomea 265, I-34136 Trieste, Italy*

We show that the quantum approximate optimization algorithm (QAOA) can construct, with polynomially scaling resources, the ground state of the fully connected *p*-spin Ising ferromagnet, a problem that notoriously poses severe difficulties to a vanilla quantum annealing (QA) approach due to the exponentially small gaps encountered at first-order phase transition for $p \geqslant 3$. For a target ground state at arbitrary transverse field, we find that an appropriate QAOA parameter initialization is necessary to achieve good performance of the algorithm when the number of variational parameters $2P$ is much smaller than the system size N because of the large number of suboptimal local minima. Instead, when P exceeds a critical value $P_N^* \propto N$, the structure of the parameter space simplifies, as all minima become degenerate. This allows achieving the ground state with perfect fidelity with a number of parameters scaling extensively with N and with resources scaling polynomially with N.

I. INTRODUCTION

Efficient optimization and ground-state (GS) preparation are two of the most prominent issues in the growing field of quantum technology [1,2]. Optimization is a long-standing problem in physics and computer science and lies at the root of the efforts to show a possible "quantum supremacy" [3–5] over classical algorithms. A robust-state preparation strategy, in turn, would be a crucial tool for quantum technologies and would also allow "solving," using quantum hardware, many long-standing problems in condensed-matter theory and quantum chemistry [6–8]. The two are intimately connected, as many optimization tasks can be reformulated in terms of finding the classical ground state of an appropriate spin-glass Hamiltonian [9].

A traditional tool in this field has been quantum annealing [10–13] (QA), also known as adiabatic quantum computation [14,15], which relies on the adiabatic theorem to find the ground state of a target Hamiltonian, starting from a trivial initial state. Although QA appeared to be more efficient than its classical counterpart for certain problems [13,16–19], it is limited by the smallest gap encountered during the evolution, which vanishes, in the thermodynamic limit, when the system crosses a phase transition. In this context, the fully connected *p*-spin Ising ferromagnet is a simple but useful benchmark for optimization methods because a vanilla QA approach fails due to the exponentially small gap at the first-order phase transition encountered for $p \geqslant 3$ [20–22]. Several techniques have been advocated to overcome the slowness induced by such an exponentially small gap, including the

introduction of nonstoquastic terms [23,24], pausing [25], dissipative effects [26,27], and approximated counterdiabatic driving [28]. Their successful application, however, often depends on the knowledge of the spectrum or on the phase diagram of the model, thus making these techniques highly problem specific.

Recent alternative ground-state preparation approaches [29–31] rely on hybrid quantum-classical variational techniques [32] to tackle such problems, avoiding the possible limitations imposed by a QA adiabatic evolution. In this work, we will focus on one such scheme, the quantum approximate optimization algorithm (QAOA) [30,33,34].

The core idea of QAOA is to write a trial wave function as a product of many unitary operators, each depending on a classical variational parameter, applied to a state simple to construct, usually a product state with spins aligned in the *x* direction. Quantum hardware performs the discrete quantum dynamics and measures the expectation value of the target Hamiltonian, which is then minimized by an external classical algorithm, as a real function in a high-dimensional parameter space.

Although QAOA is a universal computational scheme [35], its performance strongly depends on the details of the target Hamiltonian. QAOA seems to perform rather well on Max-Cut problems [34] and on short-range spin systems [36,37]. The Grover search problem has also been studied within QAOA, showing that it leads to the optimal square-root speedup with respect to classical algorithms [38]. For generic long-range Hamiltonians, however, many open questions remain. The questions concern, in particular, the efficiency of

the algorithm when a large number of unitaries are employed, the ability to deal with first-order phase transitions, and the existence of "smooth" sets of optimal parameters [34,39,40]. Addressing these issues, an essential step towards experimental implementations of QAOA in realistic problems, will be the goal of our work. We will show that QAOA can construct, with polynomially scaling resources, the ground state of the fully connected p-spin Ising ferromagnet for all $p \geqslant 2$, hence including the case where a first-order phase transition occurs. For a generic target state, we find that an appropriate QAOA parameter initialization is necessary to achieve good performance of the algorithm when the number of variational parameters $2P$ is much smaller than the system size N because of the large number of suboptimal local minima. Finally, we show that when $P > P_N^* \propto N$, the structure of the parameter space simplifies, and all minima become degenerate. This allows us to achieve the ground state with perfect fidelity with a number of parameters scaling extensively with N and with resources scaling polynomially with N.

The rest of the paper is organized as follows. In Sec. II we present the model and describe the QAOA algorithm. In Sec. III we present our main analytical and numerical results: some technical analytical details are reported in Appendix A. Finally, we draw our conclusions and discuss future outlooks in Sec. IV.

II. MODEL AND QAOA ALGORITHM

As a benchmark for QAOA on long-range models we focus on the ferromagnetic fully connected p-spin model [20–22,41,42]:

$$\widehat{H}_{\text{target}} = -\frac{1}{N^{p-1}} \left(\sum_{j=1}^{N} \hat{\sigma}_j^z \right)^p - h \left(\sum_{j=1}^{N} \hat{\sigma}_j^x \right), \quad (1)$$

where $\hat{\sigma}_j^{x,z}$ are Pauli matrices at site j, N is the total number of sites, and h is a transverse field. This model displays, for $p = 2$, a second-order quantum phase transition at a critical transverse field $h_c = 2$ from a paramagnetic ($h > h_c$) to a symmetry-broken ferromagnetic phase ($h < h_c$). The transition becomes first order for $p > 2$, and h_c decreases for increasing p, with $h_c \to 1$ for $p \to \infty$ [21].

The QAOA algorithm [30] is a variational method to find the ground state of a target Hamiltonian $\widehat{H}_{\text{target}}$. Starting from an initial spin state polarized along the $\hat{\mathbf{x}}$ direction $|+\rangle = 2^{-N/2}(|\uparrow\rangle + |\downarrow\rangle)^{\otimes N}$, QAOA writes the following variational ansatz:

$$|\psi_P(\boldsymbol{\gamma}, \boldsymbol{\beta})\rangle = e^{-i\beta_P \widehat{H}_x} e^{-i\gamma_P \widehat{H}_z} \cdots e^{-i\beta_1 \widehat{H}_x} e^{-i\gamma_1 \widehat{H}_z} |+\rangle \quad (2)$$

in terms of $2P$ variational parameters $\boldsymbol{\gamma} = (\gamma_1 \cdots \gamma_P)$ and $\boldsymbol{\beta} = (\beta_1 \cdots \beta_P)$, where $\widehat{H}_z$ and $\widehat{H}_x$ are noncommuting Hamiltonians depending on the problem we wish to solve. Here we take $\widehat{H}_x = -\sum_j \hat{\sigma}_j^x$, the standard transverse field term, and an interaction term $\widehat{H}_z$,

$$\widehat{H}_z = -\left(\sum_{j=1}^{N} \hat{\sigma}_j^z \right)^p, \quad (3)$$

chosen for convenience to have a superextensive form with an integer spectrum. These choices allow us to restrict the

parameter space for γ_m and β_m to the interval $[0, \pi]$. In each QAOA run the variational energy cost function

$$E_P(\boldsymbol{\gamma}, \boldsymbol{\beta}) = \langle \psi_P(\boldsymbol{\gamma}, \boldsymbol{\beta})| \widehat{H}_{\text{target}} |\psi_P(\boldsymbol{\gamma}, \boldsymbol{\beta})\rangle \quad (4)$$

is minimized, until convergence to a local minimum $(\boldsymbol{\gamma}^*, \boldsymbol{\beta}^*)$ is obtained. The quality of the variational solution is gauged by computing the residual energy density [37]

$$\epsilon_P^{\text{res}}(\boldsymbol{\gamma}^*, \boldsymbol{\beta}^*) = \frac{E_P(\boldsymbol{\gamma}^*, \boldsymbol{\beta}^*) - E_{\min}}{E_{\max} - E_{\min}}, \quad (5)$$

where $E_{\min}$ and $E_{\max}$ are the smallest and largest eigenvalues, respectively, of the target Hamiltonian.

The connection with a QA approach is interesting [37]. In QA one would write an interpolating Hamiltonian [15] $\widehat{H}(s) = s\widehat{H}_{\text{target}} + (1 - s)\widehat{H}_x$, with $s(t)$ driven from $s(0) = 0$ to $s(\tau) = 1$ in a sufficiently large annealing time τ. A lowest-order Trotter decomposition of the corresponding step-discretized evolution operator—with $s_{m=1\cdots P}$ being constant for a time interval $\Delta t_{m=1\cdots P}$—would then result in a state of the form of Eq. (2) with

$$\gamma_m = \frac{s_m \Delta t_m}{\hbar} \frac{1}{N^{p-1}}, \quad \beta_m = \frac{\Delta t_m}{\hbar}[1 - s_m(1 - h)], \quad (6)$$

where the total evolution time, i.e., the run time associated with the set of $2P$ evolution operators, would be given by

$$\frac{\tau}{\hbar} = \sum_{m=1}^{P} \frac{\Delta t_m}{\hbar} = \sum_{m=1}^{P} [\beta_m + (1 - h)\gamma_m N^{p-1}]. \quad (7)$$

While an optimization of the parameters s_m and Δt_m is, in principle, possible, the standard linear schedule $s(t) = t/\tau$ would result in a digitized-QA scheme where $s_m = m/P$ and $\Delta t_m = \Delta t = \tau/P$ [43,44]. With these choices, a convenient starting point for the QAOA optimization algorithm would be to take $\gamma_m^0 = \frac{\Delta t}{\hbar} \frac{m}{P} \frac{1}{N^{p-1}}$ and $\beta_m^0 = \frac{\Delta t}{\hbar}[1 - \frac{m}{P}(1 - h)]$ with the possible addition of a small noise term. Alternatively, we might choose a completely random initial point with $\gamma_m^0, \beta_m^0 \in [0, \pi]$. Limiting the variational parameters in the interval $[0, \pi]$ is motivated by the symmetries of the function $E_P(\boldsymbol{\gamma}, \boldsymbol{\beta})$, which we describe in Appendix B. Thus, the total evolution time, Eq. (7), is limited by $\tau \leqslant \pi P(1 + N^{p-1})$. These two alternative choices will be henceforth referred to as l-init and r-init.

III. RESULTS

Reference [45] showed that the target ground state of the $p = 2$ fully connected Ising ferromagnet with $h = 0$, the so-called Lipkin-Meshov-Glick [46] model, can be perfectly constructed, with unit fidelity, with the shortest QAOA circuit, $P = 1$, if the number of sites N is odd. For even N instead, $P = 2$ is required to reach exactly the GS. Reference [47] recently showed that a whole class of spin-glass models can be constructed in which QAOA shows such a property. Here we show (see the detailed proof in Appendix A) that the general p-spin model in Eq. (1) belongs, for $h = 0$, to the class of $P = 1$ QAOA-solvable problems for odd N. The proof is based on finding a set of sufficient conditions to reach unit fidelity $\mathcal{F}(\gamma, \beta) = |\langle \psi_{\text{targ}}|\psi_{P=1}(\gamma, \beta)\rangle|^2 = 1$. This provides a set of parameters (γ, β) that can be used to prepare the exact

ground state for $h = 0$. For $P = 1$ the target state fidelity reads

$$\mathcal{F}(\gamma, \beta) = |\langle\psi_{\text{targ}}|e^{-i\beta\hat{H}_x}e^{-i\gamma\hat{H}_z}|+\rangle|^2$$

$$= |\frac{1}{\sqrt{2^N}}\sum_l e^{-i\gamma E_l}\langle\psi_{\text{targ}}|e^{-i\beta\hat{H}_x}|l\rangle|^2, \quad (8)$$

where $|\psi_{\text{targ}}\rangle$ is the $h = 0$ target ground state and the sum in the second line runs over the 2^N basis states $|l\rangle$ of the computational basis, with $\hat{H}_z|l\rangle = E_l|l\rangle$. Equation (8) shows that $\mathcal{F}$ is the scalar product of two 2^N-dimensional unit vectors of components

$$\mathcal{F} = |\mathbf{v}^\dagger \cdot \mathbf{u}|^2, \quad (9)$$

with

$$[\mathbf{v}(\gamma)]_l = \frac{1}{\sqrt{2^N}}e^{i\gamma E_l}, \quad [\mathbf{u}(\beta)]_l = \langle\psi_{\text{targ}}|e^{-i\beta\hat{H}_x}|l\rangle. \quad (10)$$

To ensure $\mathcal{F} = 1$, the Cauchy-Schwarz inequality requires $\mathbf{v}(\gamma)$ and $\mathbf{u}(\beta)$ to be parallel up to an overall phase factor. As discussed in Appendix A, this requires $\beta = \frac{\pi}{4}$. A unit fidelity further imposes [47] that all terms appearing in the sum in Eq. (2) are pure phase factors, which have to be identical for all l, modulo 2π. In Appendix A we perform the calculation explicitly, showing that the pair $(\beta = \frac{\pi}{4}, \gamma = \frac{\pi}{4})$ attains unit fidelity $\mathcal{F} = 1$ for odd p, while for even p the precise value of γ depends on p. As a remark, notice that in the theoretical proof we used, for convenience, the fidelity, rather than the residual energy. In the following numerics, however, we will prefer the residual energy as a figure of merit since it is directly linked to the variational minimization of the expectation value of the target Hamiltonian. Moreover, computing the fidelity requires full knowledge of the target ground state, which, in general, is not available for large systems. The variational energy, instead, is computed more easily, and it is accessible also in experimental implementation of QAOA [40], without performing full tomography of the variational state. Finally, when the residual energy vanishes, also perfect fidelity $\mathcal{F} = 1$ is attained, so both figures of merit show when the target state is prepared exactly.

The possibility of preparing exactly the GS with $P = 1$ is noteworthy, as it suggests that one can construct the exact $h = 0$ classical ground state with an algorithm whose equivalent computational time [see Eq. (7)] scales as N^{p-1}. On the contrary, for any finite N, a vanilla QA algorithm would need to cope with a minimum spectral gap at the transition point [20–22,42,48], which scales as $\Delta \sim N^{-1/3}$ if $p = 2$ and $\Delta \sim e^{-\alpha_p N}$ if $p \geqslant 3$: with a linear-schedule annealing, this implies a total annealing time $\tau \propto \Delta^{-2}$, and hence, $\tau \sim N^{2/3}$ for $p = 2$, and $\tau \sim e^{2\alpha_p N}$ for $p > 2$. Therefore, QAOA shows an *exponential speedup* with respect to a linear-schedule QA for $p > 2$ without exploiting any knowledge of the spectrum or of the phase diagram.

Such a remarkable property is, however, lost as soon as one targets a ground state with $h \neq 0$, where QAOA is no longer able to find the exact ground state with a small parameter number, $P = 1$ or 2. We find that the energy landscape $E_P(\boldsymbol{\gamma}, \boldsymbol{\beta})$ is extremely rugged for $P > 2$, making local optimizations—specifically, we use the Broyden-Fletcher-Goldfard-Shanno (BFGS) algorithm [49]—highly dependent on the initial set of parameters $(\boldsymbol{\gamma}^0, \boldsymbol{\beta}^0)$. We observe a very different behavior if
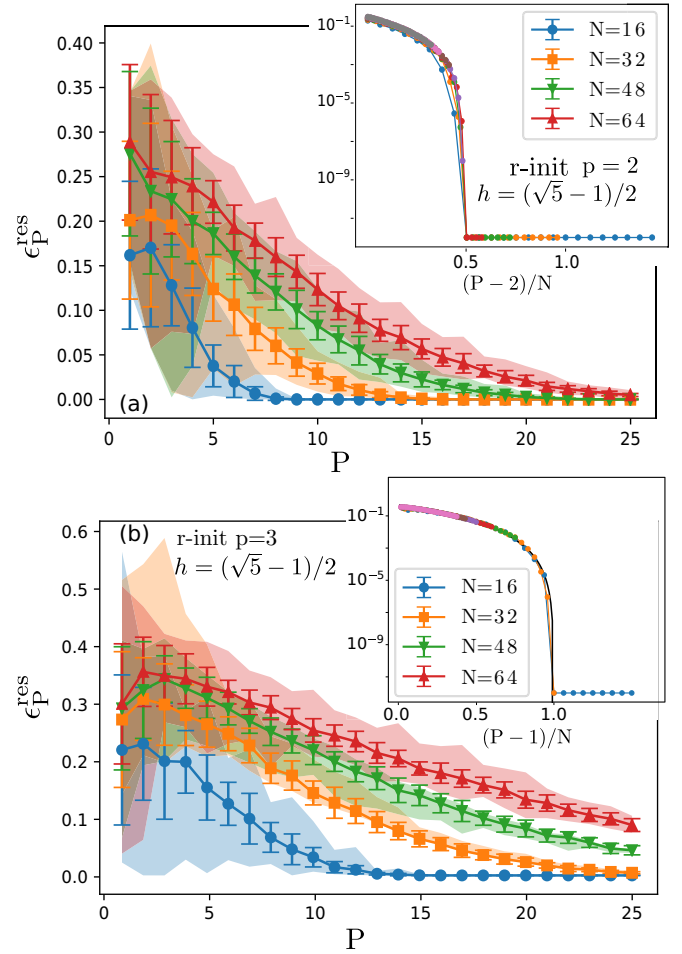

FIG. 1. (a) Results of local optimizations with $(\boldsymbol{\gamma}^0, \boldsymbol{\beta}^0)$ initialized randomly in $[0, \pi]$ (r-init) averaged over 100 different realizations for several values of N and $h = \frac{\sqrt{5}-1}{2}$ for $p = 2$. The shaded areas show the range between the best and worst results obtained for each set of data. The inset shows the collapsed data (in log scale) after rescaling $P \rightarrow (P - 2)/N$. (b) Same data for $p = 3$. The rescaling in the inset now is $P \rightarrow (P - 1)/N$, and the black line is the curve $(1 - P/P_N^*)^3$. Results for larger $p > 3$ are qualitatively similar.

the minimization is initialized with parameters γ_m^0 and β_m^0 chosen randomly in $[0, \pi]$ (r-init) or, rather, with an initial guess based on a linear schedule, $\gamma_m^0 = \frac{\Delta t}{\hbar}\frac{m}{P}\frac{1}{N^{p-1}}$ and $\beta_m^0 = \frac{\Delta t}{\hbar}[1 - \frac{m}{P}(1 - h)]$ (l-init). The results for the random initialization are summarized in Fig. 1, where we show the normalized residual energy [Eq. (5)] versus the number of QAOA steps P for $h = \frac{\sqrt{5}-1}{2} < h_c$, whose target state lies in the ferromagnetic phase for any value of p. Data for different system sizes N collapse perfectly after rescaling $P \rightarrow (P - 2)/N$ [see the inset in Fig. 1(a)] and drop below machine precision at $P = P_N^* = \frac{N}{2} + 2$. Correspondingly, the variance of the residual energy distribution, which is rather large for $P < P_N^*$, as witnessed by the error bars, drops to zero at P_N^*, implying that all local minima become degenerate. The colored shaded area around each curve shows the range between the lowest and highest residual energies obtained for each value of P and N. The distribution of individual optimizations is symmetric around the average, with the exception of small values of P where the
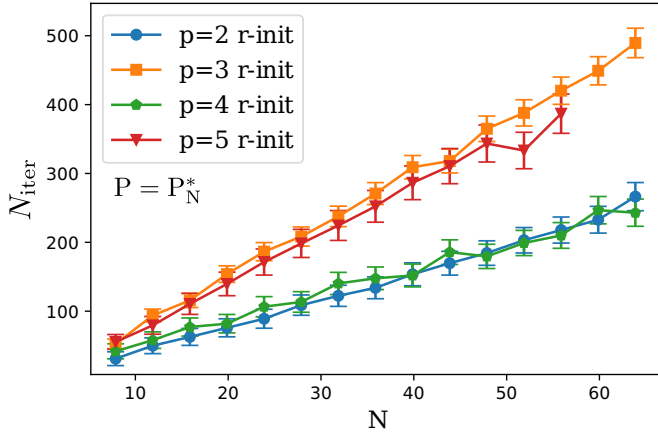
FIG. 2. Number of iterations required for the BFGS minimization algorithm to converge, averaged over 20 optimization runs with random initializations of QAOA parameters. The corresponding value of P is $N/2 + 2$ for even p and $N + 1$ for odd p, which is sufficient to obtain a residual energy below the numerical error.

r-init QAOA occasionally finds a local minimum with very small residual energy. For better readability, in the following figures we report only error bars corresponding to the standard deviation of our data.

The scaling shown in Fig. 1 holds for any value of the transverse field h if the QAOA minimization is initialized with random parameters. In general, we find that the residual energy follows

$$\epsilon_P^{\text{res}} = \begin{cases} \left(1 - \frac{P}{P_N^*}\right)^b & \text{if } P < P_N^*, \\ 0 & \text{if } P \geqslant P_N^*, \end{cases} \quad (11)$$

with $b \simeq 3$. Remarkably, this scaling holds also for $p > 2$, with similar values of b, with the only difference being that $P_N^* = N + 1$ for odd p because of the lack of the $\mathbb{Z}_2$ symmetry. This in turn suggests that for *finite N* one can attain perfect control of the state with $P = P_N^* \propto N$, physically corresponding to a total evolution time that scales as a power law with N. Our data for $p = 3$ are reported in Fig. 1(b), where, in the inset, we also highlight the curve described by Eq. (11) (solid black line). Once again, this is at variance with a standard linear-schedule QA, where the total evolution time has to scale *exponentially* with N when the transition is first order, i.e., for $p > 2$.

We have shown that a QAOA circuit with $P = P_N^* \propto N$ is sufficient to prepare the exact ground state of the p-spin model for an arbitrary target h. However, to estimate the total computational complexity of running the QAOA algorithm to solve the p-spin model, we must include the computational cost of finding the QAOA variational parameters. Indeed, during the optimization process, the quantum device is used N_{iter} times to sample the optimization landscape associated with QAOA circuits of $P = P_N^*$. In Fig. 2 we show the number of iterations N_{iter} required for convergence as a function of N. N_{iter} appears to increase linearly with N, with a slope that depends on only the parity of p. Hence, the total computational time needed for converging to the exact ground state, at an arbitrary transverse field, is, at most, polynomial in N, namely, $O(N^2)$, since it requires an order $O(N)$ of iterations
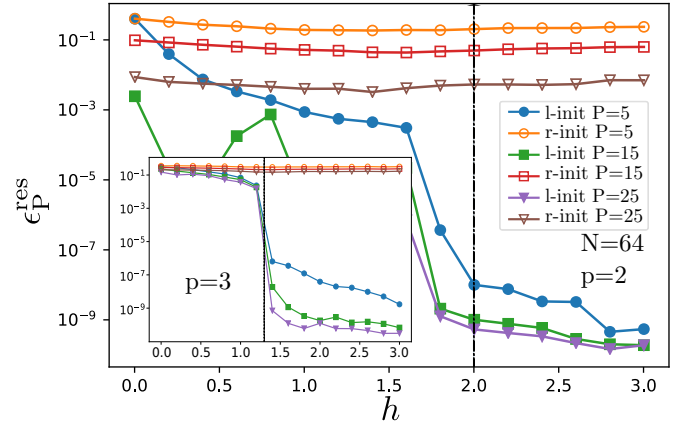

FIG. 3. QAOA residual energy versus the transverse field h for a system with $N = 64$ and $P = 5, 15, 25$ for random (r-init) and linear (l-init) initialization of the QAOA parameters. Notice that $P_N^* = N/2 + 2 = 34$ for $p = 2$ and $P_N^* = N + 1 = 65$ for $p = 3$. The vertical dashed lines show the critical transverse field: $h_c = 2$ for $p = 2$ (main plot) and $h_c \simeq 1.3$ for $p = 3$ (inset). The linear initialization (l-init) corresponds to Eq. (6) multiplied elementwise by a noise factor $(1 + r)$, with $r \in [-0.05, 0.05]^{2P}$ being a vector of uniformly distributed random numbers. Data are averaged over 100 different instances of r.

and a similar number of variational parameters, all in the range $[0, \pi]$. In an experimental setup, one has to take into account also the measurement process, which will increase the total computational cost of the QAOA optimization. As explained in Ref. [50], this extra cost is again polynomial in N, and for the p-spin model it can be estimated as

$$M \leqslant \frac{(1 + |h|)^2}{\epsilon_{\text{ms}}^2} N^2; \quad (12)$$

see Appendix C for details. Here M is the number of measurements required to estimate the energy $E_P(\boldsymbol{\gamma}, \boldsymbol{\beta})$ with an error ϵ_{ms}. The number of iterations required for the optimization then scales as $O(N^3)$, taking into account both Eq. (12) and the scaling shown in Fig. 2, leading to a total computational cost still polynomial in N.

Let us now focus on the dependence of the QAOA performance on the initial choice of the variational parameters. A linear initialization of QAOA parameters, with small noise (see the caption of Fig. 3 for details), improves drastically the QAOA performance. This is illustrated in Fig. 3, where the results of the two competing schemes, random (r-init) versus linear (l-init) initialization, are shown for a system with $N = 64$ for both $p = 2$ (main plot) and $p = 3$ (inset) and three fixed values of $P = 5, 15, 25$. Notice how the linear initialization is able to "detect" the quantum paramagnetic phase for $h > h_c$ as being "easy," with the QAOA minima found to have vanishingly small residual energy, almost to machine precision, even if $P < P_N^*$. This occurs not only in the second-order transition case with $p = 2$ but also in the more "difficult" first-order case with $p = 3$. At variance with that, a random initialization performs, on average, quite independently of the target transverse field h and knows nothing about the location of the critical field. Interestingly, this suggests that the QAOA for small P is sensitive to the phase diagram of the target
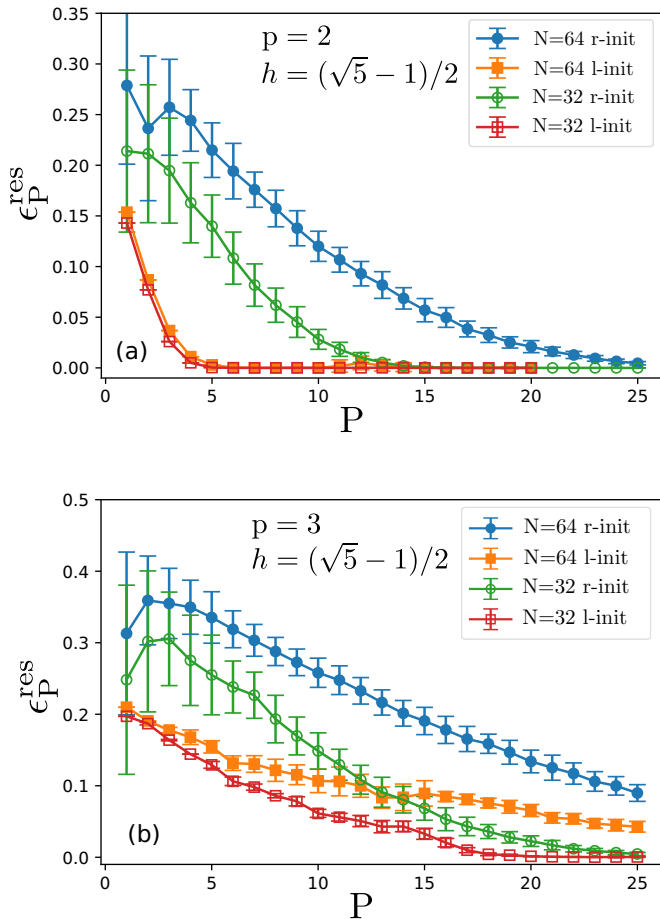
FIG. 4. Comparison between the optimized residual energy obtained from a linear initial guess plus small noise (l-init) and from random initialization (r-init) for two system sizes, $N = 32$ and $N = 64$. In (a) $p = 2$, and in (b) $p = 3$.

Hamiltonian: choosing a good ansatz for the initial parameter set $(\boldsymbol{\gamma}^0, \boldsymbol{\beta}^0)$ is fundamental to initialize the variational wave function in a good basin of attraction, where the minimization leads to small values of $\epsilon_P^{\mathrm{res}}$. Whether this feature is unique to infinite-range models or is a common property of long-range Hamiltonians is an interesting issue to pursue in future works.

The linear initialization also displays better efficiency, compared to the random one, when the target state belongs to the ferromagnetic phase ($h < h_c$) and $P < P_N^*$. This is illustrated in Fig. 4 for $p = 2$ [Fig. 4(a)] and $p = 3$ [Fig. 4(b)]. Here, however, the improvement is only quantitative—$\epsilon_P^{\mathrm{res}}$ decreases faster and scales better with system size—since the actual change in the landscape, with degenerate global minima, occurs only at P_N^*. Moreover, the system displays a large roughness of the variational energy landscape, which makes the task of obtaining good variational minima extremely demanding, especially for $p \geqslant 3$, hence justifying the poorer improvement of l-init over r-init observed in Fig. 4(b).

A smooth change in the control parameters is useful for experimental implementations of QAOA algorithms [40]. Besides allowing for easier control of external fields, they might also lead to faster convergence to a local minimum [37], hence reducing the number of measurements to be performed.

Finding local minima $(\boldsymbol{\gamma}^*, \boldsymbol{\beta}^*)$ which can be seen as the discretization of some continuous function proves, however, to be a difficult task for the present model. In contrast to Refs. [34,37], an iterative procedure that initializes $(\boldsymbol{\gamma}^0, \boldsymbol{\beta}^0)$ from an interpolation of a smooth set, obtained for a smaller parameter space, does not seem to work in a straightforward way. The linear initialization we have adopted is able to find reasonably smooth $(\boldsymbol{\gamma}^*, \boldsymbol{\beta}^*)$ only for small values of P. As the dimensionality of the parameter space increases and so does the roughness and the number of local minima, the optimal parameters obtained starting from a linear initialization scheme appear to be increasingly irregular (data not shown). Our failed attempts do not exclude the possibility that smart smooth choices for $(\boldsymbol{\gamma}^0, \boldsymbol{\beta}^0)$ can be constructed: they signal only that finding them is a nontrivial task due to the extreme roughness of the energy landscape. Recent results of reinforcement-learning-assisted QAOA show that smooth protocols can be constructed [51]; whether or not they describe the discretization of an adiabatic schedule is still an open question.

IV. CONCLUSIONS

We analyzed the performance of QAOA on the fully connected p-spin model, showing that it is able to find exactly the ferromagnetic ground state with polynomial resources, even when the system encounters a first-order phase transition. In particular, the algorithm prepares the ground state of $\widehat{H}_z$ with only $P = 1$ (if N is odd) or $P = 2$ (if N is even) steps, with a corresponding evolution time that scales as N^{p-1}, while a vanilla QA would require an exponentially long annealing time. This exact minimum, however, exists only for zero transverse field, $h = 0$. Interestingly, the exact minimum, which clearly survives for $P \geqslant 2$, is very hard to find with gradient-based optimization schemes due to the extreme roughness of the energy landscape, especially for $p > 2$. The "hardness" of the problem for $p > 2$ is thus reflected in the difficulty in finding the correct absolute minimum, rather than in the resources (i.e., the computational time) needed.

The performance of the optimization itself strongly depends on the initialization of the variational parameters $(\boldsymbol{\gamma}^0, \boldsymbol{\beta}^0)$. For a random initialization, the residual energy drops below machine precision as $(P_N^* - P)^b$, with $b \sim 3$ and P_N^* growing linearly with N. This behavior is independent of the target transverse field h and from p, with the only difference being that $P_N^* = N/2 + 2$ for even p and $P_N^* = N + 1$ for odd p. With a linear initialization, the algorithm performs much better and is able to detect the presence of a phase transition, although the improvement deteriorates rapidly as P increases because of the growing number of "bad" local minima.

For future developments, it would be interesting to understand whether infinite- or long-range Hamiltonians can be used to boost QAOA performance on short-range models. The idea is to add a further unitary $e^{-i\epsilon \hat{H}_z'}$ in Eq. (2), generated by a long-range Hamiltonian $\hat{H}_z'$, unrelated to the problem to be solved. This enlarges the portion of Hilbert space approximated with a QAOA ansatz for a fixed number of Trotter steps P at the price, however, of increasing the number of variational parameters.

At variance with Refs. [34,37,40], we are unable to construct minima in the energy landscape associated with smooth parameters $(\boldsymbol{\gamma}^*, \boldsymbol{\beta}^*)$. These regular parameter choices are often desired since they might be linked to adiabatic schedules which would improve the protocol $s(t)$ in a continuous-time QA framework and allow, within QAOA, for a faster minimum search in the $2P$-dimensional parameter space once a solution for $P' < P$ is known [37,40]. Preliminary results [51] with reinforcement-learning [52] methods applied to the QAOA evolution suggest that smooth choices of $(\boldsymbol{\gamma}^*, \boldsymbol{\beta}^*)$ do indeed exist, but they are hard to find with local optimizations. Whether or not *global* minima are related to smooth values of $(\boldsymbol{\gamma}^*, \boldsymbol{\beta}^*)$ remains an open and interesting question.

ACKNOWLEDGMENTS

APPENDIX A: EXACT GROUND-STATE PREPARATION FOR $P = 1$

In this Appendix we will show that one can get the exact target ground state of the p-spin model with a single QAOA step, $P = 1$, starting from the fully x-polarized state $|+\rangle = \frac{1}{\sqrt{2^N}} \bigotimes_{j=1}^{N} (|\uparrow\rangle_j + |\downarrow\rangle_j)$, provided the number of sites N is *odd*. This holds true for all possible values of p and generalizes the result of Ref. [45] to $p > 2$.

1. $P = 1$: Requirements for β

For $P = 1$ the QAOA state has only two parameters, which we will denote by γ and β, without an index. Let $|\psi_{\text{targ}}\rangle$ denote the (target) ground state of the model, and let us define the fidelity

$$
\begin{aligned}
\mathcal{F}(\gamma, \beta) &= |\langle \psi_{\text{targ}} | \psi_{P=1}(\gamma, \beta)\rangle|^2 \\
&= |\langle \psi_{\text{targ}} | e^{-i\beta \hat{H}_x} e^{-i\gamma \hat{H}_z} |+\rangle|^2 \\
&= \left| \frac{1}{\sqrt{2^N}} \sum_l e^{-i\gamma E_l} \langle \psi_{\text{targ}} | e^{-i\beta \hat{H}_x} |l\rangle \right|^2, \quad \text{(A1)}
\end{aligned}
$$

where we have expanded the initial state $|+\rangle = \frac{1}{\sqrt{2^N}} \sum_l |l\rangle$ as an equal superposition of all possible 2^N classical $\mathbf{z}$-basis configurations $|l\rangle$ and we used $\hat{H}_z |l\rangle = E_l |l\rangle$, where E_l is the energy of the configuration $|l\rangle$. Let us now define the following two 2^N-dimensional complex vectors:

$$
[\mathbf{v}(\gamma)]_l = \frac{1}{\sqrt{2^N}} e^{i\gamma E_l}, \qquad [\mathbf{u}(\beta)]_l = \langle \psi_{\text{targ}} | e^{-i\beta \hat{H}_x} |l\rangle. \quad \text{(A2)}
$$

Simple algebra shows that they have unit norm, $\|\mathbf{v}(\gamma)\| = 1$ and $\|\mathbf{u}(\beta)\| = 1$, and that the fidelity can be expressed as a scalar product of them: $\mathcal{F}(\gamma, \beta) = |\mathbf{v}^\dagger(\gamma) \cdot \mathbf{u}(\beta)|^2$. Hence, by the Cauchy-Schwarz inequality,

$$
1 = \mathcal{F}(\gamma, \beta) = |\mathbf{v}^\dagger(\gamma) \cdot \mathbf{u}(\beta)|^2 \iff \exists \theta \in \mathbb{R},
$$

$$
\text{such that} \quad \mathbf{u}(\beta) = e^{i\theta} \mathbf{v}(\gamma), \quad \text{(A3)}
$$

i.e., the two vectors coincide, up to an overall phase factor. Since $|[\mathbf{v}(\gamma)]_l|^2 = \frac{1}{2^N}$, this in turn implies that we must have

$$
|\langle \psi_{\text{targ}} | e^{-i\beta \hat{H}_x} |l\rangle|^2 = |[\mathbf{u}(\beta)]_l|^2 = \frac{1}{2^N} \quad \forall l. \quad \text{(A4)}
$$

So far, our arguments have been rather general. We now specialize our discussion to the case where $|\psi_{\text{targ}}\rangle$ is the ground state of the classical p-spin ferromagnet.

For odd p, we have $|\psi_{\text{targ}}\rangle = |\uparrow \cdots \uparrow\rangle$, and a simple calculation shows that

$$
\begin{aligned}
\langle \psi_{\text{targ}} | e^{-i\beta \hat{H}_x} |l\rangle &= \prod_{j=1}^{N} \langle \uparrow | \cos \beta \, \hat{1}_j + i \sin \beta \, \hat{\sigma}_j^x |l_j\rangle \\
&= (\cos \beta)^{N_l^\uparrow} (i \sin \beta)^{N_l^\downarrow}, \quad \text{(A5)}
\end{aligned}
$$

where $N_l^\uparrow$ and $N_l^\downarrow$ denote the number of $\uparrow$ and $\downarrow$ spins in the configuration $|l\rangle$. Hence, the requirement given by Eq. (A4) is satisfied only if

$$
\cos^2 \beta = \sin^2 \beta = \frac{1}{2} \implies \beta = \frac{\pi}{4}, \frac{3\pi}{4}, \frac{5\pi}{4}, \frac{7\pi}{4}. \quad \text{(A6)}
$$

Similar arguments have been used (see Ref. [47]) for the more general case in which $|\psi_{\text{targ}}\rangle$ is the classical ground state of a generic spin-glass Hamiltonian.

For even p the calculation is slightly more involved since the target state is now a nonclassical superposition of the two degenerate ferromagnetic states,

$$
|\psi_{\text{targ}}\rangle = \frac{1}{\sqrt{2}} (|\uparrow \cdots \uparrow\rangle + |\downarrow \cdots \downarrow\rangle). \quad \text{(A7)}
$$

Hence,

$$
\begin{aligned}
|\langle \psi_{\text{targ}} | e^{-i\beta \hat{H}_x} |l\rangle|^2 = \frac{1}{2} |(\cos \beta)^{N_l^\uparrow} (i \sin \beta)^{N_l^\downarrow} \\
+ (\cos \beta)^{N_l^\downarrow} (i \sin \beta)^{N_l^\uparrow} |^2. \quad \text{(A8)}
\end{aligned}
$$

Once again, one easily verifies that $\beta = \frac{\pi}{4}$ satisfies the requirement (A4), *provided N is odd*, so that $N_l^\uparrow$ and $N_l^\downarrow$ have opposite parities, and therefore, $|i^{N_l^\uparrow} + i^{N_l^\downarrow}|^2 = 2$.

From now on we will therefore restrict our choice of β to $\beta = \frac{\pi}{4}$, a necessary condition for unit fidelity, and study the conditions that γ has to verify. Essentially, the value of γ will have to be chosen in such a way that the various phase factors interfere constructively in a way that is independent of l. For this goal, we notice that the energy E_l of the configuration $|l\rangle$ can be expressed as

$$
E_l = -\langle l | \left(\sum_j \hat{\sigma}_j^z \right)^p |l\rangle = -(N_l^\uparrow - N_l^\downarrow)^p = -M_l^p, \quad \text{(A9)}
$$

where $M_l = N_l^\uparrow - N_l^\downarrow$ is the total magnetization of the configuration.

2. $P = 1$ and odd p: Requirements for γ

Here we prove that for odd values of p and N, the p-spin QAOA circuit of depth $P = 1$ and parameters $(\gamma = \frac{\pi}{4}, \beta = \frac{\pi}{4})$ is sufficient to prepare the ferromagnetic target state $|\psi_{\text{targ}}\rangle = |\uparrow \cdots \uparrow\rangle$. Substituting $\beta = \frac{\pi}{4}$ and $i => i^{\frac{\pi}{2}}$ in

Eq. (A5) and using $E_l = -M_l^p$ in Eq. (A1), we get

$$\mathcal{F}\left(\gamma, \frac{\pi}{4}\right) = \left| \frac{1}{2^N} \sum_l e^{i\gamma M_l^p} e^{i\frac{\pi}{2}N_l^\downarrow} \right|^2. \quad \text{(A10)}$$

Taking $\gamma = \frac{\pi}{4}$ gives

$$\mathcal{F}\left(\frac{\pi}{4}, \frac{\pi}{4}\right) = \left| \frac{1}{2^N} \sum_l \exp\left\{ i\frac{2\pi}{8} \left[(M_l^p + 2N_l^\downarrow) \mod 8 \right] \right\} \right|^2$$

$$= \left| \frac{1}{2^N} \sum_l \exp\left\{ i\frac{2\pi}{8} [(M_l + 2N_l^\downarrow) \mod 8] \right\} \right|^2$$

$$= \left| \frac{1}{2^N} \sum_l \exp\left\{ i\frac{2\pi}{8} [N \mod 8] \right\} \right|^2 = 1, \quad \text{(A11)}$$

where we have used the fact that for odd N, $M_l = N - 2N_l^\downarrow$ is also odd and the following property of arithmetic congruences holds:

$$M_l^{p-1} = 1 \mod 8 \implies M_l^p = M_l \mod 8 \quad \text{if } p \text{ is odd.} \quad \text{(A12)}$$

Equation (A11) proves our initial claim that the QAOA protocol ($\gamma = \frac{\pi}{4}, \beta = \frac{\pi}{4}$) prepares the target ground state of $\hat{H}_z$ for the p-spin ferromagnet with unit fidelity, provided N and p are both odd.

3. $P = 1$ and even p: Requirements for γ

For even values of p, the system is $\mathbb{Z}_2$ symmetric. The p-spin QAOA circuit preserves such symmetry. Therefore, the targeted ground state of $\hat{H}_z$ is $|\psi_{\text{targ}}\rangle$ in Eq. (A7). As we did in the previous section, we compute the fidelity between the output $|\psi_{P=1}(\gamma, \beta = \frac{\pi}{4})\rangle$ of the QAOA circuit and the (nonclassical) target state $|\psi_{\text{targ}}\rangle$ in Eq. (A7):

$$\mathcal{F}\left(\gamma, \beta = \frac{\pi}{4}\right) = \left| \frac{1}{2^N} \sum_l e^{i\gamma M_l^p} \left(\frac{e^{i\frac{\pi}{2}N_l^\downarrow} + e^{i\frac{\pi}{2}N_l^\uparrow}}{\sqrt{2}} \right) \right|^2. \quad \text{(A13)}$$

We observe that for odd $N = N_l^\uparrow + N_l^\downarrow$, $N_l^\uparrow$ and $N_l^\downarrow$ must have *opposite parities*, and the term inside the parentheses is a pure phase factor, which can be expressed as

$$\left(\frac{e^{i\frac{\pi}{2}N_l^\downarrow} + e^{i\frac{\pi}{2}N_l^\uparrow}}{\sqrt{2}} \right) = e^{i\frac{\pi}{4}N} e^{-i\pi f(M_l)}, \quad \text{(A14)}$$

where

$$f(M) = \begin{cases} 0 & \text{for } M \mod 8 = \pm 1, \\ 1 & \text{for } M \mod 8 = \pm 3. \end{cases} \quad \text{(A15)}$$

Hence, omitting the irrelevant l-independent common factor $e^{i\frac{\pi}{4}N}$, we can rewrite the fidelity as

$$\mathcal{F}\left(\gamma, \frac{\pi}{4}\right) = \left| \frac{1}{2^N} \sum_l e^{i[\gamma M_l^p - \pi f(M_l)]} \right|^2. \quad \text{(A16)}$$

The arithmetic to prove that the various phase factors can be made l independent for a judicious choice of γ is now, for even p, slightly more involved. By experimenting with

this expression for $p \leqslant 10$, we come out with the following unconventional parametrization of an even value of p: for every even p, two natural numbers, n and k, can be found such that

$$p = 2^{k+1} + n2^k. \quad \text{(A17)}$$

Correspondingly, given the value of k in Eq. (A17), we will set the value of γ to

$$\gamma_k = \frac{2\pi}{2^{k+4}}. \quad \text{(A18)}$$

The crucial arithmetic identity which we will use (see Sec. A 3 for a proof) is the following:

$$m^{2^{k+1}+n2^k} \mod 2^{k+4} = f(m) 2^{k+3} + 1 \quad \forall m \in \mathbb{Z} \text{ with odd } m, \quad \text{(A19)}$$

where $f(m)$ is the function given in Eq. (A15).

With these definitions, it is immediate possible to verify that

$$\mathcal{F}\left(\gamma_k, \frac{\pi}{4}\right)$$

$$= \left| \frac{1}{2^N} \sum_l e^{-i\pi f(M_l)} \exp\left[i\frac{2\pi}{2^{k+4}} (M_l)^{2^{k+1}+n2^k} \right] \right|^2$$

$$= \left| \frac{1}{2^N} \sum_l e^{-i\pi f(M_l)} \exp\left\{ i\frac{2\pi}{2^{k+4}} [(M_l)^{2^{k+1}+n2^k} \mod 2^{k+4}] \right\} \right|^2$$

$$= \left| \frac{1}{2^N} \sum_l e^{-i\pi f(M_l)} \exp\left\{ i\frac{2\pi}{2^{k+4}} [f(M_l) 2^{k+3} + 1] \right\} \right|^2$$

$$= \left| \frac{1}{2^N} \sum_l e^{-i\pi f(M_l)} e^{i\pi f(M_l) + i\frac{2\pi}{2^{k+4}}} \right|^2 = 1. \quad \text{(A20)}$$

Proof of identity in Eq. (A19). For completeness, we also present a proof of the arithmetic identity equation (A19). To prove Eq. (A19), it is sufficient to show that

$$\forall k \in \mathbb{N}, \ m \in \mathbb{Z}, \ m \text{ odd}, \quad \left(m^{2^{k+1}} - 1\right) \mod 2^{k+4}$$

$$= f(m) 2^{k+3}. \quad \text{(A21)}$$

We prove Eq. (A21) by induction over k:

(i) We show that Eq. (A21) holds for $k = 0$.
For $k = 0$, a direct computation, for odd m, gives

$$\left(m^{2^{0+1}} - 1\right) \mod 2^{0+4} = (m^2 - 1) \mod 16$$

$$= (m - 1)(m + 1) \mod 16$$

$$= f(m) 2^3. \quad \text{(A22)}$$

(ii) We show that if Eq. (A21) holds for a given $k \in \mathbb{N}$ and for all odd $m \in \mathbb{N}$, then it holds also for $k + 1$.

Using Eq. (A21), we write

$$m^{2^{k+1}} = a_m 2^{k+4} + f(m) 2^{k+3} + 1, \quad \text{(A23)}$$

with $a_m \in \mathbb{Z}$. Then, we have

$$
\begin{aligned}
\left(m^{2^{(k+1)+1}} - 1\right) & \\
&= \left(m^{2^{k+1}} - 1\right)\left(m^{2^{k+1}} + 1\right) \\
&= [a_m 2^{k+4} + f(m) 2^{k+3}][a_m 2^{k+4} + f(m) 2^{k+3} + 2] \\
&= [a_m 2^{k+5} + f(m) 2^{k+4}][a_m 2^{k+3} + f(m) 2^{k+2} + 1].
\end{aligned}
\tag{A24}
$$

From this, we derive

$$
\begin{aligned}
\left(m^{2^{(k+1)+1}} - 1\right) \quad &\mathrm{mod}\ 2^{(k+1)+4} \\
&= f(m) 2^{k+4}[a_m 2^{k+3} + f(m) 2^{k+2} + 1] \quad \mathrm{mod}\ 2^{k+5}
\end{aligned}
\tag{A25}
$$

$$
= f(m) 2^{k+4},
\tag{A26}
$$

where we have used that $f(m) = 0, 1$ for all odd $m \in \mathbb{Z}$. This indeed implies that for all $k \in \mathbb{N}$,

$$
\begin{aligned}
\left(m^{2^{k+1}} - 1\right) \quad &\mathrm{mod}\ 2^{k+4} = f(m) 2^{k+3} \Rightarrow \left(m^{2^{(k+1)+1}} - 1\right) \\
&\mathrm{mod}\ 2^{(k+1)+4} = f(m) 2^{(k+1)+3}.
\end{aligned}
\tag{A27}
$$

This concludes the proof by induction of Eq. (A21).

Incidentally, as an immediate consequence of Eq. (A21) we get that, for any $n \in \mathbb{N}$,

$$
m^{2^{k+1}} \qquad \mathrm{mod}\ 2^{k+4} = f(m) 2^{k+3} + 1, \tag{A28}
$$

$$
m^{2^{k+2}} \qquad \mathrm{mod}\ 2^{k+4} = 1, \tag{A29}
$$

$$
m^{2^{k+1}+n2^k} \quad \mathrm{mod}\ 2^{k+4} = f(m) 2^{k+3} + 1. \tag{A30}
$$

Notice that Eq. (A29) also follows from the properties of the multiplicative group of integers modulo 2^k discussed in Refs. [53,54] [e.g., $(\mathbb{Z}/2^{k+4}\mathbb{Z})^\times \cong C_2 \times C_{2^{k+2}}$].

APPENDIX B: SYMMETRIES OF THE PARAMETER SPACE FOR GENERAL P, N, AND p

We discuss here the symmetries in the parameter space of Eq. (4), which we recall here for convenience,

$$
E_P(\boldsymbol{\gamma}, \boldsymbol{\beta}) = \langle \psi_P(\boldsymbol{\gamma}, \boldsymbol{\beta}) | \widehat{H}_{\mathrm{target}} | \psi_P(\boldsymbol{\gamma}, \boldsymbol{\beta}) \rangle. \tag{B1}
$$

A first trivial operation that leaves the energy unaltered is the inversion $(\boldsymbol{\gamma}, \boldsymbol{\beta}) \rightarrow -(\boldsymbol{\gamma}, \boldsymbol{\beta})$, which corresponds to the complex conjugate of Eq. (4). Indeed, it is immediately possible to see that

$$
|\psi_P(-\boldsymbol{\gamma}, -\boldsymbol{\beta})\rangle = \prod_{m=1}^{P} e^{i\beta_m \widehat{H}_x} e^{i\gamma_m \widehat{H}_z} |\psi_0\rangle = |\psi_P(\boldsymbol{\gamma}, \boldsymbol{\beta})\rangle^*, \tag{B2}
$$

given that $|\psi_0\rangle = |+\rangle$ is a real wave function in the basis of $\hat{S}_z$.

The symmetries on the $\boldsymbol{\beta}$ parameters are shared by all QAOA wave functions where quantum fluctuations are induced by a magnetic field transverse to the computational basis. We can write a single evolution operator $e^{-i\beta_m \widehat{H}_x}$ as a set of rotations on each individual spin,

$$
e^{-i\beta_m \widehat{H}_x} = e^{i\beta_m \sum_{j=1}^{N} \hat{\sigma}_j^x} = \bigotimes_{j=1}^{N} \left(\cos\beta_m + \hat{\sigma}_j^x \sin\beta_m \right). \tag{B3}
$$

TABLE I. Symmetry operations for the QAOA process of the p-spin model. It is understood that any component of $\boldsymbol{\gamma}$ or $\boldsymbol{\beta}$ can be modified.

	Symmetry operation
$\forall N,\ p$	$E(-\boldsymbol{\gamma}, -\boldsymbol{\beta}) = E(\boldsymbol{\gamma}, \boldsymbol{\beta})$
p odd	$E(\boldsymbol{\gamma}, \boldsymbol{\beta} + \pi) = E(\boldsymbol{\gamma}, \boldsymbol{\beta})$
p even	$E(\boldsymbol{\gamma}, \boldsymbol{\beta} + \frac{\pi}{2}) = E(\boldsymbol{\gamma}, \boldsymbol{\beta})$
N odd	$E(\boldsymbol{\gamma} + \pi, \boldsymbol{\beta}) = E(\boldsymbol{\gamma}, \boldsymbol{\beta})$
N even	$E(\boldsymbol{\gamma} + \frac{\pi}{2^{p-1}}, \boldsymbol{\beta}) = E(\boldsymbol{\gamma}, \boldsymbol{\beta})$

A shift $\beta_m \rightarrow \beta_m + \pi$ changes the sign of each term in the product, leading to

$$
\begin{aligned}
e^{-i(\beta_m+\pi)\widehat{H}_x} &= \bigotimes_{j=1}^{N} \left(-\cos\beta_m - \hat{\sigma}_j^x \sin\beta_m \right) \\
&= (-1)^N \bigotimes_{j=1}^{N} \left(\cos\beta_m + \hat{\sigma}_j^x \sin\beta_m \right),
\end{aligned}
\tag{B4}
$$

which is a trivial global phase that does not change the energy in Eq. (4). Moreover, if p is even, the target Hamiltonian is $\mathbb{Z}_2$ symmetric. Recall that $\widehat{H}_x = -\hat{S}_x$ (twice the total spin), which implies that

$$
e^{i\frac{\pi}{2}\hat{S}_x} \widehat{H}_{\mathrm{target}} e^{-i\frac{\pi}{2}\hat{S}_x} = \widehat{H}_{\mathrm{target}} \tag{B5}
$$

because $e^{-i\frac{\pi}{2}\hat{S}_x}$ is a π rotation around the **x** direction, which gives a global spin flip $\hat{\sigma}_j^z \rightarrow -\hat{\sigma}_j^z$, leading to $E_P(\boldsymbol{\gamma}, \boldsymbol{\beta} + \frac{\pi}{2}) = E_P(\boldsymbol{\gamma}, \boldsymbol{\beta})$.

The symmetry for $\boldsymbol{\gamma}$ is subtler and is model specific. Notice first that $\hat{S}_z = \sum_j \hat{\sigma}_j^z$ has integer eigenvalues, even or odd depending on N, and so does $\widehat{H}_z = -\hat{S}_z^p$. Following the same notation introduced previously, we write a single QAOA evolution operator as

$$
e^{-i\gamma_m \hat{H}_z} = \sum_l e^{i\gamma_m M_l^p} |l\rangle\langle l|. \tag{B6}
$$

If N is odd, the eigenvalues M_l^p of $\hat{S}_z^p$ are also odd, and the periodicity of γ_m is π because

$$
(\gamma_m + \pi)M_l^p = \gamma_m M_l^p + \pi \quad \mathrm{mod}\ 2\pi. \tag{B7}
$$

Hence, the shift $\gamma_m \rightarrow \gamma_m + \pi$ introduces a global phase $e^{-i(\gamma_m+\pi)\hat{H}_z} = -e^{-i\gamma_m \hat{H}_z}$, which is irrelevant in the expectation value of the energy. If N is even, the eigenvalues M_l^p of $\hat{S}_z^p$ are multiples of 2^p; hence,

$$
\left(\gamma_m + \frac{\pi}{2^{p-1}}\right)M_l^p = \gamma_m M_l^p \quad \mathrm{mod}\ 2\pi, \tag{B8}
$$

which means that $e^{i(\gamma_m+\frac{\pi}{2^{p-1}})\hat{S}_z^p} = e^{i\gamma_m \hat{S}_z^p}$. In Table I we summarize the symmetries we have discussed.

APPENDIX C: ACCOUNTING FOR THE MEASUREMENT PROCESS

Within the QAOA scheme, the expectation value $E_P(\boldsymbol{\gamma}, \boldsymbol{\beta})$ is usually estimated by repeated measurements. In this Appendix we take into account the cost of such measurements

and discuss how this changes the scaling of the computational time needed to solve the p-spin model.

To keep the analysis straightforward, we consider a simple but suboptimal measurement strategy [50]. We first decompose the Hamiltonian into a sum of Pauli strings; then we estimate the expectation value of each Pauli string independently by repeated measurements. In this scheme, the total number of measurements M needed to estimate $E_P(\boldsymbol{\gamma}, \boldsymbol{\beta})$ up to a target precision ϵ_{ms} is bounded by [50]

$$M \leqslant \left(\frac{\sum_l |\omega_l|}{\epsilon_{\mathrm{ms}}} \right)^2, \tag{C1}$$

where ω_l are the coefficients of the Pauli strings appearing in the Hamiltonian. For the p-spin target Hamiltonian [see Eq. (1)], the bound reads

$$M \leqslant \frac{(1 + |h|)^2}{\epsilon_{\mathrm{ms}}^2} N^2. \tag{C2}$$

Running the QAOA with $P = P^*$, we can find the solution up to precision ϵ, with a total number of measurements $M N_{\mathrm{iter}}$. The computational cost associated with the measurements is therefore $O(\frac{N^3}{\epsilon})$. The total computational cost of the QAOA algorithm hence becomes

$$t_{\mathrm{cc}} \sim N_{\mathrm{iter}} \big[c_{\mathrm{QAOA-1}} O(P^*) + c_{\mathrm{ms-1}} O(N^2/\epsilon_{\mathrm{ms}}^2) \big]$$
$$= c_{\mathrm{QAOA1}} O(N^2) + c_{\mathrm{ms1}} O(N^3/\epsilon_{\mathrm{ms}}^2). \tag{C3}$$

Here c_{QAOA1} and c_{ms1} are, respectively, the cost of performing one QAOA step on the quantum device and measuring all the qubits ones. We conclude that, even after including the cost of measurements, the computational cost of finding an approximate ground state of the p-spin Hamiltonian with QAOA remains polynomial in the system size.

[1] A. Acín *et al.*, New J. Phys. **20**, 080201 (2018).
[2] G. Kurizki, P. Bertet, Y. Kubo, K. Mølmer, D. Petrosyan, P. Rabl, and J. Schmiedmayer, Proc. Natl. Acad. Sci. USA **112**, 3866 (2015).
[3] J. Preskill, arXiv:1203.5813.
[4] E. Farhi and A. W. Harrow, arXiv:1602.07674.
[5] F. Arute *et al.*, Nature (London) **574**, 505 (2019).
[6] S. Lloyd, Science **273**, 1073 (1996).
[7] I. Buluta and F. Nori, Science **326**, 108 (2009).
[8] I. M. Georgescu, S. Ashhab, and F. Nori, Rev. Mod. Phys. **86**, 153 (2014).
[9] A. Lucas, Front. Phys. **2**, 5 (2014).
[10] A. B. Finnila, M. A. Gomez, C. Sebenik, C. Stenson, and J. D. Doll, Chem. Phys. Lett. **219**, 343 (1994).
[11] T. Kadowaki and H. Nishimori, Phys. Rev. E **58**, 5355 (1998).
[12] J. Brooke, D. Bitko, T. F. Rosenbaum, and G. Aeppli, Science **284**, 779 (1999).
[13] G. E. Santoro, R. Martoňák, E. Tosatti, and R. Car, Science **295**, 2427 (2002).
[14] E. Farhi, J. Goldstone, S. Gutmann, J. Lapan, A. Lundgren, and D. Preda, Science **292**, 472 (2001).
[15] T. Albash and D. A. Lidar, Rev. Mod. Phys. **90**, 015002 (2018).
[16] B. Heim, T. F. Rønnow, S. V. Isakov, and M. Troyer, Science **348**, 215 (2015).
[17] E. Crosson and A. W. Harrow, in *2016 IEEE 57th Annual Symposium on Foundations of Computer Science (FOCS)* (IEEE, Piscataway, NJ, 2016), pp. 714–723.
[18] V. S. Denchev, S. Boixo, S. V. Isakov, N. Ding, R. Babbush, V. Smelyanskiy, J. Martinis, and H. Neven, Phys. Rev. X **6**, 031015 (2016).
[19] T. Albash and D. A. Lidar, Phys. Rev. X **8**, 031016 (2018).
[20] T. Jörg, F. Krzakala, J. Kurchan, A. C. Maggs, and J. Pujos, Europhys. Lett. **89**, 40004 (2010).
[21] V. Bapst and G. Semerjian, J. Stat. Mech. (2012) P06007.
[22] M. M. Wauters, R. Fazio, H. Nishimori, and G. E. Santoro, Phys. Rev. A **96**, 022326 (2017).
[23] H. Nishimori and K. Takada, Front. ICT **4**, 2 (2017).

[24] Y. Susa, Y. Yamashiro, M. Yamamoto, and H. Nishimori, J. Phys. Soc. Jpn. **87**, 023002 (2018).
[25] G. Passarelli, V. Cataudella, and P. Lucignano, Phys. Rev. B **100**, 024302 (2019).
[26] G. Passarelli, G. De Filippis, V. Cataudella, and P. Lucignano, Phys. Rev. A **97**, 022319 (2018).
[27] G. Passarelli, K.-W. Yip, D. A. Lidar, H. Nishimori, and P. Lucignano, Phys. Rev. A **101**, 022331 (2020).
[28] G. Passarelli, V. Cataudella, R. Fazio, and P. Lucignano, Phys. Rev. Res. **2**, 013283 (2020).
[29] A. Peruzzo, J. McClean, P. Shadbolt, M.-H. Yung, X.-Q. Zhou, P. J. Love, A. Aspuru-Guzik, and J. L. O'brien, Nat. Commun. **5**, 4213 (2014).
[30] E. Farhi, J. Goldstone, and S. Gutmann, arXiv:1411.4028.
[31] C. Kokail *et al.*, Nature (London) **569**, 355 (2019).
[32] J. R. McClean, J. Romero, R. Babbush, and A. Aspuru-Guzik, New J. Phys. **18**, 023023 (2016).
[33] M. Y. Niu, S. Lu, and I. L. Chuang, arXiv:1905.12134.
[34] L. Zhou, S.-T. Wang, S. Choi, H. Pichler, and M. D. Lukin, Phys. Rev. X **10**, 021067 (2020).
[35] S. Lloyd, arXiv:1812.11075.
[36] Z. Wang, S. Hadfield, Z. Jiang, and E. G. Rieffel, Phys. Rev. A **97**, 022304 (2018).
[37] G. B. Mbeng, R. Fazio, and G. E. Santoro, arXiv:1906.08948.
[38] Z. Jiang, E. G. Rieffel, and Z. Wang, Phys. Rev. A **95**, 062317 (2017).
[39] G. B. Mbeng, R. Fazio, and G. E. Santoro, arXiv:1911.12259.
[40] G. Pagano *et al.*, Proc. Natl. Acad. Sci. **117**, 25396 (2020).
[41] M. Filippone, S. Dusuel, and J. Vidal, Phys. Rev. A **83**, 022327 (2011).
[42] T. Caneva, R. Fazio, and G. E. Santoro, Phys. Rev. B **76**, 144427 (2007).
[43] R. Barends *et al.*, Nature (London) **534**, 222 (2016).
[44] G. B. Mbeng, L. Arceci, and G. E. Santoro, Phys. Rev. B **100**, 224201 (2019).
[45] W. W. Ho, C. Jonay, and T. H. Hsieh, Phys. Rev. A **99**, 052332 (2019).

[46] H. Lipkin, N. Meshkov, and A. Glick, Nucl. Phys. **62**, 188 (1965).

[47] M. Streif and M. Leib, arXiv:1901.01903.

[48] N. Defenu, T. Enss, M. Kastner, and G. Morigi, Phys. Rev. Lett. **121**, 240403 (2018).

[49] J. Nocedal and S. Wright, *Numerical Optimization* (Springer, 2006).

[50] W. J. Huggins, J. McClean, N. Rubin, Z. Jiang, N. Wiebe, K. B. Whaley, and R. Babbush, arXiv:1907.13117.

[51] M. M. Wauters, E. Panizon, G. B. Mbeng, and G. E. Santoro, Phys. Rev. Res. **2**, 033446 (2020).

[52] R. S. Sutton and A. G. Barto, *Reinforcement Learning: An Introduction*, 2nd ed. (MIT Press, Cambrdige, MA, 2018).

[53] J. C. F. Gauss, *Disquisitiones arithmeticae*, Yale Paperbound Vol. 157 (Yale University Press, New Haven, CT, 1966).

[54] Wikipedia, Multiplicative group of integers modulo n, https://en.wikipedia.org/wiki/Multiplicative_group_of_integers_modulo_n.