Optimizing QAOA: Success Probability and Runtime Dependence on Circuit Depth

Murphy Yuezhen Niu,[1, 2, 3, *] Sirui Lu,[4] and Isaac L. Chuang[1, 2]

[1] *Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, Massachusetts, 02139, USA*
[2] *Department of Physics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA*
[3] *Google Inc., 340 Main Street, Venice, CA 90291*
[4] *Department of Physics, Tsinghua University, Beijing, 100084, China*

(Dated: May 30, 2019)

The quantum approximate optimization algorithm (QAOA) first proposed by Farhi et al. promises near-term applications based on its simplicity, universality, and provable optimality. A depth-p QAOA consists of p interleaved unitary transformations induced by two mutually non-commuting Hamiltonians. A long-standing question concerning the performance of QAOA is the dependence of its success probability as a function of circuit depth p. We make initial progress by analyzing the success probability of QAOA for realizing state transfer in a one-dimensional qubit chain using two-qubit XY Hamiltonians and single-qubit Hamiltonians. We provide analytic state transfer success probability dependencies on p in both low and large p limits by leveraging the unique spectral property of the XY Hamiltonian. We support our proof under a given QAOA ansatz with numerical optimizations of QAOA for up to N=20 qubits. We show that the optimized QAOA can achieve the well-known quadratic speedup, Grover speedup, over the classical alternatives. Treating the QAOA optimization as a quantum control problem, we also provide numerical evidence of how the circuit depth determines the controllability of the QAOA ansatz.

I. INTRODUCTION

Quantum approximate optimization algorithm (QAOA) promises near-term applications for quantum devices given its simplicity, universality and provable optimality. In contrast to quantum adiabatic algorithms [10, 11], QAOA adopts abrupt switching between two different Hamiltonian evolutions [12]. This simple strategy can reduce the complexity of Hamiltonian controls by obviating the need to continuously vary Hamiltonian in time. Fine-tuned controls over Hamiltonian trajectories are otherwise necessary for the traditional quantum adiabatic algorithms. Despite its simplicity, QAOA is computationally universal [22]. It also has important implications in computational complexity: the efficient classical sampling of a depth-1 QAOA will collapse the polynomial hierarchy to the third level [9]. Lastly, from Pontryagin's maximum principle [31], QAOA is optimal for solving variational problems whose cost function is a linear function of the system Hamiltonian [35].

Despite these attractive properties, to be suitable for near-term quantum devices, a long-standing question remains to be addressed: how does the success probability of QAOA depend on its circuit depth? Near-term quantum device's computation time is limited by noise and decoherence. This in turn limits the realizable quantum algorithms to relatively short circuit depth. To understand QAOA's potential or limitations for near-term applications, it is therefore critical to understand its performance when fixing the upper bound on the depth of the QAOA circuit.

It is exceedingly hard to study the QAOA performance without choosing the QAOA Hamiltonian and optimization problem. This is mainly due to the lack of a sufficient condition for QAOA to achieve optimality in a generic scenario. Since the QAOA success probability directly depends on its optimality, a bound on QAOA success probability scaling usually requires problem-specific numerical optimizations. Recently, specific properties of the chosen optimization problem and QAOA Hamiltonians are utilizied to design protocols that imitate the Grover search algorithm [17, 30], or to prepare highly entangled quantum states [15, 16]. These encouraging results spotlight the importance of the problem and hardware specific information such as the controllable system Hamiltonians in designing QAOA algorithms for improving its performance guarantee.

In this work, we make initial progress towards answering this question by analyzing the performance of QAOA for state transfer in a one-dimensional qubit chain using the XY Hamiltonians and single-qubit Hamiltonians as QAOA ansatzes. We choose quantum state transfer as our QAOA optimization task considering its simplicity and wide applications [1, 2, 5–7, 28, 36]. State transfer is a preliminary requirement for realizing quantum networks which are necessary for connecting quantum computers to form large-scale computation network [8, 18]. We choose the QAOA Hamiltonian ansatz based on the experimental availability, where XY couplings are available in existing superconducting qubit device [14]. Moreover, the XY Hamiltonian's particle number conserving nature makes it suitable for realizing state transfer within a given particle number subspace, which can mitigate unwanted information leakage into the higher excitation subspace.

We harness the analytic spectral features of the XY Hamiltonians to derive explicit success probability P_{succ}

* email: murphyniu@google.com

dependencies on circuit depth p for state transfer in two different limits. In the low circuit depth and short QAOA duration limits we have $\lim_{p\to 0,\delta\to 0} P_{\rm succ} \propto p^2$; and in the large circuit depth limit we have $\lim_{p\to\infty,\delta\to 0} P_{\rm succ} \propto p^{4p}$. Our proof reconfirms the achievable Grover speedup with QAOA ansatz [17]. As a compliment to the system-specific analysis, we apply the existing results of the Lieb-Robinson bound to the QAOA success probability with any 2-local bounded-norm qubit Hamiltonians for one-dimensional state transfer. To verify the optimality of our scaling proof, we numerically optimize the associated QAOA ansatz for different circuit depth and overall runtime. Our numerical results confirm the expected quadratic Grover-like scaling. We also demonstrate an interesting connection between the achievable success probability and its controllability dependency on the circuit depth: once the circuit depth is too low, the QAOA is no longer controllable when the control landscape is full of local optima that are not globally optimal.

The structure of the paper is as follows: in Sec. II we introduce the basic setup of the QAOA for state transfer using the XY Hamiltonian; in Sec. III we derive the QAOA success probability as a function of circuit depth in both low and large circuit depth limit; we discuss the associated quantum speed limit in Sec. IV; we summarize the performance of the QAOA in regard to circuit depth, runtime and number of qubits in Sec. V, and conclude in Sec. VI.

II. QAOA FOR STATE TRANSFER

We introduce in this section the basic concept of quantum state transfer and its realization through QAOA. Quantum state transfer has been proposed in both quantum optical systems and condensed matter system [2, 36]. Different state transfer schemes include the use of quantum disorder [5, 6], optimal control [37], long-range interaction [28] and adiabatic evolution under a moving potential [1]. Unlike the majority of these existing approaches, QAOA resorts to a discrete set of operations of a size given by the QAOA circuit depth. Such a circuit-based model is naturally suitable for near-term quantum devices such as superconducting qubits and ion traps, but more flexible than the traditional circuit-based model using gates only from a predefined universal gate set.

The state transfer problem of interest is defined in a one-dimensional qubit chain of length N. We use $|\overline{n}\rangle$ to represent a product state of a positive eigenstate of the local Pauli-z operator at the nth site and the negative eigenstates of the local Pauli-z of other sites: $\sigma_n^z|\overline{n}\rangle = |\overline{n}\rangle$, $\sigma_{i,i\neq n}^z|\overline{n}\rangle = -|\overline{n}\rangle$, e.g. $|\overline{n}\rangle = |0\rangle_1|0\rangle_2\cdots|0\rangle_{n-1}|1\rangle_n|0\rangle_{n+1}\cdots|0\rangle_N$. We denote $|\overline{0}\rangle$ as the product state of the negative eigenvalue eigenstate of the local Pauli-z operators: $|\overline{0}\rangle = |0\rangle_1|0\rangle_2\cdots|0\rangle_N$. If we treat qubits as spins, and negative eigenvalue of Pauli-z operator as an excitation of the spin state, the state transfer problem we will solve lies in the span of zero and single excitation subspaces. In this subspace, a quantum state with boundary excitation is represented as

$$|\psi_i\rangle = \alpha|\overline{1}\rangle + \beta|\overline{0}\rangle, \tag{1}$$

with $|\alpha|^2+|\beta|^2 = 1$. Starting from the state $|\psi_i\rangle$, the task of state transfer is therefore to realize a unitary transformation U such that:

$$|\psi_f\rangle = U|\psi_i\rangle = \alpha|\overline{N}\rangle + \beta|\overline{0}\rangle. \tag{2}$$

We choose the two Hamiltonians used for QAOA iteration to be

$$\hat{H}_C = |\overline{N}\rangle\langle\overline{N}| = \frac{1}{2}(\sigma_N^z + I_N), \tag{3}$$

$$\hat{H}_B = \sum_{i=1}^{N}(\sigma_i^x\sigma_{i+1}^x + \sigma_i^y\sigma_{i+1}^y). \tag{4}$$

The reasons for our choice of QAOA Hamiltonians are two-fold: (1) $\hat{H}_C$'s plus one eigenstate is our transferred target state $|\overline{N}\rangle$ and can thus serve as a Grover-like oracle by assigning a phase to the target state; (2) $\hat{H}_B$ is off-diagonal and induces a swap operation between neighboring qubits, and thus can move the excitation around for the purpose of state transfer.

Since the total qubit-z operator $S_z = \sum_{i=1}^N \sigma_i^z$ commutes with both $\hat{H}_C$ and $\hat{H}_B$ and $|\overline{1}\rangle$ is an eigenstate of the total qubit-z operator: $S_z|\overline{1}\rangle = (1-N)|\overline{1}\rangle$, the total excitation is conserved throughout the QAOA simulation. We can therefore solve the quantum dynamics in the subspace spanned by $\{|\overline{0}\rangle, |\overline{1}\rangle, |\overline{2}\rangle, \cdots, |\overline{N}\rangle\}$.

Denoting the unitary evolution under $\hat{H}_C$ for a time duration t as $U(\hat{H}_C, t)$ and the unitary evolution under $\hat{H}_B$ for time duration t' as $U(\hat{H}_B, t')$, a depth p QAOA realizes the following unitary transformation:

$$U_p = \prod_{k=1}^{p} U(\hat{H}_C, \delta_k^C)U(\hat{H}_B, \delta_k^B), \tag{5}$$

where the durations of evolutions under given Hamiltonians are represented by δ_k^B and δ_k^C. Here, each kth QAOA iteration consists of a unitary evolution under $\hat{H}_B$ for time δ_k^B then followed by a unitary evolution under $\hat{H}_C$ for time δ_k^C. Since the eigenvalues of $\hat{H}_B$ are not rational, unlike the original QAOA, the rotation angle δ_k^B of $\hat{H}_B$ is not restricted to $(0, 2\pi)$.

Since the zero excitation state of the system is invariant under the unitary evolution of both Hamiltonians, the state transfer task is equivalent to realizing the unitary transformation: $|\overline{1}\rangle \to |\overline{N}\rangle$. Thus, we can quantify the fidelity of state transfer by the fidelity between $U_p|\overline{1}\rangle$ and $|\overline{N}\rangle$:

$$F = |\langle\overline{N}|U_p|\overline{1}\rangle|^2 = \langle\overline{1}|U_p^\dagger\hat{H}_C U_p|\overline{1}\rangle. \tag{6}$$

This is equivalent to the success probability, so we will use them interchangeably henceforth. We can then treat

state transfer as a special kind of maximization satisfaction problem, except that the cost function F here contains only one clause as opposed to many clauses in traditional QAOAs [12].

It is shown in [7] that if we have complete control over the amplitudes of XY coupling at different sites, perfect state transfer can be realized through a single unitary evolution under the XY Hamiltonian. This is, however, unrealistic for near-term devices, where the system calibration for such fine-tuned interactions is costly and the maximum interaction strength is limited.

III. SUCCESS PROBABILITY SCALING AS A FUNCTION OF CIRCUIT DEPTH

In this section, we derive the success probability scaling as a function of circuit depth. Our analysis is based on an iterative procedure using the knowledge from the spectrum of the QAOA ansatz Hamiltonian.

To simplify our analysis, we adopt the following QAOA ansatz: the duration under the evolution of dispersion Hamiltonian $\hat{H}_B$, δ, is short and the same for different iterations, and the evolution under the diagonal Hamiltonian $\hat{H}_C$ is of angle π, resembling a Grover oracle: $U_p = \left(e^{-i\pi|\overline{N}\rangle\langle\overline{N}|}U(\hat{H}_B,\delta)\right)^p$. With this ansatz, our result can be connected to the scaling analysis in conventional Grover search and thus serves as a lower bound on the success probability for the optimized QAOA to be discussed in the subsequent sections.

We first diagonalize the dispersion Hamiltonian $\hat{H}_B$ in the single excitation subspace to obtain its kth eigenstate:

$$|\phi_k\rangle = \frac{1}{\sqrt{N/2}}\sum_{n=1}^{N/2}\left(\sin\left[\frac{kn\pi}{N/4+1}\right]|\overline{2n}\rangle + \sin\left[\frac{k(n+1/2)\pi}{N/4+1}\right]|\overline{2n-1}\rangle\right), \tag{7}$$

with the kth eigenvalue being $E_k = 2\cos[\frac{k\pi}{N/2+1}]$.

Given the initial state of the system as $|\psi_0\rangle = |\overline{1}\rangle$, the system evolves to $|\psi_1\rangle = e^{-i\hat{H}_B\delta}|\overline{1}\rangle$ after a depth one QAOA. Then the success probability of transferring the excitation to the other end of chain after a depth one QAOA is

$$P_{\text{succ}}(1) = \langle\overline{1}|e^{i\hat{H}_B\delta}|\overline{N}\rangle\langle\overline{N}|e^{-i\hat{H}_B\delta}|\overline{1}\rangle = |f_1^N(\delta)|^2, \tag{8}$$

where we use $f_1^N(\delta) = \langle\overline{N}|e^{-i\hat{H}_B\delta}|\overline{1}\rangle$ to represent the amplitude of target state. Now we apply another QAOA iteration to update the quantum system to

$$\begin{aligned}|\psi_2\rangle &= U(\hat{H}_B,\delta)U(\hat{H}_C,\pi)|\psi_1\rangle \\ &= e^{-i2\hat{H}_B\delta}|\overline{1}\rangle - 2e^{-i\hat{H}_B\delta}|\overline{N}\rangle\langle\overline{N}|e^{-i\hat{H}_B\delta}|\overline{1}\rangle\end{aligned} \tag{9}$$

This gives the success probability of the state transfer after a depth two QAOA as:

$$P_{\text{succ}}(2) = \langle\psi_2|\overline{N}\rangle\langle\overline{N}|\psi_2\rangle = |f_1^N(2\delta) - 2f_1^N(\delta)f_N^N(\delta)|^2, \tag{10}$$

where $f_N^N(\delta) = \langle\overline{N}|e^{-i\hat{H}_B\delta}|\overline{N}\rangle$ denotes the amplitude of the state $|\overline{N}\rangle$ remaining in state $|\overline{N}\rangle$ after Hamiltonian evolution under $\hat{H}_B$ for time δ. Similarly, we continue the iteration to obtain the success probability after a depth three and depth four QAOA:

$$P_{\text{succ}}(3) = f_1^N(3\delta) - 2f_1^N(2\delta)f_N^N(\delta) - 2f_1^N(\delta)f_N^N(2\delta) + 4\left(f_N^N(\delta)\right)^2 f_1^N(\delta), \tag{11}$$

$$\begin{aligned}P_{\text{succ}}(4) = {} & f_1^N(4\delta) - 2f_1^N(3\delta)f_N^N(\delta) - 2f_1^N(2\delta)f_N^N(2\delta) - 2f_1^N(\delta)f_N^N(3\delta) + 4\left(f_N^N(\delta)\right)^2 f_1^N(2\delta) \\ & + 4\left(f_N^N(\delta)\right)^2 f_N^N(2\delta) + 4f_N^N(2\delta)f_N^N(\delta)f_1^N(\delta) - 8\left(f_N^N(\delta)\right)^3 f_1^N(\delta).\end{aligned} \tag{12}$$

Now we provide expression for transition amplitude used above given the exact eigenstates of the dispersion Hamiltonian $\hat{H}_B$ in Eq. 7:

$$
\begin{aligned}
f_1^N(\delta) &= \sum_{k=1}^{N/2} \langle \overline{N}|\phi_k\rangle\langle\phi_k|\overline{1}\rangle e^{-iE_k\delta}, \\
&\approx -\frac{i\delta}{2\sqrt{N+1}}\left(\cos\left(\frac{2\pi(N/2)^2+4\pi N/2+\pi}{2(N/4+1)}\right)\csc\left(\frac{\pi N/2+2\pi}{2(N/4+1)}\right)+\cos\left(\frac{\pi}{2(N/4+1)}\right)\csc\left(\frac{\pi N/2+2\pi}{2(N/4+1)}\right)\right) \\
&= -iF(N)\delta,
\end{aligned}
\tag{13}
$$

$$
\begin{aligned}
f_N^N(\delta) =& \sum_{k=1}^{N/2} \langle \overline{N}|\phi_k\rangle\langle\phi_k|\overline{N}\rangle e^{-iE_k\delta}, \\
\approx& \frac{i\delta}{2\sqrt{N+1}}\left\{\frac{1}{2}\left[-\cos\left(\frac{4\pi N^2+\pi N-\pi}{2(N+1)}\right)\csc\left(\frac{\pi N}{N+1}\right)+2N+\cos\left(\frac{\pi(N-1)}{2(N+1)}\right)\csc\left(\frac{\pi N}{N+1}\right)\right]\right. \\
&\times\left[\cos\left(\frac{\pi N}{2(N+1)}\right)\csc\left(\frac{\pi}{2(N+1)}\right)+\cos\left(\frac{\pi(N+2)}{2(N+1)}\right)\csc\left(\frac{\pi}{2(N+1)}\right)\right] \\
&-\left[\cos\left(\frac{3\pi N}{2(N+1)}\right)\csc\left(\frac{\pi-2\pi}{2(N+1)}\right)-\cos\left(\frac{-4\pi N^2-\pi N}{2(N+1)}\right)\csc\left(\frac{\pi-2\pi N}{2(N+1)}\right)\right] \\
&\left.-\left[\cos\left(\frac{\pi N}{2(N+1)}\right)\csc\left(\frac{2\pi N+\pi}{2(N+1)}\right)-\cos\left(\frac{4\pi N^2+3\pi N}{2(N+1)}\right)\csc\left(\frac{2\pi N+\pi}{2(N+1)}\right)\right]\right\} \\
=& -iG(N)\delta
\end{aligned}
\tag{14}
$$

where the approximation is made to include only the terms that are of either zero or first order in δ, under the limit $\delta \to 0$. And both $F(N)$ and $G(N)$ are real-valued. For a QAOA of depth p, we deduce the success probability dependency on transition probabilities $f_1^N(\delta)$ and $f_N^N(\delta)$ as

$$
P_{\text{succ}}(p) = \sum_{j=1}^{p}(-1)^j \sum_{\vec{v}_j\in\mathcal{V}_j} f_1^N(\vec{v}_j(1)\delta)\prod_{k=1}^{j-1} f_N^N(\vec{v}_j(k+1)\delta) \tag{15}
$$

where $\vec{v}_j$ is a vector with each element representing the value of a j partition of p that belongs to the set $\mathcal{V}_j = \{\vec{v}_j|\sum_{k=1}^{j}\vec{v}_j(k)=p\}$. The success probability can be in turn be expressed as

$$
P_{\text{succ}}(p) \approx \left|\sum_{j=1}^{p+1} A_j\delta^j\right|^2, \tag{16}
$$

with the each amplitude A_j given by

$$
A_1 = -ipF(N), \tag{17}
$$

$$
A_2 = -F(N)G(N)\frac{p(p+1)(p+2)}{3}, \tag{18}
$$

$$
\lim_{n\to\infty} A_n \approx -F(N)G(N)^n p^{2n-1}. \tag{19}
$$

Eq. (19) is derived from the asymptotic value of the product of all possible values of n integers $p_1, p_2, \ldots, p_n$ whose sum equals p: $\sum_{i=1}^{n} p_i = p$.

So far we have only kept the leading order in $O(\delta)$ together with all orders of $O\left(\left(p^2\delta\right)^n\right)$ which will be non-negligible when $p^2\delta \sim 1$ or $p^2\delta \gg 1$. For the scaling analysis, we neglect constant terms in the sum and find the success amplitude to be

$$\begin{aligned}\sum_{j=1}^{p+1} A_j\delta^j &= -i(p+1)F(N)\delta - \left[F(N)G(N)p^3\delta^2 + \cdots + F(N)G(N)^p p^{2p+1}\delta^{p+1}\right] \\ &= -i(p+1)F(N)\delta - \frac{F(N)G(N)p^3\delta^2\left[(G(N)p^2\delta)^p - 1\right]}{(G(N)p^2\delta)^2 - 1}.\end{aligned} \tag{20}$$

Since the amplitude is composed of imaginary and real parts, the success probability of a depth-p QAOA is thus found to be

$$P_{\text{succ}}(p) \approx (p+1)^2F(N)^2\delta^2 + \frac{F(N)^2G(N)^2p^6\delta^4\left[(G(N)p^2\delta)^p - 1\right]^2}{\left[(G(N)p^2\delta)^2 - 1\right]^2}. \tag{21}$$

This success probability dependence can be used as a lower bound on the QAOA performance after optimization where the duration of each evolution can be of flexible value. In the large depth limit, the term with the largest power of p dominates:

$$\lim_{p\to\infty} P_{\text{succ}}(p) \propto p^{4p+2}. \tag{22}$$

This exponential growth in success probability is based on the assumption that δ is a small constant (which does not change with the circuit depth p), and the dominant contribution to the transition amplitudes is of the lowest order in δ. Such exponential dependence is also found by increasing the speed of adiabatic Hamiltonian evolution, see [27]. We also observe such exponential growth in our numerically optimized QAOA (see Sec. III).

In the low-depth limit, only the lowest order of δ terms dominates:

$$\lim_{p\to 1} P_{\text{succ}}(p) \propto F(N)^2\delta p^2. \tag{23}$$

Then the total number of steps required to achieve the target state is of order $O(1/(\delta F(N))) = O(\sqrt{N})$. This quadratic Grover-like dependence on circuit depth can be understood by mapping a Grover algorithm to a QAOA routine. The dispersion step of Grover iteration, which is a rotation of angle π around the equal superposition state:

$$U_s = H^{\otimes N}(-2|0\rangle\langle 0| + I)H^{\otimes N}, \tag{24}$$

with H representing the Hadamard gate, can be generalized to a rotation around any state that is not parallel to the target state $|\psi\rangle$ [23]:

$$U_s = -2|\psi\rangle\langle\psi| + I. \tag{25}$$

If we choose $|\psi\rangle = e^{-i\delta\hat{H}_B}|\overline{1}\rangle$, the corresponding pth Grover iteration for searching the transferred state $|\overline{N}\rangle$ starting from the initial state $|\overline{1}\rangle$ can then be represented by the unitary realized by a depth p QAOA circuit as

$$U^p_{\text{Grover}} = \left(e^{-i\hat{H}^1_c\pi}e^{i\hat{H}_B\delta_1}e^{-i\hat{H}^2_c\pi}e^{-i\hat{H}_B\delta_1}\right)^p, \tag{26}$$

where $\hat{H}^1_c = \frac{1}{2}(\sigma^z_N + I_N)$ and $\hat{H}^2_c = \frac{1}{2}(\sigma^z_1 + I_1)$.

IV. QUANTUM SPEED LIMIT

As a supplement to the success probability scaling analysis presented in the previous section that is limited to our specific choices of the QAOA Hamiltonians, we review in this section the general constraints on the QAOA performance using spatially local Hamiltonians imposed by the Lieb-Robinson bound. The Lieb-Robinson bound [21] is a powerful tool to study the propagation of quantum correlation and thus quantum information in many-body quantum systems [13]. It serves as a lower bound on the success probability for the QAOA performance of the same total runtime. Although such a bound is not tight, nor does it directly depends on the circuit depth of QAOA, it provides a basic reference of the optimality of QAOA in regard to its success probability scaling as a function of physical time. And in fact, the theoretical insights of Lieb-Robinson bound help us to understand the performance of numerically optimized QAOAs in the next section.

We rewrite our QAOA iterations as an evolution under the time-dependent Schródinger equation with the time-dependent Hamiltonian:

$$\hat{H}(t) = s(t)\hat{H}_C + [1 - s(t)]\hat{H}_B, \tag{27}$$

where $s(t)$ is the time varying control parameter that can only take on the values zero and one. Thus it realizes the bang-bang form of QAOA iterations.

Note $\hat{H}(t)$ can be written as the sum of nearest-neighbor interacion terms: $\hat{H}(t) = \sum_i h_{i,i+1}(t)$. Thus, by Lieb-Robinson bound, the maximum speed of quantum information propagation in this system is bounded. This speed determines how fast operations on the first qubit can affect observables on the last qubit at some later moment of time, and thus upper bounds the speed of state transfer. For convenience, let us call the first qubit A, the last qubit B, and the rest part C (see

Fig. 1). The Lieb-Robinson bound determines the maximum operator norm of the commutator between any operators O_A and O_B on the first and the last qubit at a later time t. More specifically, denoting $L = N - 1$ as the distance between the first and the last qubit and $J = \max_i \max_t \|h_{i,i+1}(t)\|$ as the maximum interaction strength, the Lieb-Robinson bound for a nearest-neighbour Hamiltonian on a D-dimensional square lattice [13, 28] is given by

$$\|[O_A(t), O_B(0)]\| \leq 2\|O_A\|\|O_B\| \sum_{k=L}^{\infty} \frac{(2Jt(4D-1))^k}{k!}. \tag{28}$$

In the one-dimensional system, the above bound simplifies to [28]:

$$\|[O_A(t), O_B(0)]\| \leq 2\|O_A\|\|O_B\| \sum_{k=L}^{\infty} \frac{(6Jt)^k}{k!} \tag{29}$$

$$\leq 2\|O_A\|\|O_B\| \exp(6eJt - L) \tag{30}$$

$$= 2\|O_A\|\|O_B\| \exp(vt - L), \tag{31}$$

where the Lieb-Robinson velocity $v = 6eJ$ is approximately 32.616 because $J = \|\sigma_i^x \sigma_i^x + \sigma_i^z \sigma_i^z\| = 2$. Consequently, $1/v \approx 0.03$.

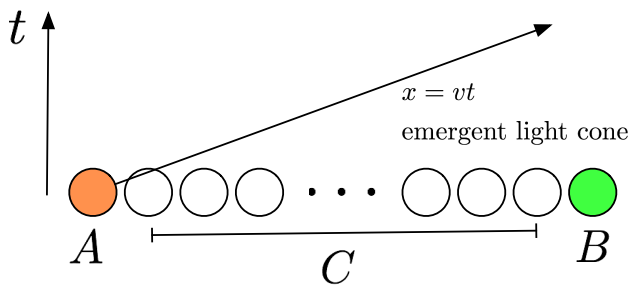


FIG. 1. The emergent light cone in state transfer problem, where the quantum state localized at A (the first qubit) is transfered to B (the last qubit) through the quantum channel C consisting of qubits in the middle. In the short-range two-local system, a non-relativistic light-cone $x = vt$ emerges. The amount of information that can be transferred outside of the lightcone is exponentially small.

Let $U_A^0 = I_A$ and $U_A^1 = \sigma_A^x$. The unitary transformation of the whole system is induced by the time-dependent Hamiltonian as $U_{ABC}(t) = \mathcal{T} \exp(\int_0^t -i\hat{H}(t')\mathrm{d}t')$. Specifically, with initial state of the system given by $\rho_0 = |\overline{0}\rangle\langle\overline{0}|$, we can interpret this procedure as a quantum channel where the input state is $\rho_{ABC}^k = U_A^k \rho_0 U_A^{k\dagger}$, and the output state, the reduced density matrix of B, is $\sigma_B^k(t) = \mathrm{Tr}_{AC}(U_{ABC}(t)\rho_{ABC}^k U_{ABC}^\dagger(t))$.

Ref. [4] shows that $\sigma_B^k(t)$ depend on k as follows. For any observable O_B and associated time evolved operator $O_B(t) = U_{ABC}^\dagger(t) O_B U_{ABC}(t)$, we have:

$$\mathrm{Tr}_B \left[O_B(\sigma_B^0(t) - \sigma_B^1(t))\right] \leq \epsilon \|O_B(t)\|, \tag{32}$$

where $\epsilon = 2\exp(vt - L)$ is given by the Lieb-Robinson bound. Here, we have used $U_A^{1\dagger} U_A^1 = I$, the definition of operator norm, and the Lieb-Robinson bound.

Therefore $\sigma_B^1(t)$ and $\sigma_B^0(t)$ are ϵ close in the trace norm: $\|\sigma_B^1(t) - \sigma_B^0(t)\|_1 \leq \epsilon$. By using the following inequality between the trace distance and the fidelity, $F(\rho, \sigma) \geq 1 - \frac{1}{2}\|\rho - \sigma\|_1$, we obtain a bound on the fidelity between $\sigma_B^1(t)$ and $\sigma_B^0(t)$:

$$F(\sigma_B^1(t), \sigma_B^0(t)) \geq 1 - \frac{1}{2}\epsilon. \tag{33}$$

Now we would like to bound the success probability $P(t)$ of a excitation of qubit state transferred from site 1 to the site N after the evolution under $\hat{H}(t)$ for time t which depends on the fidelity defined above as:

$$1 - P(t) = F^2(\sigma_B^1(t), \sigma_B^0(t)) \geq (1 - \frac{1}{2}\epsilon)^2. \tag{34}$$

Rearranging, we obtain an upper bound for the success probability of the QAOA as a function of time and the length of the qubit chain:

$$P(t) \leq \epsilon - \frac{1}{4}\epsilon^2, \tag{35}$$

From the above expression, we can identify three different regions of temporal dynamics. At the early time when $t \ll L/v$, we have $\epsilon = 2\exp(vt - L)) \ll 1$ and the probability of success is nearly zero. In this first region, the success probability is *exponentially suppressed* and remains almost zero in time. When $t \approx L/v$ we have $\epsilon = c\exp(vt - L) < 1$ and the first term of the right-hand side of Eq. (35) dominates, which gives rise to an *exponentially growing* success probability. Finally when $t > L/v$, the second term of Eq. (35) starts balancing out the first term, and gives rise to a *steady growing* region. A rough estimation of perfect state transfer time can be given by setting $\epsilon - \frac{1}{4}\epsilon^2 = 1 \Rightarrow \epsilon = 2$, which gives

$$t \approx L/v. \tag{36}$$

The main weakness of the Lieb-Robinson method is the lack of dependence on the specific form of Hamiltonian and the circuit depth. Nevertheless, it offers useful insights into the difficulty of state transfer problems. In the later numerical section, we confirm the existences of the exponentially suppressed region, the exponentially growing region (see Fig. 9), the steady growing region, and the linear dependence between t_f required for state transfer and the number of qubits N (see Fig. 11).

V. NUMERICAL OPTIMIZATION OF THE QAOA

Our analytic success probability versus circuit depth scaling analyses so far do not assume the optimality of the QAOA solution. To verify the tightness of these results for optimized QAOAs, we explore in this section

the numerically optimized QAOA performance in regard to its success probability scaling as a function of the circuit depth and the physical runtime. We start by introducing briefly our numerical optimization method, and then describe and analyze the optimized QAOA performance obtained from our numerical method. We show that the quadratic Grover-like speedup shown in our analytic spectral analysis is also present in the numerically optimized QAOA solutions. We also show that when the circuit depth is too low, the given QAOA protocol becomes uncontrollable: its control landscape no longer possesses only global optimal but also many local optimum points. Consequently, the optimized QAOA no longer necessarily guarantees the existence of a high fidelity state transfer scheme. This finding unveils the relation between the $F - p$ scaling and the controllability of the underlying physical system.

We optimize QAOA parameters under two different constraints, one with limited physical run time and the other one with unlimited physical run time. Both situations are experimentally relevant. If the coherence time is sufficiently large, we may want to achieve state transfer with a minimum number of switches between $\hat{H}_B$ and $\hat{H}_C$. In contrast, if the switching operation is easy, the total physical run time should be minimized. The practical implementation of QAOAs on near-term quantum computing hardware is outside the scope of this article.

A. Numerical methods

We use a gradient descent method to numerically determine the maximum achievable fidelity given t_f and p. We choose the optimization parameters as the duration for each QAOA Hamiltonian evolution δ_B^k and δ_C^k of each kth iteration with $k \in [p]$ for a depth-p QAOA. The total physical runtime is the sum of all QAOA iterations: $t_f = \sum_k (\delta_B^k + \delta_C^k)$. We chose the parameter ranges for the physical run time t_f and the circuit depth p by preliminary numerical experiments. Table II and **??** summarize the parameters we performed grid search on. To increase the reliability of the total number of random restarts for gradient descent iterations, we ran preliminary experiments with a varying size of random uniform (over 1-simplex) initial conditions, and discovered that 200 random initial conditions were enough for $N = 2 \to 15$ qubits cases to find approximate global optimal solutions. For $N = 16 \to 20$ qubits, we increased the number of random restarts to 400. We use L-BFGS Matlab toolbox for the QAOA optimization, which can be efficiently parallelized for large-scale experiments. These numerical results to be discussed below are summarized in Table III.

B. Numerical results for unlimited t_f

We start with the optimized QAOA performance when the physical run time t_f is not fixed. The analytic result at low circuit depth limit (Eq. (23)) coincides with our numerically result in Section V B 1. And the connection between QAOA circuit depth and the controllability is demonstrated in the numerical results on the control landscape in Section V B 2.

1. *Maximum achievable fidelity versus circuit depth p*

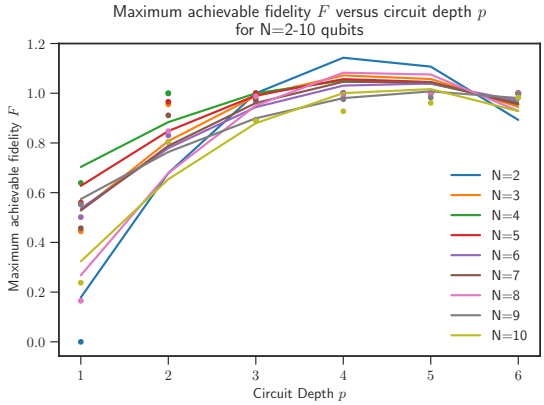
Figure 2 shows maximum achievable fidelity F as a function of circuit depth p with no constraints on t_f for $N = 2 \to 20$ qubits. The circuit depth dependence of success probability in Fig. 2 agrees with our analytic result: at the very beginning the quadratic dependence dominates (Eq. (23)), and a while later the exponential slow down dominates (Eq. (35)). The time duration ansatz used in our analytical result in Sec. III is also supported in our numerically optimized solution; see Fig. 4 for example, where the intervals for $s(t) = 1$ corresponding to the evolution under $\hat{H}_C$ is shorter on average than the duration in which $s(t) = 0$.

2. *Control landscape*

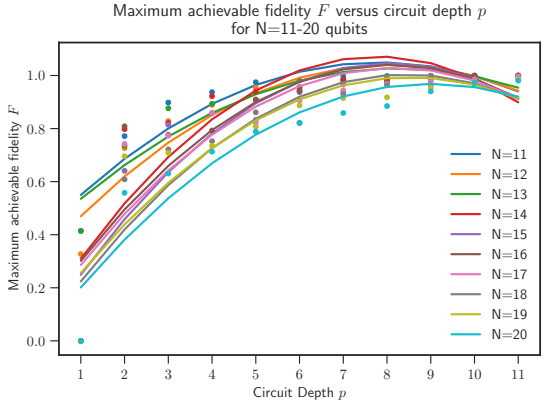
The optimization of QAOA can be regarded as a quantum control problem, where the durations of different QAOA Hamiltonian evolutions are the control parameters, and the fidelity of the state transfer is the control cost function to be maximized. Under this analogy, when the system is controllable, the landscape of the control cost function over parameter space generically has only global minima [26, 29, 34]. When the system is uncontrollable, the quantum control landscape admits many local minima [33]. In our case, if we allow the circuit depth to be infinite, the system is controllable; but the controllability for an intermediate number of circuit depths demands further investigation. As a simple example, we plot the control landscape for $N = 3$ qubits in Fig. 5 with $p = 2, 4$. Fig. 5 shows that the QAOA ansatz with $p = 2$ is uncontrollable, as there are many local minima. In contrast, the case $p = 4$ admits only global minima (at least for the plotted part); thus more controllable. We review the control viewpoint of optimizing QAOA in Appendix A.

C. Numerical results for fixed t_f: optimized QAOA

We now discuss the optimized QAOA performance with a fixed physical run time t_f. In Section V C 1, we investigate the controllability dependence on t_f. Second, in Section V C 2, we discuss the Lieb-Robinson type scaling emerged in the optimized QAOA performance.

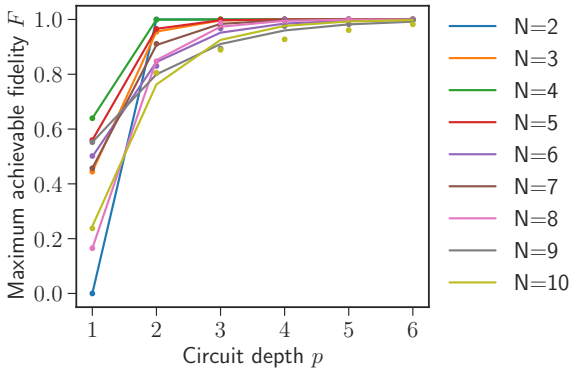


(a) $N = 2 \to 10$ qubits

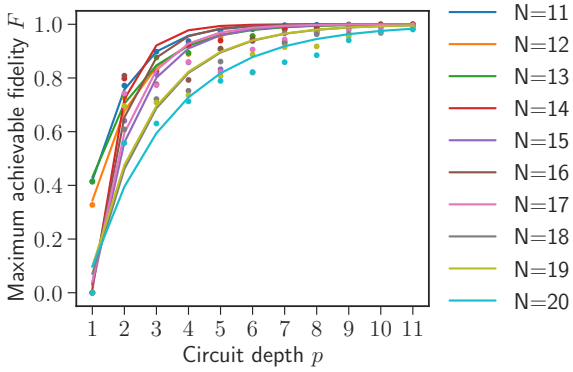


(b) $N = 11 \to 20$ qubits

FIG. 2. Maximum achievable fidelity F using QAOAs as a function of the circuit depth p with no constraints on t_f for $N = 2 \to 20$ qubits. The dots are numerical points. We fit the results with quadratic function $F(p) = ap^2 + bp + c$, as represented by the lines. We observe that the fidelity grows at low circuit depth, and then slowly converges to 1.0.



(a) $N = 2 \to 10$ qubits



(b) $N = 11 \to 20$ qubits

FIG. 3. Maximum achievable fidelity F as a function of the circuit depth p with no constraints on t_f for $N = 2 \to 20$ qubits.The dots are numerical points. We fit the results with inverted exponential function $F(p) = 1 - \exp(-a(p-b))$. We can observe fidelity grows rapidly at low circuit depth, and then the fidelity slowly converge to unity.

1. *Maximum achievable fidelity F versus circuit depth p*

In this section, we numerically study the maximum achievable fidelity F versus the circuit depth p for a fixed t_f in Figs. 6 to 7. Generally speaking, the larger circuit depth QAOA should always perform better than lower depth ones. However, if p is too large, the difficulty of the QAOA optimization increases and the optimization can get stuck in local optima. This results in a non-monotonic behavior in numerically optimized fidelity as a function of circuit depth. For fixed t_f, there is a circuit depth p beyond which fidelity can no longer be improved. As shown in Fig. 6, for $t_f = 6$, the maximum achievable fidelity does not increase for circuit depth larger than $p = 3$, and for $t_f = 13$, the maximum achievable fidelity does not increase for circuit depth larger than $p = 4$. This observation is also intimately related to the controllability of the QAOA: for the fixed run time, there exists a threshold circuit depth below which the QAOA is no longer controllable.

2. *Maximum achievable fidelity F versus physical run time t_f*

In this subsection, we investigate the performances of the QAOA with a fixed physical runtime in Figure 8. We identify three different temporal dependencies of fidelity as predicted by the Lieb-Robinson bound, as depicted in Fig. 9: *exponentially suppressed region*; *exponentially growing region*; and *steady growing region*. We find that the longer the physical run time t_f is, the better achievable fidelity will be under the condition that the circuit depth p is sufficiently large and t_f is outside of the highly suppressed region. For a low depth circuit, the oscillating in success probability occurs (Fig. 8(a)), which is a sign of uncontrollability. Such oscillation disappears for sufficiently large p, see Fig. 8(b). The three regions of the growth region become more apparent as circuit depth in-

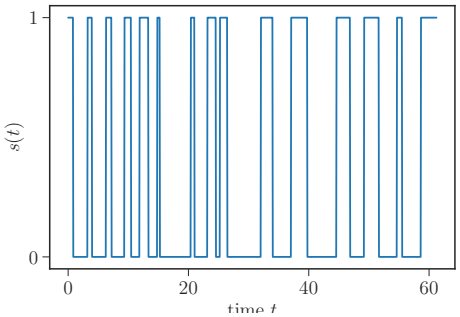
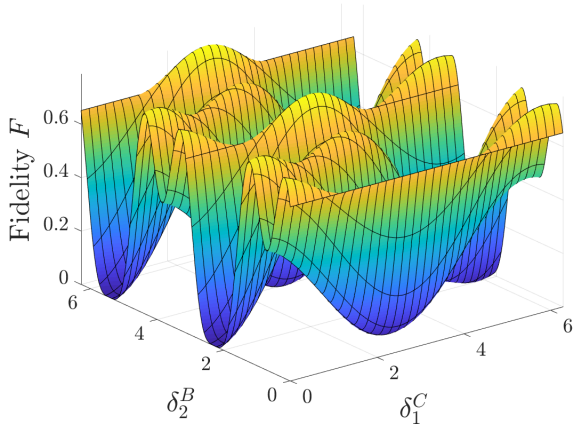


FIG. 4. An optimal bang-bang solution obtained through numerical optimizations for $N = 10$ qubits and circuit depth $p = 15$. In this case, the optimal solution favors much shorter duration for the evolution under $\hat{H}_B$ than that under $\hat{H}_C$. This is in accord with our analytical result in Sec. III.

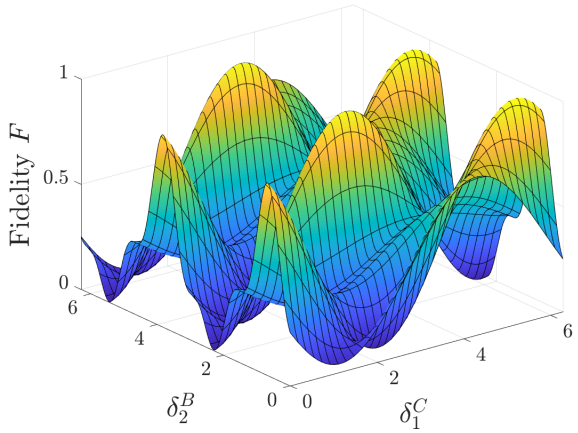
creases, see in Fig. 10.

In Fig. 11, we fit the minimum required run time t_f for achieving fidelity $F = 0.99$ as a function of the number of qubits ($N = 2 \to 19$). The linear dependence from the Lieb-Robinson bound Eq. (36) is seen with $t_f \sim 2.439N$. Given the same amount of run time, we show in Fig. 12 that the QAOA with a higher circuit depth necessarily achieves higher success probability.

The Lieb-Robinson bound gives a prediction about the size of the exponentially suppression region: $t_s \sim N/(6eJ) = 0.03N$. Practically, we define the exponentially suppressed time as the time needed to make fidelity higher than 0.01. To see if it agrees with our numerical result, we plot the exponentially suppressed time as a function of the number of qubits in Fig. 13. Our numerical result is $t_s \sim 0.246N$ with a coefficient of determination $r^2 = 0.997$. We remark that the discrepancy between $0.246N$ and $0.03N$ is because our QAOAs only operate in the span of zero and single excitation subspaces, while the Lieb-Robinson bound considered the full N-qubit Hilbert space.



(a)

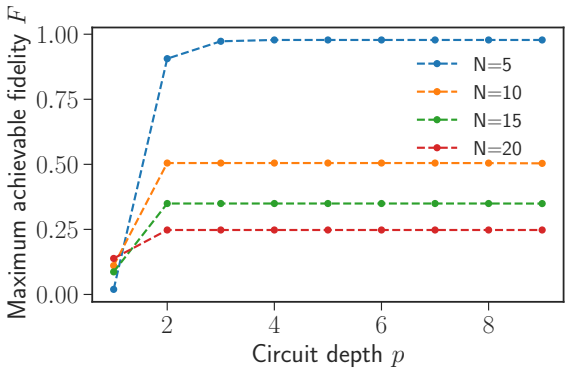


(b)

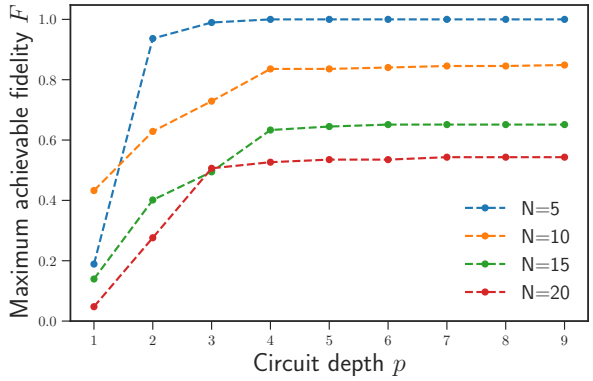
FIG. 5. A comparison of the control landscapes of the QAOA with low circuit depth ($p = 2$) and of that with a larger circuit depth ($p = 4$) for a three-qubit system. (a) The control landscape for two chosen variables of a depth-2 QAOA, which admits a maximum achievable fidelity of 0.787. As we observed many local minima, the system is uncontrollable. (b) The control landscape for two chosen variables of a depth-4 QAOA, which admits a maximum achievable fidelity of 1.000. Since all local minima are global minima, the system is controllable.

VI. SUMMARY

We study the QAOA's success probability scaling as a function of circuit depth and the physical runtime for implementing state transfer problems. By carefully utilizing the spectral properties of the QAOA Hamiltonians, we obtain analytic expressions for the success probability scaling as a function of the circuit depth. At the low-circuit-depth and short-physical-duration limit, our analytic results reproduce the Grover-like quadratic speedup. We further study the success probability scaling in numerically optimized QAOAs for a chain of up to $N = 20$ qubits (limited by computational resources). These numerical experiments confirm the quadratic speed up and match with the Lieb-Robinson analysis of quantum speed limit, i.e., when the circuit depth p is sufficiently large, with the increase of t_f, there are three different scenarios of success probability scaling: (1) exponentially suppressed region, (2) exponentially growing region and (3) steadily growing region. Treating QAOA optimization as a quantum control problem, we demonstrate the relation between the circuit depth and the controllability of QAOA: when the circuit depth is too low for a fixed distance state transfer, the control landscape possesses many locally optimal solutions that are not globally optimal and the QAOA becomes uncontrollable. Although the state transfer problem we considered here is relatively simple, , our results offer valuable insights into the performance of QAOAs by connecting its optimality to the



(a) Plots for a small $t_f = 6$.



(b) Plots for a medium $t_f = 13$.

FIG. 6. Maximum achievable fidelity F versus circuit depth p with the same fixed t_f for $N = 5, 10, 15, 19$ qubits. For fixed t_f, we observed that there exists a circuit depth p beyond which there will be no improvement of fidelity. We find a depth-3 circuit is sufficient for $t_f = 6$ while a depth-4 circuit is needed for $t_f = 13$.

Grover speedup and its success probability dependence on circuit-depth to the controllability the QAOA ansatz. To fully explore the application of QAOA, however, more work remains to be done to study the effect of realistic quantum noise on QAOA implementations.

[1] V. Balachandran and J. Gong. Adiabatic quantum transport in a spin chain with a moving potential. *Phys. Rev. A*, 77:012303, Jan 2008.

[2] S. Bose. Quantum communication through an unmodulated spin chain. *Phys. Rev. Lett.*, 91:207901, Nov 2003.

[3] F. G. Brandao, M. Broughton, E. Farhi, S. Gutmann, and H. Neven. For fixed control parameters the quantum approximate optimization algorithm's objective function value concentrates for typical instances. *arXiv preprint arXiv:1812.04170*, 2018.

[4] S. Bravyi, M. B. Hastings, and F. Verstraete. Lieb-robinson bounds and the generation of correlations and topological quantum order. *Phys. Rev. Lett.*, 97:050401, Jul 2006.

[5] C. K. Burrell and T. J. Osborne. Bounds on the speed of information propagation in disordered quantum spin chains. *Phys. Rev. Lett.*, 99:167201, Oct 2007.

[6] C. K. Burrell, J. Eisert, and T. J. Osborne. Information propagation through quantum chains with fluctuating disorder. *Phys. Rev. A*, 80:052319, Nov 2009.

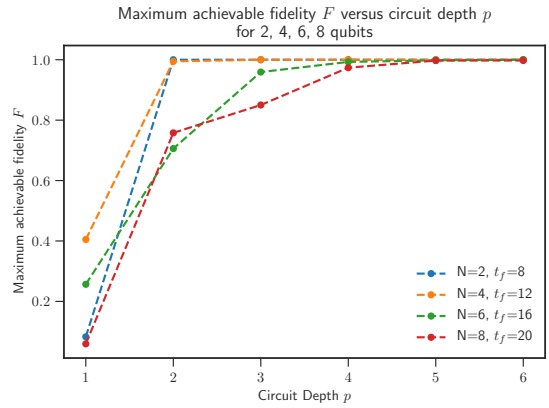
[7] M. Christandl, N. Datta, T. C. Dorlas, A. Ekert, A. Kay, and A. J. Landahl. Perfect transfer of arbitrary states in quantum spin networks. *Phys. Rev. A*, 71:032312, Mar 2005.

[8] J. I. Cirac, P. Zoller, H. J. Kimble, and H. Mabuchi. Quantum state transfer and entanglement distribution among distant nodes in a quantum network. *Phys. Rev. Lett.*, 78:3221–3224, Apr 1997.

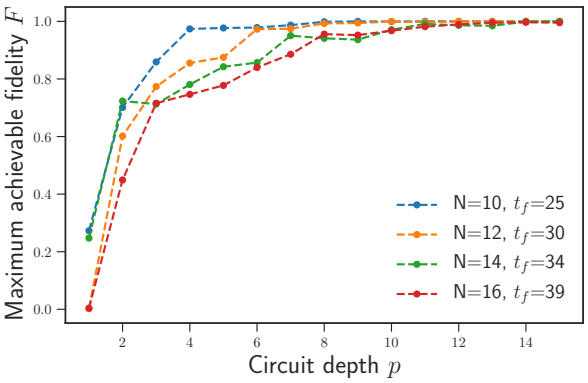
[9] E. Farhi and A. W. Harrow. Quantum supremacy through the quantum approximate optimization algorithm. *arXiv preprint arXiv:1602.07674*, 2016.

[10] E. Farhi, J. Goldstone, S. Gutmann, and M. Sipser. Quantum computation by adiabatic evolution. *arXiv preprint quant-ph/0001106*, 2000.

[11] 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(5516):472–475, 2001.

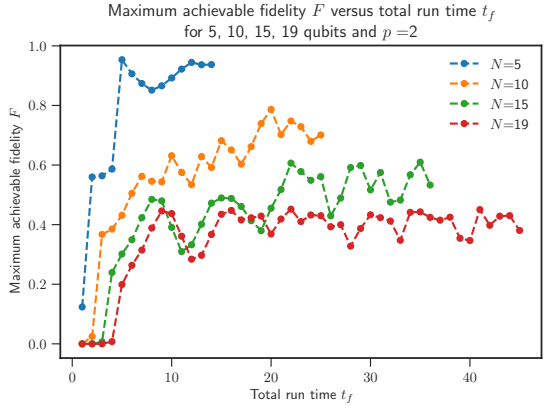


(a) $N = 2, 4, 6, 8$ qubits.

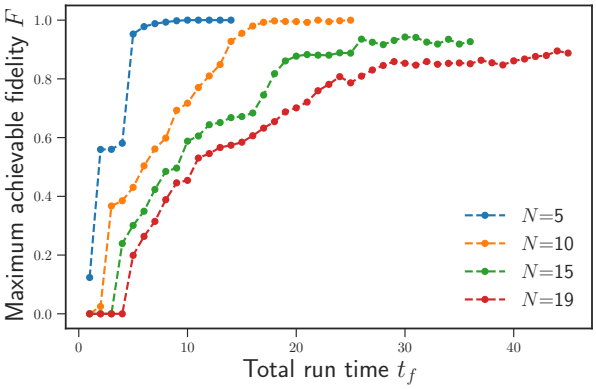


(b) $N = 10, 12, 14, 16$ qubits.

FIG. 7. Maximum achievable fidelity F versus circuit depth p with fixed t_f. The physical run time t_f is chosen through a hyper-parameter grid search separately.



(a) Plots for $p = 2$ (uncontrollable).



(b) Plots for $p = 9$ (controllable).

FIG. 8. Maximum achievable fidelity F versus physical run time t_f with a fixed circuit depth p for $N = 5, 10, 15, 19$ qubits. The oscillating behavior is due to the low circuit depth; such behaviour disappears when the circuit depth p grows larger (equivalently, the system gets more controllable).

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

[13] M. B. Hastings. *Locality in quantum systems*, volume 95. Oxford University Press, 2010.

[14] U. L. Heras, A. Mezzacapo, L. Lamata, S. Filipp, A. Wallraff, and E. Solano. Digital quantum simulation of spin systems in superconducting circuits. *Phys. Rev. Lett.*, 112:200501, May 2014.

[15] W. W. Ho and T. H. Hsieh. Efficient variational simulation of non-trivial quantum states. *SciPost Phys.*, 6:29, 2019.

[16] W. W. Ho, C. Jonay, and T. H. Hsieh. Ultrafast state preparation via the quantum approximate optimization algorithm with long range interactions. *arXiv preprint arXiv:1810.04817*, 2018.

[17] 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, Jun 2017.

[18] H. J. Kimble. The quantum internet. *Nature*, 453(7198): 1023, 2008.

[19] V. Krotov. *Global methods in optimal control theory*, volume 195. CRC Press, 1995.

[20] X. Li et al. Perfect quantum state transfer in a superconducting qubit chain with parametrically tunable couplings. *Phys. Rev. Applied*, 10:054009, Nov 2018.

[21] E. H. Lieb and D. W. Robinson. The finite group velocity of quantum spin systems. In *Statistical Mechanics*, pages 425–431. Springer, 1972.

[22] S. Lloyd. Quantum approximate optimization is computationally universal. *arXiv preprint arXiv:1812.11075*, 2018.

[23] M. A. Nielsen and I. Chuang. *Quantum computation and quantum information*. Cambridge University Press, 2011.

[24] H. Pichler, S.-T. Wang, L. Zhou, S. Choi, and M. D. Lukin. Computational complexity of the rydberg blockade in two dimensions. *arXiv preprint arXiv:1809.04954*, 2018.

[25] H. Pichler, S.-T. Wang, L. Zhou, S. Choi, and M. D. Lukin. Quantum optimization for maximum independent set using rydberg atom arrays. *arXiv preprint arXiv:1808.10816*, 2018.

[26] H. A. Rabitz, M. M. Hsieh, and C. M. Rosenthal. Quantum optimally controlled transition landscapes. *Science*, 303(5666):1998–2001, 2004.

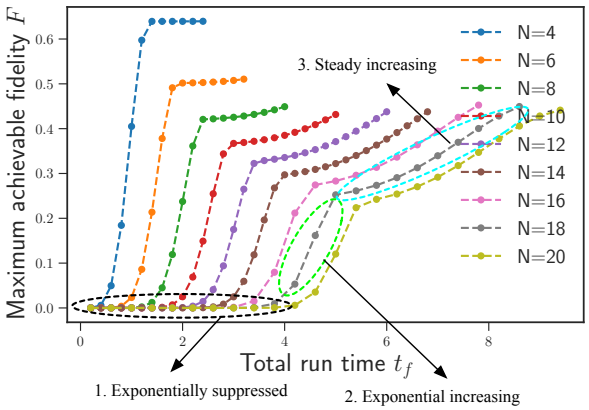
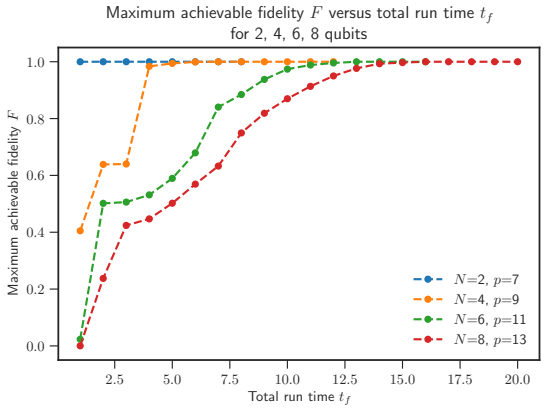
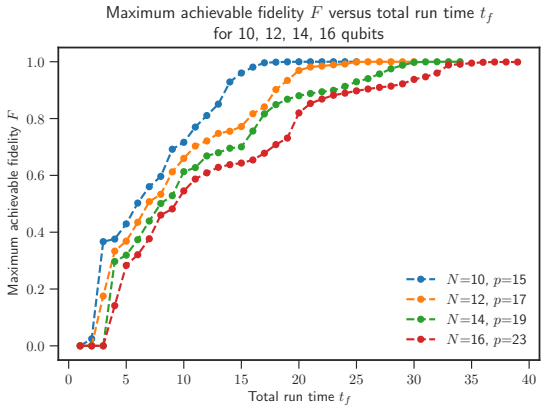


FIG. 9. Maximum achievable fidelity F versus small physical run time t_f with a sufficiently large circuit depth p for $N = 4, 6, 8, 10, 12, 14, 16, 18, 20$ qubits. In general, we identify three different growing patterns: (1) *exponentially suppressed region;* (2) *exponentially increasing region;* (3) *steady increasing region.*



(a) $N = 2, 4, 6, 8$ qubits.



(b) $N = 10, 12, 14, 16$ qubits.

FIG. 10. Maximum achievable fidelity F versus physical run time t_f with a sufficiently large circuit depth p.

[27] A. T. Rezakhani, A. K. Pimachev, and D. A. Lidar. Accuracy versus run time in an adiabatic quantum search. *Phys. Rev. A*, 82:052305, Nov 2010.

[28] P. Richerme, Z.-X. Gong, A. Lee, C. Senko, J. Smith, M. Foss-Feig, S. Michalakis, A. V. Gorshkov, and C. Monroe. Non-local propagation of correlations in quantum systems with long-range interactions. *Nature*, 511(7508):198–201, July 2014.

[29] B. Russell, H. Rabitz, and R. Wu. Quantum control landscapes are almost always trap free. *arXiv preprint arXiv:1608.06198*, 2016.

[30] V. N. Smelyanskiy, K. Kechedzhi, S. Boixo, S. V. Isakov, H. Neven, and B. Altshuler. Non-ergodic delocalized states for efficient population transfer within a narrow band of the energy landscape. *arXiv preprint arXiv:1802.09542*, 2018.

[31] R. F. Stengel. *Optimal control and estimation.* Courier Corporation, 2012.

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

[33] R.-B. Wu, M. A. Hsieh, and H. Rabitz. Role of controllability in optimizing quantum dynamics. *Phys. Rev. A*, 83:062306, Jun 2011.

[34] R.-B. Wu, R. Long, J. Dominy, T.-S. Ho, and H. Rabitz. Singularities of quantum control landscapes. *Phys. Rev. A*, 86:013405, Jul 2012.

[35] Z.-C. Yang, A. Rahmani, A. Shabani, H. Neven, and C. Chamon. Optimizing variational quantum algorithms using pontryagin's minimum principle. *Phys. Rev. X*, 7: 021027, May 2017.

[36] N. Y. Yao, L. Jiang, A. V. Gorshkov, Z.-X. Gong, A. Zhai, L.-M. Duan, and M. D. Lukin. Robust quantum state transfer in random unpolarized spin chains. *Phys. Rev. Lett.*, 106:040505, Jan 2011.

[37] X.-P. Zhang, B. Shao, S. Hu, J. Zou, and L.-A. Wu. Optimal control of fast and high-fidelity quantum state transfer in spin-1/2 chains. *Annals of Physics*, 375:435–443, 2016.

[38] L. Zhou, S.-T. Wang, S. Choi, H. Pichler, and M. D. Lukin. Quantum approximate optimization algorithm: Performance, mechanism, and implementation on near-term devices. *arXiv preprint arXiv:1812.01041*, 2018.

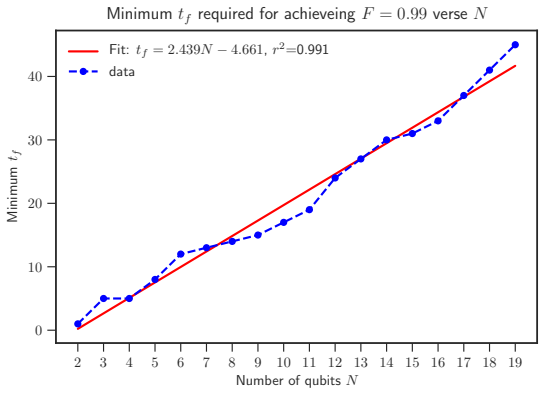


FIG. 11. Minimum required run time t_f for achieving fidelity $F = 0.99$ versus the number of qubits ($N = 2 \to 19$). The Lieb-Robinson bound gives a lower bound of approximately $t_f = 0.03N + c$.

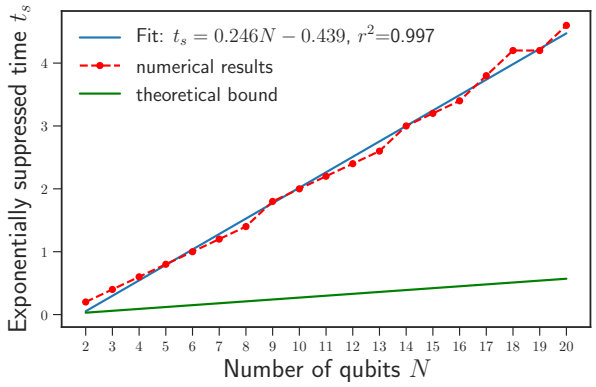
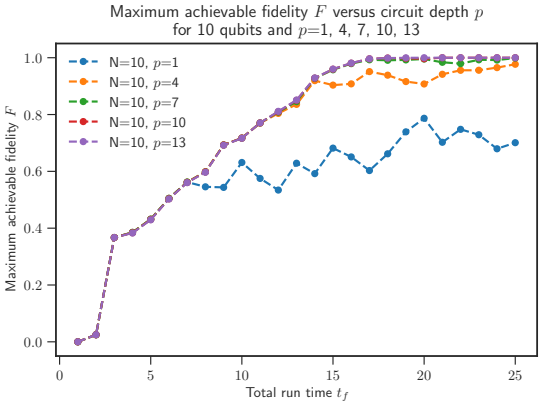
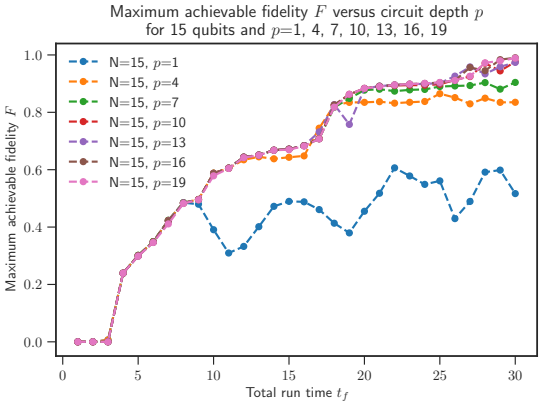


FIG. 13. Exponentially suppressed time t_f (achieved fidelity $F < 0.01$) versus the number of qubits ($N = 2 \to 20$). The Lieb-Robinson bound gives a bound approximately of $t_f \sim 0.03N + c$. The exponentially suppressed time is defined as the time needed to make fidelity higher than 0.01.



(a) $N = 10$ qubits.



(b) $N = 15$ qubits.

FIG. 12. Maximum achievable fidelity F versus physical run time t_f using QAOA as a function of different different circuit depth p. The lines with oscillating behaviors are uncontrollable.

Appendix A: Optimal control solution and Pontryagin's maximum principle

In this appendix, we first solve the dynamical equation of our system in the span of zero and single excitation subspaces. Then we apply optimal control theory [19, 31] to help design our state transfer protocol.

The dynamic of the system is governed by the Schrödinger's equation (with $\hbar = 1$):

$$\frac{d|\psi(t)\rangle}{dt} = -i\hat{H}(t)|\psi(t)\rangle. \tag{A1}$$

Let $|\psi(t)\rangle = \sum_{i=1}^{N} C_i(t)|\bar{i}\rangle$, where $C_i(t)$s are the complex amplitudes of the wave function defined in computational basis. We choose $\vec{c}(t) = \{C_1(t), C_2(t), \cdots, C_N(t)\}$ as the dynamic variables for our problem. Substituting $|\psi(t)\rangle = \sum_{i=1}^{N} C_i(t)|\bar{i}\rangle$ into Eq. (A1), we have

$$\begin{aligned}
\sum_{j=1}^{N} \frac{d}{dt} C_j(t)|\bar{j}\rangle &= -i\left\{s(t)\hat{H}_c + [1-s(t)]\hat{H}_B\right\}\sum_{j=1}^{N} C_j(t)|\bar{j}\rangle \\
&= -i\left\{s(t)|\overline{N}\rangle\langle\overline{N}| + [1-s(t)]\sum_{h=1}^{N-1}[\sigma_h^x\sigma_{h+1}^x + \sigma_h^y\sigma_{h+1}^y]\right\}\sum_{j=1}^{N} C_j(t)|\bar{j}\rangle \\
&= -is(t)C_N(t)|\overline{N}\rangle + \left\{-i[1-s(t)]\sum_{h=1}^{N-1}\left[\sigma_h^x\sigma_{h+1}^x + \sigma_h^y\sigma_{h+1}^y\right]\right\}\sum_{j=1}^{N} C_j(t)|\bar{j}\rangle.
\end{aligned} \tag{A2}$$

Left multiplying both sides by $\langle\bar{j}|$, we get

$$\begin{aligned}
\frac{d}{dt}C_j(t)\langle\bar{j}|\bar{j}\rangle &= -is(t)C_N\langle j|\overline{N}\rangle + \left\{-i[1-s(t)]\langle\bar{j}|\sum_{h=1}^{N-1}[\sigma_h^x\sigma_{h+1}^x + \sigma_h^y\sigma_{h+1}^y]\right\}\sum_{k=1}^{N} C_k(t)|\bar{k}\rangle \\
&= -i\left\{s(t)C_N\langle\bar{j}|\overline{N}\rangle + [1-s(t)]\sum_{k=1}^{N} B_{j,k}C_k\right\},
\end{aligned} \tag{A3}$$

where

$$\begin{aligned}
B_{j,k} &= \langle\bar{j}|\sum_{h=1}^{N-1}[\sigma_h^x\sigma_{h+1}^x + \sigma_h^y\sigma_{h+1}^y]|\bar{k}\rangle \\
&= 2\delta(k-j-1).
\end{aligned} \tag{A4}$$

Then we arrive in the dynamics equation in the form $\dot{\vec{c}}(t) = f(\vec{c}(t), s(t))$.

$$\begin{aligned}
\frac{d}{dt}C_j(t) &= -i\left\{s(t)C_N\langle\bar{j}|\overline{N}\rangle + [1-s(t)]\sum_{k=1}^{N} B_{j,k}C_k(t)\right\} \\
&= -i\left\{s(t)C_N\delta_{jN} + 2[1-s(t)]\sum_{k=1}^{N}\delta(k-j-1)C_k\right\} \\
&= \sum_{k=1}^{N} A_{j,k}C_k,
\end{aligned} \tag{A5}$$

where

$$A_{j,k} = -i\left\{s(t)\delta_{jN} + 2[1-s(t)]\delta(k-j-1)\right\}. \tag{A6}$$

The cost function (action) for our state transfer problem is given by

$$J[\vec{c}(t_f)] = -|C_N(t_f)|^2, \tag{A7}$$

which only depends on the final state. Thus the problem we are solving is of Mayer type [31]. Then the control Hamiltonian is linearly dependent to the conjugate momentum $\vec{p}$ and control dynamics is

$$H_{\text{control}} = \vec{p}^T \cdot f(\vec{c}(t), s(t)) = \vec{p}^T \cdot \mathbf{A} \cdot \vec{c}, \tag{A8}$$

such that the conjugate momentum is determined by the control Hamiltonian in the same way as that in the classical mechanics:

$$\dot{\vec{p}} = -\partial_{\vec{c}} H_{\text{control}}(\vec{c}, \vec{p}, s). \tag{A9}$$

With the fixed total time t_f, the boundary condition for conjugate variable is given by

$$\vec{p}(t_f) = \left. \frac{\partial J(t)}{\partial \vec{c}} \right|_{t=t_f}. \tag{A10}$$

Denote the components of $\vec{p}$ to be $P_i(t)$. $P_i(t)$ should satisfy:

$$\frac{d}{dt} P_j(t) = -is(t) P_N \delta_{jN} + 2[1 - s(t)] P_{j+1}. \tag{A11}$$

Subsequently, the necessary and sufficient conditions for an optimal control $s(t)$ is determined by:

$$\frac{\partial \hat{H}}{\partial s} = 0, \frac{\partial^2 \hat{H}}{\partial s^2} \geq 0. \tag{A12}$$

However, this cannot be applied to linear control problem where $\frac{\partial \hat{H}}{\partial s}$ is not a function of s. Pontryagin's principle comes to rescue, which replace two criteria with three new ones that are necessary and sufficient.

$$H_{\text{control}}(\vec{c}^*, \vec{p}^*, s^*) \leq H_{\text{control}}(\vec{c}^*, \vec{p}^*, s), \forall t \in [0, t_f] \tag{A13}$$
$$\vec{p}(t_f) = \partial_{\vec{c}} J[\vec{c}(t_f)], \tag{A14}$$
$$\partial_t J[\vec{c}] + H_{\text{control}}(\vec{c}^*, \vec{p}^*, s^*)|_{t=t_f} = 0 \tag{A15}$$

Since the control Hamiltonian is a linear function of the control parameter $s(t) (0 \leq s(t) \leq 1)$, $s(t)$ should be maximized when $\partial_s H_{\text{control}} < 0$ and should be minimized when $\partial_s H_{\text{control}} > 0$. The optimal control for QAOA is therefore determined from Pontryagin's theorem as follows:

$$s(t) = \begin{cases} 0 & \frac{\partial \hat{H}_{\text{control}}}{\partial s} > 0 \\ 1 & \frac{\partial \hat{H}_{\text{control}}}{\partial s} < 0 \end{cases} \tag{A16}$$

Then, the best control solution $s(t)$ is of bang-bang form, which corresponds to switching between two constant controls for each time duration. The bang-bang form of control contains abrupt switch between two values of $s(t)$ at time t_0 determined by $\partial_s H_{\text{control}}|_{t=t_0} = 0$, or specifically as

$$\vec{p}^T \cdot \mathbf{F} \cdot \vec{c} = 0, \tag{A17}$$

where the matrix elements of F are given by

$$F_{ij} = \delta_{iN} - 2\delta(j - i - 1). \tag{A18}$$

It is therefore trivial to verify whether a given control $s(t)$ as a function of t is optimal or not. However, finding the 'optimal' control is generally hard due to the mutual dependency of control and system dynamics. A brute force search on switching time is already computationally formidable but is still unable to find the optimal control without specifying the number of bangs.

Appendix B: Details of our numerical optimization

In this appendix, we provide more details on our numerical optimizations. Table II summarizes the parameters t_f and p we performed grid search on. In the first round of grid search (Run 1), we investigate the general performance of optimized QAOAs with different fixed total run time t_f and circuit depth p. To investigate the perfomance of optimized QAOA within or near the exponentially suppressed region, we performed the second round of grid search which scanned more densely spaced total run time over a smaller interval. Table III overviews our numerical results and their implications.

In these tables, we adopt the MATLAB style to represent arrays, for example, a "0.2:0.2:1.6" corresponds to an array of "[0.2 0.4 0.6 0.8 1.0 1.2 1.4 1.6] and a "1:5" corresponds an array of "[1 2 3 4 5]".

TABLE I. Parameters of the first and the second round of grid searches over t_f and p. The first round were used to determine the general performance of optimized QAOAs. The second round were used to investigate the performance of optimized QAOAs near or within the exponentially suppressed region that we have identified.

Runs	N	2	3	4	5	6	7	8	9	10	11
Run 1:	p	1:7	1:8	1:9	1:10	1:11	1:12	1:13	1:14	1:15	1:16
	t_f	1:8	1:10	1:12	1:14	1:16	1:18	1:20	1:22	1:25	1:27
Run 2:	p	1:7	1:8	1:9	1:10	1:11	1:12	1:13	1:14	1:15	1:16
	t_f	0.2:0.2:1.6	0.2:0.2:2.0	0.2:0.2:2.4	0.2:0.2:2.8	0.2:0.2:3.2	0.2:0.2:3.6	0.2:0.2:4.0	0.2:0.2:4.4	0.2:0.2:5.0	0.2:0.2:5.4
	N	12	13	14	15	16	17	18	19	20	
Run 1:	p	1:17	1:18	1:19	1:20	1:2:23	1:2:25	1:2:27	1:2:29	1:2:31	
	t_f	1:30	1:32	1:34	1:36	1:2:39	1:2:41	1:2:43	1:2:45	1:2:47	
Run 2:	p	1:17	1:18	1:19	1:20	1:2:23	1:2:25	1:2:27	1:2:29	1:2:31	
	t_f	0.2:0.2:6.0	0.2:0.2:6.4	0.2:0.2:6.8	0.2:0.2:7.2	0.2:0.4:7.8	0.2:0.4:8.2	0.2:0.4:8.6	0.2:0.4:9.0	0.2:0.4:9.4	

TABLE II. Parameters of the grid search over p with no constraints on t_f. The obtained numerical results are presented and analyzed in Section V B.

N	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
p	1:7	1:8	1:9	1:10	1:11	1:12	1:13	1:14	1:15	1:16	1:17	1:18	1:19	1:20	1:22	1:24	1:26	1:28	1:30

TABLE III. Overview of our numerical results. We consider the performance of optimized QAOAs with unlimited t_f in Section V B and that with fixed t_f in Section V C. In Section V C 1, we study the maximum achievable fidelity F as a function of circuit depth p , and investigated the controllability of QAOAs. In Section V C 2, we study the maximum achievable fidelity F as a function of total run time t_f, and investigate the Lieb-Robinson type quantum speed limit emerged in QAOAs.

	Figures	Max achi. F	Physical run time t_f	Circuit depth p	Number of qubits N	Goal
	Figure 2	Max achi. F	unlimited t_f	(a): p =[1:6]; (b): p =[1:11]	$N = 2 \to 20$	QAOA
Section V B	Figure 4	N/A	unlimited t_f	$p = 15$	N=10	bang-bang sol.
unlimited t_f	Figure 5	z variable	unlimited t_f	$p = 2, 4$	$N = 3$	Landscape
	Figure 6(a)	y variable	$t_f = 6$	x variable: [1:9]	[5,10,15,19]	Controllability
Section V C 1	Figure 6(b)	y variable	$t_f = 13$	x variable: [1:9]	[5,10,15,19]	
	Figure 7(a)	y variable	sufficiently large	x variable: [1:6]	[2,4,6,8]	
Fixed t_f	Figure 7(b)	y variable	sufficiently large	x variable: [1:15]	[10,12,14,16]	
	Figure 8(a)	y variable	x variable: all	$p = 2$	[5,10,15,19]	LR-bound
	Figure 8(b)	y variable	x variable: all	$p = 9$	[5,10,15,19]	
	Figure 10(a)	y variable	x variable: all	sufficiently large	[2,4,6,8]	
Section V C 2	Figure 10(b)	y variable	x variable: all	sufficiently large	[10,12,14,16]	
Fixed p	Figure 12(a)	y variable	x variable: [1:25]	sufficiently large	$N = 10$	
	Figure 12(b)	y variable	x variable: [1:30]	sufficiently large	$N = 15$	
	Figure 11	$F > 0.99$	y variable	sufficiently large	x variable: [2:20]	
	Figure 13	$F < 0.01$	y variable	sufficiently large	x variable: [2:20]	