Accelerating quantum optimal control through iterative gradient-ascent pulse engineering

Yuquan Chen,[1,2,*] Yajie Hao,[3,4,*] Ze Wu,[1,2] Bi-Ying Wang,[5] Ran Liu,[1,2] Yanjun Hou,[1,2] Jiangyu Cui,[5,†] Man-Hong Yung,[3,4,5,6,7,‡] and Xinhua Peng[1,2,8,§]

[1]*CAS Key Laboratory of Microscale Magnetic Resonance and School of Physical Sciences, University of Science and Technology of China, Hefei 230026, China*

[2]*CAS Center for Excellence in Quantum Information and Quantum Physics, University of Science and Technology of China, Hefei 230026, China*

[3]*Department of Physics, Southern University of Science and Technology, Shenzhen 518055, China*

[4]*Shenzhen Institute for Quantum Science and Engineering, Southern University of Science and Technology, Shenzhen 518055, China*

[5]*Central Research Institute, Huawei Technologies, Shenzhen 518129, China*

[6]*Guangdong Provincial Key Laboratory of Quantum Science and Engineering, Southern University of Science and Technology, Shenzhen 518055, China*

[7]*Shenzhen Key Laboratory of Quantum Science and Engineering, Southern University of Science and Technology, Shenzhen 518055, China*

[8]*Hefei National Laboratory, University of Science and Technology of China, Hefei 230088, China*

(Received 28 February 2023; revised 22 September 2023; accepted 27 September 2023; published 7 November 2023)

Quantum optimal control, a powerful toolbox for engineering an optimal control field modulation that most precisely implements a desired quantum operation in the best way possible, has evolved into one of the cornerstones for enabling quantum technologies. The gradient ascent pulse engineering (GRAPE) algorithm is a widely used method in quantum optimal control, which has achieved great success in different physical platforms. However, its computational complexity increases exponentially with the number of qubits, making it challenging to be implemented for large-scale quantum systems. To mitigate this issue, we present the iterative GRAPE algorithm (iGRAPE), which reduces the optimization problem into a series of lower-dimensional subproblems by incorporating disentanglement operations. Our numerical simulations on physical platforms such as nuclear magnetic resonance and superconducting quantum systems demonstrate that iGRAPE significantly enhances state preparation speed. Specifically, compared to GRAPE, iGRAPE achieves up to a five-fold acceleration in preparing Greenberger–Horne–Zeilinger states using a 12-qubit implementation, and up to a 13-fold acceleration for arbitrary state preparation with eight qubits. To further validate our findings, we conduct experimental validation of iGRAPE on a four-qubit nuclear magnetic resonance system. Overall, iGRAPE offers an efficient solution for implementing optimal control in large-scale quantum systems, holding great potential for advancing quantum technologies during the noisy intermediate-scale quantum era.

DOI: 10.1103/PhysRevA.108.052603

I. INTRODUCTION

In the past few decades, quantum optimal control (QOC) theory [1–5] has been well developed and has stimulated a lot of interest in the field of quantum technology. Specifically, this theory focused on achieving the optimal implementation of a target quantum state or desired quantum operation in quantum simulation and sensing as well as scalable quantum computation devices [6–12]. To realize this goal in practice, an efficient optimization algorithm is essential. So far, various numerical optimization algorithms have been developed including gradient-based methods such as gradient ascent pulse engineering (GRAPE) [13], stochastic gradient Descent [14,15], Krotov's algorithm [16], reinforcement learning algorithms and their variants [17–23], as well as nongradient-based methods such as the chopped random basis [24,25] and Nelder-Mead approaches [26].

The GRAPE algorithm has attracted much attention among those algorithms. Since it utilizes a direct analytical expression for the gradient, GRAPE can efficiently find a suitable solution in the parameter space with fast convergence speed. Moreover, in combination with the advanced optimizers (such as BFGS, AdaGrad, and Adam), as well as the automatic differentiation technique implemented on GPU [27], many variants of GRAPE were proposed to improve its. performance [28,29]. The GRAPE algorithm was originally developed in nuclear magnetic resonance (NMR) systems [13], which has now been widely applied to many other quantum platforms [30–34], such as superconducting quantum circuits [34], circuit QED [31], trapped ions [35], and nitrogen-vacancy (NV) centers [36]. However, this technique relied on dynamic simulations of the quantum systems on classical computers, which essentially limited its application on large quantum systems.

*These authors contribute equally in this work.
†cuijiangyu@huawei.com
‡yung@sustech.edu.cn
§xhpeng@ustc.edu.cn

In response to the pressing need to alleviate this issue, we introduce an algorithm named iterative gradient ascent pulse engineering (iGRAPE). Unlike its predecessor, GRAPE, which focuses on the dynamics of the entire system, the iGRAPE algorithm revolutionizes the approach by leveraging the inverse evolution from the target state to the initial state. By decomposing the optimization problem into low-dimensional components using disentanglement partition systems, the iGRAPE algorithm remarkably reduces the computational resources required for optimization. To demonstrate the effectiveness of the iGRAPE algorithm, we apply it to two widely recognized physical platforms, namely, NMR and superconducting quantum systems, to realize the desired states. Our findings showcase the remarkable potential of iGRAPE, revealing that its implementation with 12 qubits can accelerate the Greenberger–Horne–Zeilinger (GHZ)-state preparation process by up to fivefold. Furthermore, in the task of arbitrary state preparation using eight qubits, iGRAPE outperforms GRAPE by achieving a remarkable 13-fold acceleration. In an effort to validate our findings experimentally, we successfully prepare the GHZ state using the iGRAPE method on a four-qubit NMR platform, achieving an impressive experimental fidelity of 98.25%. These results not only demonstrate the practical feasibility of the iGRAPE algorithm, but also highlight its superiority over GRAPE for large-scale quantum systems. Consequently, iGRAPE presents a promising and robust solution for QOC in noisy intermediate-scale quantum systems.

II. iGRAPE ALGORITHM

A quantum system generally can be described by the Hamiltonian

$$H(t) = H_{\rm s}(t) + H_{\rm c}(t), \tag{1}$$

where the system Hamiltonian is

$$H_{\rm s}(t) = H_{\rm L} + \sum_{i,j} g_{ij}(t) H_{ij}, \tag{2}$$

and the control-field Hamiltonian reads

$$H_{\rm c}(t) = \sum_{\alpha} u_{\alpha}(t) H_{\alpha}. \tag{3}$$

Here H_{ij} denotes the coupling between qubits i and j with the coupling strength $g_{ij}(t)$ and $H_{\rm L}$ represents the local terms of the system Hamiltonian; H_{α} represents the control-field Hamiltonian with the amplitude $u_{\alpha}(t)$.

The goal of GRAPE for state preparation is to design a suitable set of $u_{\alpha}(t)$ control pulses to transfer an initial state $|\psi(0)\rangle$, such as $|\psi(0)\rangle = |0\cdots 0\rangle = |\mathbb{0}\rangle$, to a given target state $|\Psi\rangle$ in a specified time T. We discretize the entire time evolution linearly into K segments, i.e., $\Delta t = T/K$. For the kth segment, the induced temporal-evolution propagator $\mathcal{U}_k$ is

$$\mathcal{U}_k(\Delta t) = \exp\left\{-i\Delta t\left(H_{\rm s} + \sum_{\alpha} u_{\alpha}(k) H_{\alpha}\right)\right\}. \tag{4}$$

Here, we assume the system Hamiltonian $H_{\rm s}$ and the control parameter within the kth segment $u_{\alpha}(k)$ are time independent. Consequently, the unitary evolution of the GRAPE algorithm can be written as $U = \Pi_{k=1}^{K} \mathcal{U}_k(\Delta t)$, which is optimized to transfer $|\psi(0)\rangle$ to $|\Psi\rangle$.

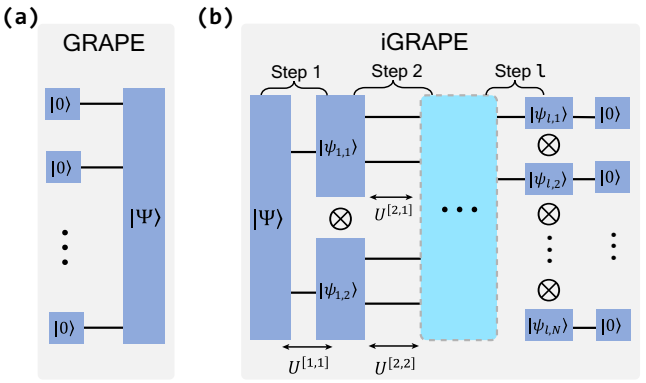

FIG. 1. Scheme of the GRAPE and iGRAPE algorithm for state preparation problems when the initial state is $|\mathbb{0}\rangle$. (a) The GRAPE algorithm designs the control field from the initial state $|\mathbb{0}\rangle$ to the target state $|\Psi\rangle$. (b) The iGRAPE algorithm calculates the reverse evolution and designs the control field from the target state $|\Psi\rangle$ to the initial state $|\mathbb{0}\rangle$.

A. Algorithm process

The central concept behind the iGRAPE algorithm is to leverage the inverse evolution process to convert the target state $|\Psi\rangle$ to the initial state $|\psi(0)\rangle$. This is achieved by breaking down $|\Psi\rangle$ into a sequence of product states in subsystems until the final product state $|\psi(0)\rangle$ is obtained. As shown in Fig. 1(b), the operator $U^{[1,1]}$ optimized in algorithm *Step* 1 transfers the target state $|\Psi\rangle$ to the product of the states of its two subsystems $|\psi_{1,1}\rangle$ and $|\psi_{1,2}\rangle$; then in *Step* 2, the optimized operator $U^{[2,1]}$ further transfers $|\psi_{1,1}\rangle$ to the product of the states of its two subsystems $|\psi_{2,1}\rangle \otimes |\psi_{2,2}\rangle$, and likely $U^{[2,2]}$ transfers $|\psi_{1,2}\rangle$ to $|\psi_{2,3}\rangle \otimes |\psi_{2,4}\rangle$; until in *Step* l, all the subsystem states become single-qubit states and can be finally transferred to $|\mathbb{0}\rangle$ by applying single-qubit rotations. Each subsystem state $|\psi_{n,m}\rangle$ can be an arbitrary state in the corresponding Hilbert space. Let n labels the *Step* index, and m labels the subsystem index in the *Step*. The entire propagator U can be written as

$$U = \left(\bigotimes_{m} U^{[l+1,m]}\right) \cdots \left(\bigotimes_{m=1}^{2} U^{[2,m]}\right) \cdot U^{[1,1]}, \tag{5}$$

where the target $U|\Psi\rangle = |\mathbb{0}\rangle$, thus $|\Psi\rangle = U^{\dagger}|\mathbb{0}\rangle$.

The numerical optimization is applied to each step in engineering the pulse sequence to implement the operation $U^{[n,m]}$. Similar to GRAPE, we discretize linearly the entire time evolution of $U^{[n,m]}$ into K_n segments. The kth temporal-evolution propagator $\mathcal{U}_k^{[n,m]}$ is

$$\mathcal{U}_k^{[n,m]} = \exp\left\{i\Delta t\left(H_{\rm s} + \sum_{\alpha} u_{\alpha}^{[n,m]}(k) H_{\alpha}\right)\right\}. \tag{6}$$

Hence, $U^{[n,m]} = \Pi_{k=1}^{K_n} \mathcal{U}_k^{[n,m]}$ and $u_{\alpha}^{[n,m]}(k)$ are the optimized control parameters. Note that $\mathcal{U}_k^{[n,m]}$ is different from Eq. (4). Here we set $-\Delta t$ to Δt to guarantee the physical implementation of $U^{\dagger}$ in the final pulse sequence. Consequently, $U^{[n,m]\dagger}$

is directly realized by reversing the pulse sequence engineered for $U^{[n,m]}$. Via the operation $U^{[n,m]}$, the system is disentangled into two subsystems

$$U^{[n,m]}|\psi_{n-1,m}\rangle = |\psi_{n,2m-1}\rangle \otimes |\psi_{n,2m}\rangle. \tag{7}$$

To further drive each subsystem independently in the following *Steps*, one needs to turn off all couplings between these subsystems at the end of each *Step*. By labeling the two subsystems in Eq. (7) as subsystem A and subsystem B and setting the cost function as $L_t = 1 - \mathrm{tr}(\rho^A \rho^A)$, with $\rho^A = \mathrm{tr}_B(U^{[n,m]}|\psi_{n-1,m}\rangle\langle\psi_{n-1,m}|U^{[n,m]\dagger})$, one can obtain a pure state ρ^A by minimizing L_t. The parameters $u_\alpha(k)$ are optimized through gradient descent.

B. Gradient for the cost function

To find the specific set $u_\alpha^{[n,m]}(k)$, the previous cost function L_t can be written as

$$\begin{aligned} L_t &= 1 - \mathrm{tr}\big[\mathrm{tr}_B^2(|\phi_{n,m}\rangle\langle\phi_{n,m}|)\big], \\ |\phi_{n,m}\rangle &= U^{[n,m]}|\psi_{n-1,m}\rangle, \end{aligned} \tag{8}$$

minimizing L_t such that $\rho^{\mathrm{A}} = \mathrm{tr}_B(|\phi_{n,m}\rangle\langle\phi_{n,m}|)$ becomes the pure state. The parameters $u_\alpha^{[n,m]}(k)$ could be adjusted by the updated rule

$$u_\alpha^{[n,m]}(k) \leftarrow u_\alpha^{[n,m]}(k) - \omega \frac{\partial L_t}{\partial u_\alpha^{[n,m]}(k)}. \tag{9}$$

Expand ρ^{A} such that

$$\rho^{\mathrm{A}} = \sum_j (\mathbb{1}_{\mathrm{A}} \otimes \langle j|_{\mathrm{B}})|\phi_{n,m}\rangle\langle\phi_{n,m}|(\mathbb{1}_{\mathrm{A}} \otimes |j\rangle_{\mathrm{B}}), \tag{10}$$

and the basis of subsystems A and B are denoted as $|i\rangle_{\mathrm{A}}$ and $|j\rangle_{\mathrm{B}}$. The gradient can be written as

$$\begin{aligned} \frac{\partial L_t}{\partial u_\alpha^{[n,m]}(k)} &= 4\mathrm{Re}\big(\langle\lambda_{n,m}|i\Delta t \mathcal{U}_{\mathrm{left}}^{[n,m]} H_\alpha \mathcal{U}_{\mathrm{right}}^{[n,m]}|\psi_{n-1,m}\rangle\big), \\ \mathcal{U}_{\mathrm{left}}^{[n,m]} &= \mathcal{U}_{K_n}^{[n,m]} \ldots \mathcal{U}_{k+1}^{[n,m]}, \\ \mathcal{U}_{\mathrm{right}}^{[n,m]} &= \mathcal{U}_{k}^{[n,m]} \ldots \mathcal{U}_{1}^{[n,m]}, \end{aligned} \tag{11}$$

where the elements in vector $|\lambda_{n,m}\rangle$ is given as

$$\begin{aligned} (\langle i|_{\mathrm{A}}\langle j|_{\mathrm{B}})|\lambda_{n,m}\rangle = \sum_{i',j'} &\langle\phi_{n,m}|(|i'\rangle_{\mathrm{A}}|j'\rangle_{\mathrm{B}}) \\ &\times (\langle i|_{\mathrm{A}}\langle j'|_{\mathrm{B}})|\phi_{n,m}\rangle(\langle i'|_{\mathrm{A}}\langle j|_{\mathrm{B}})|\phi_{n,m}\rangle. \end{aligned} \tag{12}$$

The iGRAPE method also relies on dividing the system into different components at each *Step* that can be selected based on the desired target state and the Hamiltonian of the system. In this study, we will evaluate the performance of both the GRAPE and iGRAPE algorithms by preparing target states on two physical systems: superconducting and NMR quantum computing systems.

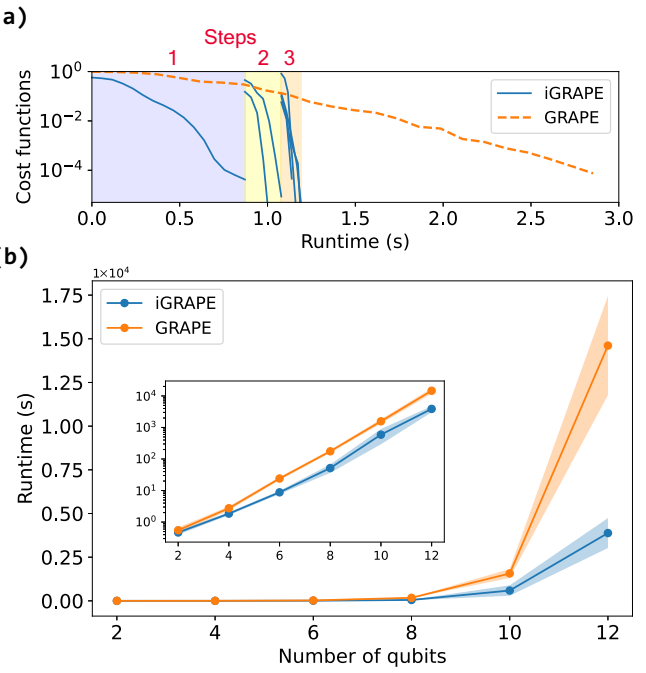

FIG. 2. Benchmark of the iGRAPE and the GRAPE algorithms for GHZ-state preparation on superconducting quantum systems. (a) Variation of cost functions during the optimization process of the GRAPE and iGRAPE algorithms on a four-qubit case. (b) Runtime versus number of qubits. Each point is an average over 20 random selections of a set of initial control parameters for different numbers of qubits. The colored band-region accounts for the statistical distribution. The inset is a semi-log plot.

III. BENCHMARK ON SUPERCONDUCTING QUANTUM SYSTEMS WITH TUNABLE COUPLINGS

Our benchmark of the iGRAPE algorithm is first performed on a one-dimensional (1D) 12-superconducting-qubit chain, where the system Hamiltonian H_{SC} and the corresponding control field Hamiltonian H_{c} can be described as [37]

$$\begin{aligned} H_{\mathrm{SC}} &= \sum_j \Big(\omega_j \hat{n}_j + \frac{\eta_j}{2}\hat{n}_j(\hat{n}_j - 1)\Big) \\ &\quad + \sum_j [g_j(t)(\hat{a}_j^\dagger \hat{a}_{j+1} + \hat{a}_j \hat{a}_{j+1}^\dagger)], \\ H_{\mathrm{c}}(t) &= \sum_j [u_{xj}(t)(\hat{a}_j + \hat{a}_j^\dagger) + u_{yj}(t) i(\hat{a}_j - \hat{a}_j^\dagger)], \end{aligned} \tag{13}$$

where $\hat{n}$ is the number operator, $\hat{a}_j^\dagger$ ($\hat{a}_j$) is the creation (annihilation) operator, ω_j and η_j are the transition frequency and the anharmonicity of the jth qubit, respectively, and $g_j(t)$ denotes the interaction strength between the jth and $(j+1)$th qubits. Each qubit can be fully controlled by individual capacitively coupled microwave control lines (XY), and u_{xj}, u_{yj} are the amplitudes of the control fields.

Figure 2(a) shows the varying of cost functions during the optimization process of iGRAPE and GRAPE algorithms on a four-qubit case for a GHZ-state preparation, which consists

of three *Steps*:

$$\begin{aligned}
&\text{Step 1: } |\text{GHZ}\rangle \xrightarrow{U^{[1,1]}} |\psi_{1,1}\rangle \otimes |\psi_{1,2}\rangle,\\
&\text{Step 2: } \begin{cases} |\psi_{1,1}\rangle & \xrightarrow{U^{[2,1]}} |\psi_{2,1}\rangle \otimes |\psi_{2,2}\rangle,\\ |\psi_{1,2}\rangle & \xrightarrow{U^{[2,2]}} |\psi_{2,3}\rangle \otimes |\psi_{2,4}\rangle, \end{cases}\\
&\text{Step 3: } |\psi_{2,i}\rangle \xrightarrow{U^{[3,i]}} |0\rangle,\ \imath \in [1,4].
\end{aligned} \tag{14}$$

Here $|\psi_{1,1}\rangle$ and $|\psi_{1,2}\rangle$ are two-qubit states and $|\psi_{2,i}\rangle (i = 1,\ldots,4)$ are single-qubit states. $U^{[1,1]}$, $U^{[2,i]} (i = 1, 2)$, and $U^{[3,i]}$, $(i = 1,\ldots,4)$ are, respectively, 16×16, 4×4, and 2×2 unitary operations. Although the cost function in Eq. (8) for each training curve is different, they all fall within the range of [0,1]. It can be seen that the iGRAPE algorithm converges quickly in each *Step* (denoted by the blue solid lines), while the GRAPE algorithm takes longer to reach convergence (denoted by the red dashed line). For *Step* 2 and *Step* 3, the optimization tasks in each *Step* can be simultaneously optimized in lower-dimensional Hilbert spaces. Note that, on superconducting systems, the coupling parameters $g_j(t)$ can be controlled off and on; thus in *Steps* 2 and 3, the couplings between the different subsystems are turned off.

Using the L-BFGS-B optimization algorithm [38], Fig. 2(b) shows that the relationship between the running time of the iGRAPE algorithm and the qubit number on the 1D 12-qubit superconducting-qubit chain, in contrast to the case of the GRAPE algorithm. It can be seen from Fig. 2(b) that the execution times of iGRAPE (in blue dots) are less than those of GRAPE (in red dots) and the advantage is more significant as the qubit number grows. We observe a five-fold iGRAPE speedup for the 12-qubit case. We also show the semi-log plot in the inset of Fig. 2(b). Although the runtime of iGRAPE still grows exponentially with the system size, the exponential index is reduced to 0.398 from 0.446 in the case of GRAPE.

IV. BENCHMARK ON NMR QUANTUM SYSTEMS WITH ALWAYS-ON COUPLINGS

For NMR quantum systems, the system Hamiltonian and the control-field Hamiltonian can be, respectively, written as

$$\begin{aligned}
H_{\text{NMR}} &= \sum_j \left(\pi \nu_j \hat{\sigma}_z^{(j)}\right) + \sum_{i<j} \left(\frac{\pi}{2} J_{ij} \hat{\sigma}_z^{(i)} \hat{\sigma}_z^{(j)}\right),\\
H_{\text{c}}(t) &= \sum_j \left(\pi u_{xj}(t) \hat{\sigma}_x^{(j)} + \pi u_{yj}(t) \hat{\sigma}_y^{(j)}\right).
\end{aligned} \tag{15}$$

Here H_{NMR} is time independent, ν_j represents the chemical shift of the jth spin, J_{ij} is the scalar coupling strength between two spins, and u_{xj} and u_{yj} denote, respectively, the control radio-frequency (rf) fields along the x and y directions. In the Appendix, we provide a table showing the values of ν_i and J_{ij} for some NMR samples. Unlike the superconducting quantum systems mentioned earlier, these coupling terms are always present in NMR systems. Therefore, they must be taken into consideration during each *Step* of the optimization process.

A. Adaptation of iGRAPE

To avoid the evolution under the couplings between two subspaces in the optimization of the following *Steps*, we find that the evolution between two subsystems can be frozen when one of these two subsystems is in the state $|\mathbb{0}\rangle$:

$$e^{-iHt}|\mathbb{0}\rangle_{\text{A}}|\psi\rangle_{\text{B}} = e^{i\theta}|\mathbb{0}\rangle_{\text{A}} e^{-iH_{\text{B}}t}|\psi\rangle_{\text{B}}, \tag{16}$$

where H denotes the whole-system Hamiltonian and H_{B} denotes the Hamiltonian for the Ssubsystem B

$$\begin{aligned}
H_{\text{B}} = \frac{\pi}{2}\left[\sum_i \left(2\nu_i + \sum_{p<i} J_{pi}\right)\hat{\sigma}_z^{(i)} + \sum_{i<j} J_{ij}\hat{\sigma}_z^{(i)}\hat{\sigma}_z^{(j)}\right]\\
+ H_{\text{B, control}},
\end{aligned} \tag{17}$$

and $e^{i\theta}$ is a global phase. When the state of subsystem A reaches $|\mathbb{0}\rangle_{\text{A}}$, the coupling Hamiltonian $J_{ij}\hat{\sigma}_z^{(i)}\hat{\sigma}_z^{(j)}$ will not change the state of the subsystem A except for a global phase. In this situation, the cost function L_f can be expressed as

$$L_f = 1 - \text{tr}(\rho_0 \rho^{\text{A}}), \tag{18}$$

where ρ_0 denotes the state $|\mathbb{0}\rangle\langle\mathbb{0}|_{\text{A}}$ in subsystem A. The gradient of L_f is

$$\begin{aligned}
\frac{\partial L_f}{\partial u_\alpha(k)} &= 2\text{Re}\left(\langle\eta_{n,m}|i\Delta t \mathcal{U}_{\text{left}}^{[n,m]} H_\alpha \mathcal{U}_{\text{right}}^{[n,m]}|\psi_{n-1,m}\rangle\right),\\
\mathcal{U}_{\text{left}}^{[n,m]} &= \mathcal{U}_{K_n}^{[n,m]} \ldots \mathcal{U}_{k+1}^{[n,m]},\\
\mathcal{U}_{\text{right}}^{[n,m]} &= \mathcal{U}_{k}^{[n,m]} \ldots \mathcal{U}_{1}^{[n,m]},
\end{aligned} \tag{19}$$

where the elements in vector $|\eta_{n,m}\rangle$ is defined as

$$(\langle\mathbb{0}|_{\text{A}}\langle j|_{\text{B}})|\eta_{n,m}\rangle = (\langle\mathbb{0}|_{\text{A}}\langle j|_{\text{B}})|\phi_{n,m}\rangle. \tag{20}$$

Therefore, we can set $|\psi_{n,2m}\rangle$ in Eq. (7) as $|\mathbb{0}\rangle$, as shown in Fig. 3(a)

$$U^{[n,m]}|\psi_{n-1,m}\rangle = |\psi_{n,2m-1}\rangle \otimes |\mathbb{0}\rangle. \tag{21}$$

In the following *Step*, the control fields to be optimized are only performed on the $|\psi_{n,2m-1}\rangle$ state. For instance, if we consider the NMR sample Crotonic acid which can be regarded as a seven-qubit quantum simulator [39], we divide the system into two subsystems: four ^{13}C spins and three ^{1}H spins. The varying of cost functions in the optimization processes for a GHZ-state preparation is shown in Fig. 3(b), and the algorithm *Steps* are designed as

$$\begin{aligned}
&\text{Step 1: } |\text{GHZ}\rangle \xrightarrow{U^{[1,1]}} |\psi_{1,1}\rangle_{\text{C}}|\mathbb{0}\rangle_{\text{H}},\\
&\text{Step 2: } |\psi_{1,1}\rangle_{\text{C}} \xrightarrow{U^{[2,1]}} |\mathbb{0}\rangle_{\text{C}}.
\end{aligned} \tag{22}$$

A similar result was observed as the case in Fig. 2(a).

We investigate the scalability of the iGRAPE algorithm for GHZ-state preparation, similar to superconducting systems. The benchmark results are presented in Fig. 3(c), indicating that the iGRAPE algorithm (represented by blue dots) outperforms the traditional GRAPE algorithm (represented by red dots). Notably, this advantage is more pronounced in larger quantum systems; for instance, a four-fold improvement over GRAPE is achieved in the seven-qubit case. Additionally, we observe that the exponential index in iGRAPE has been reduced to 0.349 from 0.449 in GRAPE as shown in the inset.

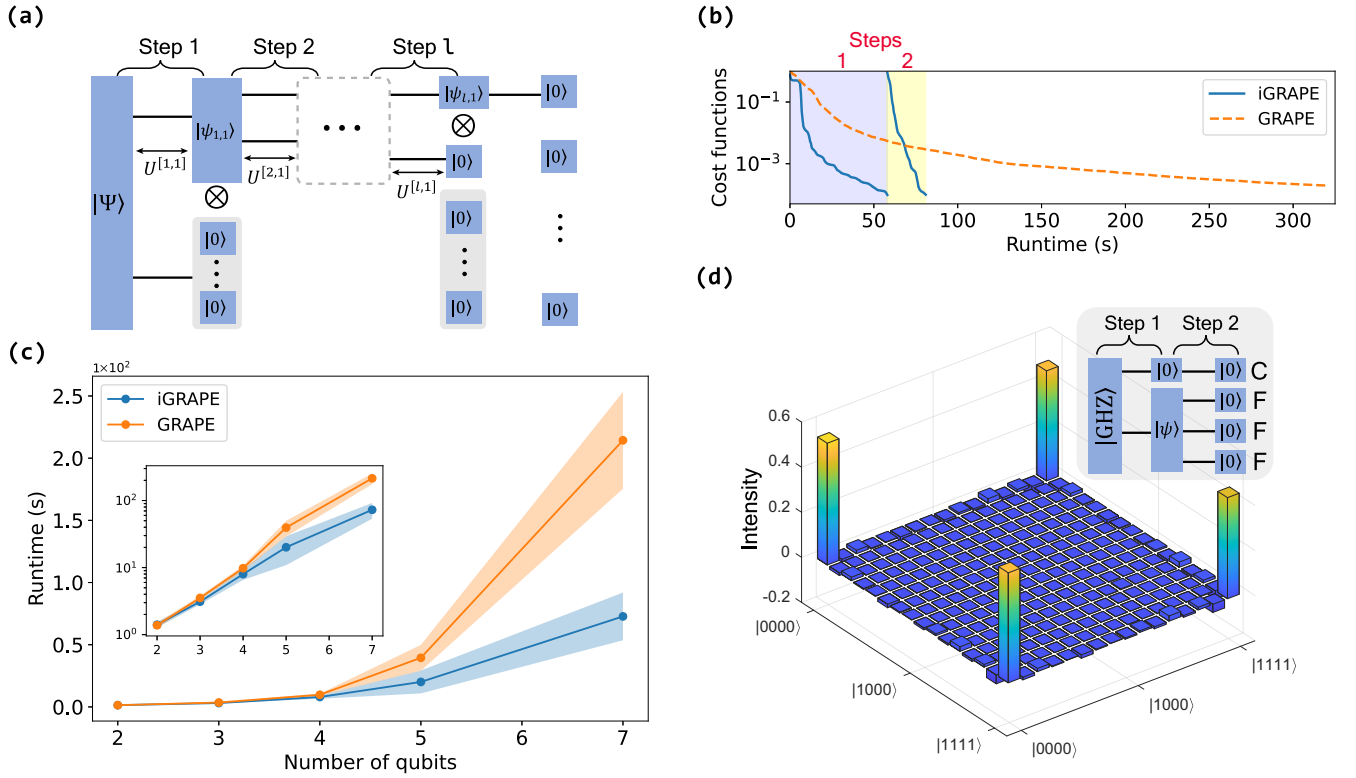

FIG. 3. Benchmark of the iGRAPE and the GRAPE algorithms for GHZ-state preparation on NMR quantum systems. (a) Revised scheme of iGRAPE for the coupling always-on NMR systems. (b) Evolution of the cost functions during the running of the GRAPE and iGRAPE algorithms for the seven-qubit NMR quantum system. (c) Runtime versus number of qubits. Each point in the plot represents the average result obtained from 20 random selections of initial control parameters on NMR systems with different numbers of qubits. See the Appendix for details. The colored band-region accounts for the statistical distribution. The inset is a semi-log plot. (d) Experimentally reconstructed density matrix (real part) for the prepared GHZ state by the iGRAPE algorithm on the four-qubit NMR system (^{13}C-iodotriuroethylene, see the Appendix). The imaginary part is almost zero (<0.06).

B. Experimental verification

To verify the feasibility of the iGRAPE algorithm in our experiments, we employ ^{13}C-iodotriuroethylene dissolved in d-chloroform as a four-qubit quantum simulator. The sample comprises one ^{13}C and three ^{19}F nuclear spins (see the Appendix for details) [40,41]. Experiments are performed on a Bruker Avance III 400 MHz spectrometer at room temperature. The system is initially at a thermal equilibrium state and first prepared to a pseudopure state (PPS) $\hat{\rho}_{\rm pps} = [(1-\epsilon)/16]\mathbb{1} + \epsilon|\mathbb{0}\rangle\langle\mathbb{0}|$ using the selective-transition approach [42] with polarization $\epsilon \approx 10^{-5}$. The experimental fidelity of $\hat{\rho}_{\rm pps}$ is about 99.29% (see the Appendix for details).

According to the iGRAPE method, the algorithm process for the experiment shown in Fig. 3(d) contains two *Steps*:

$$\text{Step 1: } |\text{GHZ}\rangle \xrightarrow{U^{[1]}} |0\rangle_{\rm C}|\psi\rangle_{\rm F},$$
$$\text{Step 2: } |\psi\rangle_{\rm F} \xrightarrow{U^{[2]}} |\mathbb{0}\rangle_{\rm F}. \tag{23}$$

We first transfer the target GHZ state into a product state $|0\rangle_{\rm C}|\psi\rangle_{\rm F}$ by designing a pulse sequence in *Step* 1. Here, $|\psi\rangle_{\rm F}$ is a three-qubit state for three ^{19}F spins. In *Step* 2, we search for a pulse sequence to transfer three ^{19}F spins from $|\psi\rangle_{\rm F}$ into the state $|\mathbb{0}\rangle$. The fidelity of a pure state $|\psi\rangle$ and a density matrix ρ is defined as $F(|\psi\rangle, \rho) = \langle\psi|\rho|\psi\rangle$, which is used to calculate the final state fidelity of the experiment. Subsequently, the final pulse sequence is obtained by reversing the complete pulse sequence and applied to PPS. The resulting state is subjected to state tomography, which produces a GHZ state fidelity of 98.25%, as depicted in Fig. 3(d). In comparison, the pulse sequences produced by the GRAPE algorithm attain a fidelity of 97.82% in this experiment.

V. BENCHMARK FOR ARBITRARY QUANTUM STATES

To further demonstrate the effectiveness of the iGRAPE algorithm, we will examine its performance in preparing arbitrary quantum states. We employ parameterized quantum circuits (PQC) to generate a range of target states that need to be prepared. As each *Step* in the iGRAPE algorithm aims to transform the target state into subsystem products, it is crucial to consider the entanglement between these subsystems, particularly during the first *Step*. By utilizing von Neumann entropy as a measure of entanglement [43], we can evaluate how well the iGRAPE algorithm performs for different levels of entanglement in the target states.

Figures 4(a) and 4(b) illustrate the PQC used to generate arbitrary states with eight qubits, along with the corresponding performance of the algorithm in a superconducting quantum system. Figures 4(c) and 4(d) depict the PQC employed for

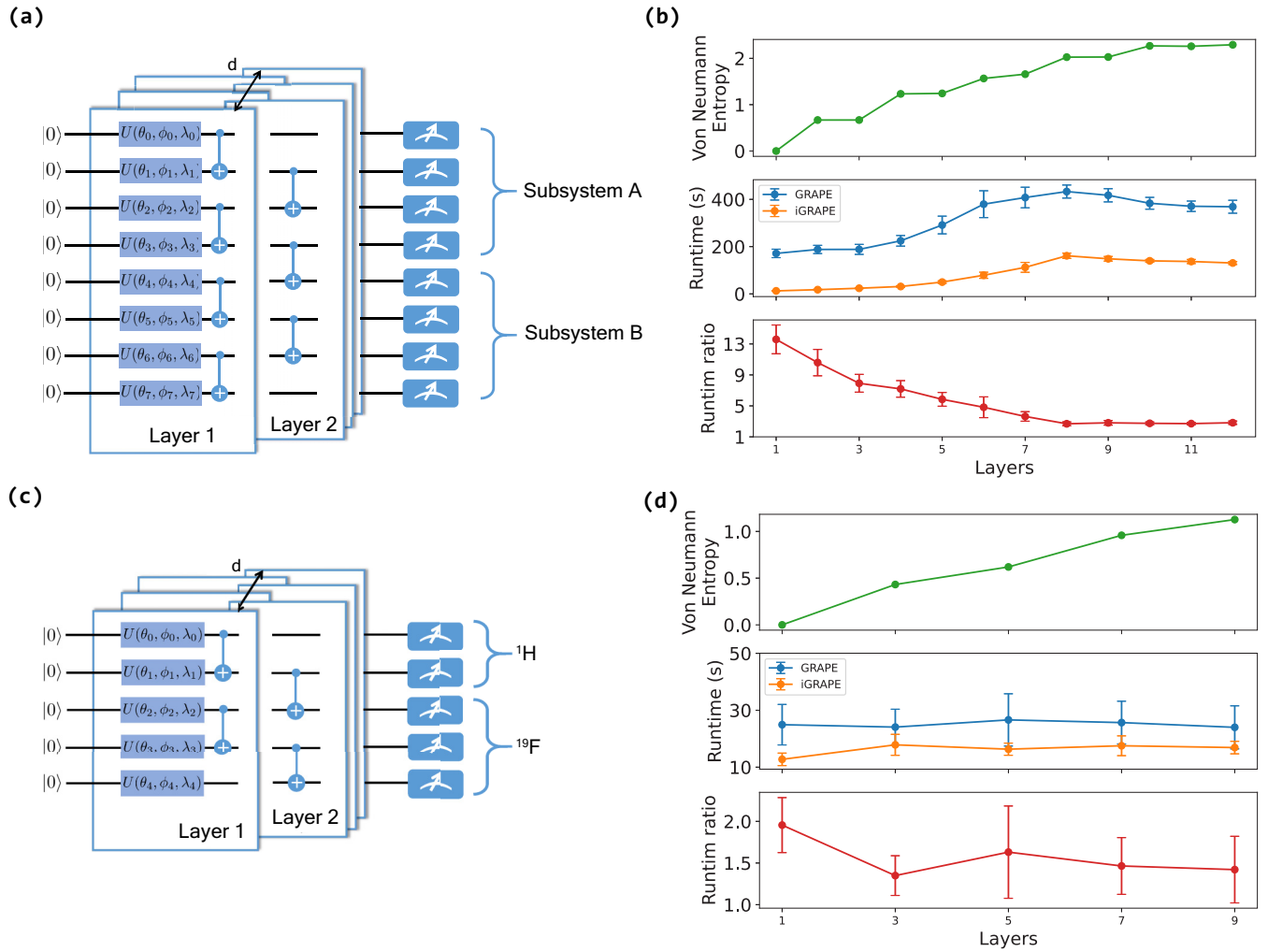

FIG. 4. Benchmark of the iGRAPE and the GRAPE algorithms for arbitrary quantum states preparation. Target states are generated by parameterized quantum circuits (PQC). (a) PQC that generates arbitrary eight-qubit quantum states for the superconducting quantum system, where d represents the number of layers. (b) The von Neumann entropy, the runtime for both algorithms and the runtime ratios (GRAPE/iGRAPE) versus different layers of the circuit. Twenty sets of different random parameters are generated for each circuit, with each point and its error bar representing the mean value and corresponding standard deviation, respectively. (c) PQC that generates arbitrary five-qubit quantum states for the NMR quantum system. (d) The corresponding von Neumann entropy, runtime, and the runtime ratios versus different layers of circuit.

generating arbitrary quantum states with five qubits, as well as the algorithm's performance on an NMR quantum platform. The general U gate in Fig. 4 has the form

$$U(\theta,\phi,\lambda)=\begin{pmatrix}\cos(\frac{\theta}{2}) & -e^{i\lambda}\sin(\frac{\theta}{2})\\ -e^{i\phi}\sin(\frac{\theta}{2}) & -e^{i\lambda+i\phi}\cos(\frac{\theta}{2})\end{pmatrix}, \tag{24}$$

where $\theta\in[0,\pi]$, $\phi\in[0,2\pi]$, $\lambda\in[0,2\pi]$, and $\{\theta_i,\phi_i,\lambda_i\}$ are randomly chosen in the range. An arbitrary pure state $|\psi\rangle$ of a composite system AB can be described by its Schmidt decomposition

$$|\psi\rangle=\sum_i \alpha_i|i_A\rangle|i_B\rangle, \tag{25}$$

where $|i_A\rangle$ and $|i_B\rangle$ are orthonormal states for subsystems A and B, respectively. For pure states, the von Neumann entropy S of the reduced states ρ_A and ρ_B is a well-defined measure of entanglement

$$S(\rho_A)=S(\rho_B)=-\sum_i |\alpha_i|^2\log(|\alpha_i|^2), \tag{26}$$

and this is zero if and only if $|\psi\rangle$ is a product state (not entangled). In Fig. 4(b), a significant improvement in speed is observed for the iGRAPE algorithm, ranging from 2.7 to 13.6 times faster, when preparing eight-qubit arbitrary quantum states of a superconducting quantum system. For the 1-bromo-2,4,5-trifluorobenz sample, we divide the system consisting of five spins [shown in Fig. 4(c)] into two subsystems: Subsystem A composed of two ^{1}H spins and subsystem B composed of three ^{19}F spins. Figure 4(d) reveals that iGRAPE achieves a speedup ranging from 1.5 to 2 times faster for preparing arbitrary quantum states of a NMR quantum system with five qubits. As the number of layers increases, the von Neumann entropy of subsystems in the final state

Algorithm 1. The GRAPE algorithm.

Input: Target state $|\Psi\rangle$
Set tolerance ϵ_0
Output: Control parameters $u_\alpha(k)$
Initialization: guess initial controls $u_\alpha(k)$
Evolution: $|\psi(T)\rangle = U_K ... U_1|\mathbb{0}\rangle$
Cost function: $L = 1 - |\langle\Psi||\psi(T)\rangle|^2$
while $L \geqslant \epsilon_0$
  Calculate gradient $\frac{\partial L}{\partial u_\alpha(k)}$
  Update: $u_\alpha(k) \leftarrow u_\alpha(k) - \omega \frac{\partial L}{\partial u_\alpha(k)}$
  Evolution: $|\psi(T)\rangle = U_K \ldots U_1|\mathbb{0}\rangle$
  Cost function: $L = 1 - |\langle\Psi||\psi(T)\rangle|^2$
end while

escalates. It is noteworthy that both algorithms exhibit similar fluctuation patterns when generating quantum states under different circuit depths. The runtime advantage provided by the iGRAPE algorithm tends to remain stable as circuit depth increases.

VI. CONCLUSION AND DISCUSSION

In conclusion, we introduce the iGRAPE technique for addressing state-transfer challenges in QOC. Our research reveals that the iGRAPE algorithm surpasses the traditional GRAPE method in terms of resource efficiency. By harnessing advanced optimization techniques and pioneering parallel computing strategies, iGRAPE significantly reduces resource usage while ensuring high computational precision. This is substantiated by numerical simulations conducted on state preparations and further validated through experiments carried out on a four-qubit NMR quantum processor.

There are several factors that contribute to the superiority of the iGRAPE over the GRAPE algorithm when it comes to state preparation problems. First, the cost functions used in the iGRAPE algorithm focus solely on disentangling the quantum state, without considering its specific form. This approach leads to multiple possible final states, which helps with optimization convergence. Second, the disentangling process simplifies the optimization problem by breaking it down into lower-dimensional subsystems at each *Step*. This allows for parallel optimization after the initial step. These findings are particularly relevant in today's noisy intermediate-scale quantum (NISQ) era, where implementing quantum algorithms on real quantum systems is a significant challenge.

If the coupling terms g_{ij} in the system's Hamiltonian are fixed and nonzero, in the case of ZZ coupling, we can set one of the subsystems into $|\mathbb{0}\rangle$ so that the other subsystem can evolve independently. If the coupled Hamiltonian H_{ij} takes on different forms [44], we can employ a similar approach to transform one of the subsystems into an eigenstate of H_{ij}. This ensures that operations on one subsystem only contribute a global phase to other subsystems. Currently, the number of operators K_n in the iGRAPE algorithm for the nth *Step* primarily relies on the system and control Hamiltonian. By considering the entanglement information of the target state, it is possible to devise an improved approach that minimizes the number of *Steps*. For transferring an arbitrary state $|\psi\rangle$ to another arbitrary state $|\psi'\rangle$, we can employ the intermediate state $|\mathbb{0}\rangle$ and utilize iGRAPE to

Algorithm 2. The iGRAPE algorithm.

Input: Target state $|\Psi\rangle$
Set tolerance ϵ_0
Output: Pulse sequence U that satisfies $U^\dagger|\mathbb{0}\rangle = |\Psi\rangle$
Set $n = 1$, $|\psi_{n-1,1}\rangle = |\Psi\rangle$
while $\exists a$ s.t. $|\psi_{n-1,a}\rangle \neq$ single qubit state **do** ▷ Outer loop: divide the quantum system
  for m in $\{m||\psi_{n-1,m}\rangle \neq$ single qubit state$\}$ **do** ▷ Inner loop: generate pulse sequence for each subsystem
    Guess initial controls $u_\alpha^{[n,m]}(k)$
    Evolution: $|\phi_{n,m}\rangle = \mathcal{U}_{K_n}^{[n,m]} \ldots \mathcal{U}_1^{[n,m]}|\psi_{n-1,m}\rangle$
    Calculate cost function L
    while $L \geqslant \epsilon_0$ **do** ▷ Minimize cost function L
      Calculate gradient: $\frac{\partial L}{\partial u_\alpha^{[n,m]}(k)}$
      Update: $u_\alpha^{[n,m]}(k) \leftarrow u_\alpha^{[n,m]}(k) - \omega \frac{\partial L}{\partial u_\alpha^{[n,m]}(k)}$
      Evolution: $|\phi_{n,m}\rangle = \mathcal{U}_{K_n}^{[n,m]} \ldots \mathcal{U}_1^{[n,m]}|\psi_{n-1,m}\rangle$
      Recalculate cost function L
    end while
    Pulse sequence for the subsystem is $U^{[n,m]} = \Pi_{k=1}^{K_n} \mathcal{U}_k^{[n,m]}$
  end for
  $n \leftarrow n + 1$
end while
for m in $\{m||\psi_{n-1,m}\rangle \neq |\mathbb{0}\rangle\}$ **do**
  Generate single qubit rotation $U^{[n,m]}$ s.t. $U^{[n,m]}|\psi_{n-1,m}\rangle = |\mathbb{0}\rangle$
end for
The whole pulse sequence is $U = (\bigotimes_m U^{[l+1,m]}) \cdots (\bigotimes_{m=1}^2 U^{[2,m]}) \cdot U^{[1,1]}$

TABLE I. NMR samples of different system sizes. The molecular parameters of each sample including ν_i (diagonal) and J_{ij} (off-diagonal).

Sample name	Molecular formula	System size	Parameters (Hz)
Diethyl fluoromalonate	$FCH(COOC_2H_5)_2$	2 qubits	H: $400M$, 2.8, 1.2 F: 47.6, $376M$, 3.1, 1.3
Diethyl fluoromalonate	$FCH(COOC_2H_5)_2$	3 qubits	C: $100M$, 2.9, 1.1 H: 160.7, $400M$, 2.8, 1.2 F: −194.4, 47.6, $376M$, 3.1, 1.3
Carbon-13-iodotrifluroethylene	C_2F_3I	4 qubits	C: 15 479.88, 7.9, 1.22 F_1: −297.71, −3, 132.45, 6.8, 0.66 F_2: −275.56, 64.74, −42 682.97, 4.4, 0.63 F_2: 39.17, 51.50, −129.08, −56 445.71, 4.8, 0.61
1-bromo-2,4,5-trifluorobenzene	$BrC_6H_2F_3$	5 qubits	F_1: −47 708, 0.8, 50 F_2: −45.5, −45 257, 0.6, 50 F_3: 135.8, 323.5, −37 734, 0.8, 50 H_1: 62.1, 1468.2, 1811.2, 2396, 1.5, 110 H_2: 1781.1, 122.9, 60.9, −10.1, 2393, 1.5, 110
Crotonic acid	$C_4H_6O_2$	7 qubits	C_1: 1750.3 C_2: 40.8, 14 930.1 C_3: 1.6, 69.5, 12 199.9 C_4: 8.47, 1.4, 71.04, 17 173.7 H_1: 4.0, 155.6, −1.8, 6.5, 2785.85 H_2: 6.64, −0.7, 162.9, 3.3, 15.81, 2320.25 H_3: 128, −7.1, 6.6, −0.9, 6.9, −1.7, 718.487

TABLE II. Shaped pulses details of states preparation on NMR quantum systems.

Number of qubits	2	3	4	5	7
iGRAPE pulse steps	500, 100	800, 200	1500, 260	2000, 400	2800, 1200
GRAPE pulse steps	600	1000	1760	2400	3000
Transfer time (ms)	3	5	8.8	12	20

generate a pulse sequence from $|\psi\rangle$ to $|\mathbb{0}\rangle$. Subsequently, this same approach can be used to generate another pulse sequence from $|\mathbb{0}\rangle$ to $|\psi'\rangle$. Currently, the utilization of the iGRAPE algorithm is restricted exclusively to problems related to state transfer in pure quantum states. Nevertheless, our aim for the future is to expand its scope and explore potential applications in broader areas of quantum optimal control. This includes the manipulation of mixed quantum states, the preparation of unitary operations, nonunitary processes in dissipative systems, and enhancing robustness, among others.

Data generated and analyzed during the current study are available from the corresponding author upon reasonable request.

ACKNOWLEDGMENTS

This work is supported by the Innovation Program for Quantum Science and Technology (Grant No. 2021ZD0303205), the National Natural Science Foundation of China (Grants No. 11661161018, No. 11927811, and No. 11875160), Anhui Initiative in Quantum Information Technologies (Grant No. AHY050000), the Guangdong Innovative and Entrepreneurial Research Team Program (2016ZT06D348), Natural Science Foundation of Guangdong Province (Grant No. 2017B030308003), and Science, Technology and Innovation Commission of Shenzhen Municipality (Grants No. ZDSYS20170303165926217, No. JCYJ20170412152620376, and No. JCYJ20170817105046702).

M.-H.Y. and X.P. conceived the project. Y.H., M.-H.Y., and Y.C. formulated the theoretical framework. Y.C. designed the experiment. Y.C., Z.W., and R.L. performed the measurements and analyzed the data. Z.W., R.L., and X.P. assisted with the experiment. X.P. supervised the experiment. Y.C., Y.H., B.-Y.W., and Z.W. wrote the manuscript. All authors contributed to analyzing the data, discussing the results, and commented on the writing.

The authors declare no competing interests.

APPENDIX A: ALGORITHM STEPS

In this paper, we employ the L-BFGS-B optimization algorithm [38] for pulse optimization in both GRAPE and iGRAPE. We successfully attain a minimum state fidelity of 99.7% across all our examples. The central processing unit (CPU) model utilized is an Intel® Xeon® CPU E5-2620 V3 operating at a frequency of 2.4 GHz. Detailed information regarding the NMR sample, including various system sizes and experimental parameters for superconducting quantum systems, will be provided in the subsequent sections.

A quantum system generally can be described by the Hamiltonian

$$H(t) = H_s(t) + H_c(t), \tag{A1}$$

which is tunable through the time-dependent control Hamiltonian $H_c(t)$. The goal of quantum optimal control is to prepare the target state or the target unitary by appropriately adjusting $H_c(t)$. For example, consider the situation where the system is initially in a given state $|\mathbb{0}\rangle = |0\ldots0\rangle$ and we are interested in dynamics that prepares the target state $|\Psi\rangle$ within a transfer time of T. To search for controls that accomplish this task, a cost function L must be introduced to quantify the degree of fulfillment

$$L = 1 - |\langle\Psi|U(T)|\mathbb{0}\rangle|^2, \tag{A2}$$

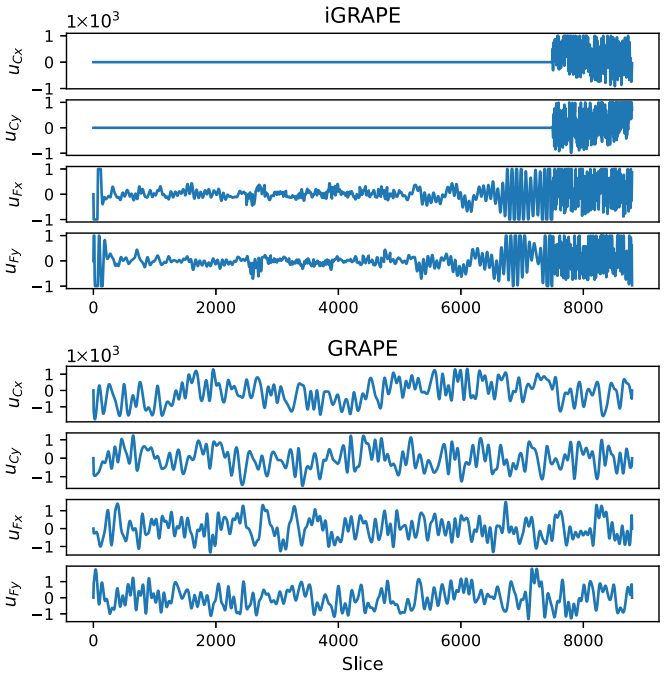

FIG. 5. The pulse sequences produced by iGRAPE and GRAPE for the four-qubit NMR experiment.

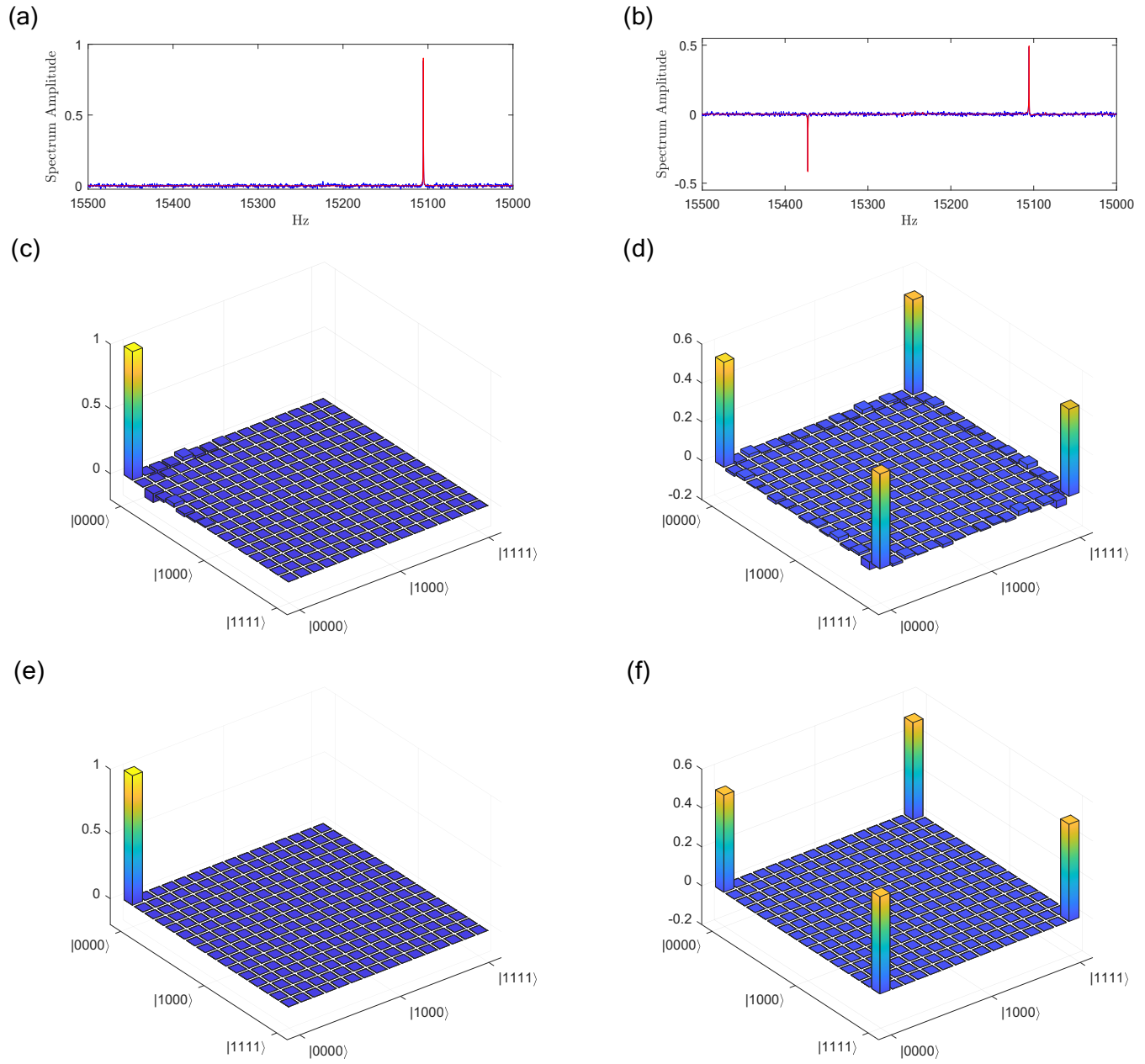

FIG. 6. Experimental data and the corresponding tomography results. (a) Experimental ^{13}C spectra of PPS ρ_{0000} state. (b) Experimental ^{13}C spectra of GHZ state with $R_y(\pi/2)$ rotation on ^{13}C as the readout operator. (c) The PPS tomography and (d) the GHZ-state tomography. (e) The ideal tomography of PPS and (f) the ideal tomography of GHZ state.

where $U(T)$ is determined by solving

$$i\frac{dU(t)}{dt} = H(t)U(t),$$
$$U(0) = \mathbb{1}, \tag{A3}$$

evaluated at $t = T$.

For the GRAPE algorithm, the control Hamiltonian can be expressed as

$$H_c(t) = \sum_\alpha u_\alpha(t)H_\alpha, \tag{A4}$$

where a set of external control fields $u_\alpha(t)$ acting on the system via control operators H_α. We will assume for simplicity that the chosen transfer time T is discretized into K equal steps of duration $\Delta t = T/K$. During each step, the control amplitudes u_α are constant. In other words, during the kth step, the amplitude u_α of the αth control Hamiltonian is given by $u_\alpha(k)$. The time evolution of the system during a time step k is given by the propagator

$$U_k = \exp\left\{-i\Delta t\left(H_s + \sum_\alpha u_\alpha(k)H_\alpha\right)\right\}, \tag{A5}$$

TABLE III. Experimental parameters for all 12 qubits in the superconducting quantum system. ω is the idle frequency of each qubit and η is the anharmonicity. T_1, the energy relaxation time, and T_2^*, the dephasing time extracted from Ramsey experiment, are measured at their respective idle frequencies.

	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11	Q12
ω (GHz)	4.978	4.183	5.192	4.352	5.110	4.226	5.030	4.300	5.142	4.140	4.996	4.260
η (MHz)	−248	−204	−246	−203	−247	−202	−246	−203	−244	−203	−246	−201
T_1 (μs)	40.1	34.7	30.8	43.2	31.8	34.3	46.5	38.1	32.2	54.6	29.6	30.3
T_2^* (μs)	7.9	1.5	6.3	2.4	4.9	2.7	6.8	2.3	5.1	3.5	5.9	3.0

and the corresponding cost function L_g becomes

$$L_g = 1 - |\langle\Psi||\psi(T)\rangle|^2, \tag{A6}$$

where $|\psi(T)\rangle = U_K \ldots U_1|\mathbb{0}\rangle$. Update $u_\alpha(k)$ to minimize the cost function L_g,

$$u_\alpha(k) \leftarrow u_\alpha(k) - \omega\frac{\partial L_g}{\partial u_\alpha(k)}. \tag{A7}$$

Here ω is the learning rate (a small step size). This forms the basis of the GRAPE algorithm. The whole process of GRAPE is shown in Algorithm 1.

The algorithmic process of iGRAPE is extensively described in the preceding text. The pseudocode for iGRAPE can be found in Algorithm 2.

APPENDIX B: PARAMETER SETTINGS

1. For NMR quantum systems

The NMR samples utilized in the journal text are presented in Table I. It is worth mentioning that the subsystem of the three-qubit NMR sample Diethyl fluoromalo corresponds to a two-qubit system, as indicated in the table.

The duration Δt for each propagator $\mathcal{U}_k^{[n,m]}$ mentioned in the journal text is 5 μs, and the parameters of shaped pulses can be found in Table II. In the iGRAPE algorithm, we determine through empirical means how to distribute the number of pulses among each subsystem.

Figure 5 depicts the pulse sequences produced by both algorithms in the four-qubit NMR experiment showcased. The corresponding numerical final-state fidelity of this sequence is 99.83%. It is worth mentioning that the pulse sequence generated by the iGRAPE algorithm at *Step* 2 in Eq. (23) is tailored to the F subsystem, hence it does not encompass the control field for C.

The spectrum of the PPS, denoted as ρ_0, is depicted in Fig. 6(a). This state is obtained by applying shaped pulses to the thermal equilibrium state ρ_{eq}. The figure presented in Fig. 6(b) illustrates the spectrum of the target (GHZ) state, achieved by applying a $R_y(\pi/2)$ rotation on ^{13}C as the readout operator. This state is prepared using a pulse sequence generated by iGRAPE and added to the PPS. Additionally, we conduct full-state tomography to reconstruct the density matrices of both the PPS and the target states. In our NMR setup, we initially obtain only the deviation of the density matrix $\rho_\Delta = \rho - \mathbb{1}/2^N$ through NMR tomography. However, it is important to note that this deviation cannot be considered as a quantum state. Hence, we employ postprocessing techniques to derive density matrices as depicted in Fig. 6(c) (the PPS) and Fig. 6(d) (the target GHZ state). This is achieved by imposing constraints on the normalization [$\text{tr}(\rho) = 1$], hermiticity ($\rho = \rho^\dagger$), and positive semi-definiteness ($\rho \geqslant 0$). All these constraints are realized using the CVX toolbox in MATLAB.

Thus, the reconstructed matrices fulfill all the requirements for a quantum-state density matrix. The theoretical tomography results for both PPS and GHZ state are depicted in Figs. 6(e) and 6(f), respectively. The fidelity between a pure state $|\psi\rangle$ and a density matrix ρ is defined as $F(|\psi\rangle, \rho) = |\langle\psi|\hat{\rho}|\psi\rangle|^2$. The reconstructed PPS in Fig. 6(c) exhibits a fidelity of 99.29%, while the reconstructed GHZ state in Fig. 6(d) demonstrates a fidelity of 98.25%.

2. For superconducting quantum systems

The superconducting quantum systems employed in this study are based on a one-dimensional chain model, where the coupling between qubits can be adjusted. The parameters, such as the idle frequency denoted by ω and the anharmonicity represented by η, are presented in Table III.

TABLE IV. Shaped pulses details of states preparation on superconducting quantum systems.

Number of qubits	2	4	6	8	10	12
iGRAPE pulse steps	320, 300	380, 320, 300	420, 360, 320, 300	440, 380, 340, 300	460, 440, 360, 320, 300	500, 420, 360, 320, 300
GRAPE pulse steps	620	1000	1400	1460	1840	1900
Transfer time (ns)	31	50	70	73	92	95

The duration Δt for each propagator $\mathcal{U}_k^{(n,m)}$ in the journal text is 0.05 ns, and the parameters of shaped pulses are shown in Table IV.

APPENDIX C: BENCHMARK FOR HAAR RANDOM STATES

We supplement a set of numerical tests here for the section of Fig. 4 in the main text, to examine the performance of both algorithms under Haar random states. We utilize IBM's QISKIT package to generate Haar random states. For the cases ofeight qubits and five qubits, we generate 500 quantum states for each scenario and prepare each generated state using two different algorithms. From Fig. 7, it can be observed that the von Neumann entropy corresponding to the generated Haar random states mostly falls within a narrow range. The runtime of both algorithms is also approximately similar to the runtime of the von Neumann entropy shown in Fig. 4 of the main text. Specifically, for eight qubits, iGRAPE has an average runtime of 138.8 s for Haar random states, while GRAPE has an average runtime of 389.4s . On average, iGRAPE provides a speedup factor of 2.8 compared to GRAPE, consistent with the performance under von Neumann entropy shown in Fig. 4 of the main text. For the case of five qubits, iGRAPE has an average runtime of 15.0 s while GRAPE has an average runtime of 25.0 s; on average, iGRAPE provides a speedup factor of 1.67 compared to GRAPE, again consistent with the performance under von Neumann entropy shown in Fig. 4. Based on the numerical simulations, we find that the performance improvement achieved by iGRAPE is primarily determined by the value of von Neumann entropy associated with target states rather than how those target states are obtained initially.

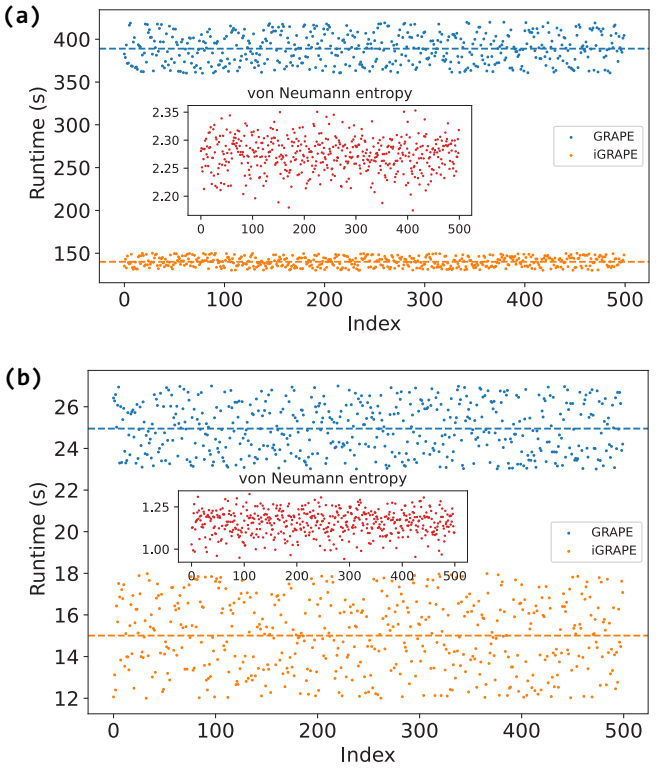

FIG. 7. Benchmark of the iGRAPE and the GRAPE algorithms for Haar random states preparation. We generate 500 Haar random states for each of the two cases corresponding to Fig. 4 and the dashed line represents the average runtime of the algorithm. (a) The runtime of the two algorithms in the case of eight qubits. (b) The runtime of the two algorithms in the case of five qubits.

[1] V. Krotov, *Global Methods in Optimal Control Theory* (CRC Press, Boca Raton, FL, 1995), Vol. 195.

[2] A. E. Bryson and Y.-C. Ho, *Applied Optimal Control: Optimization, Estimation, and Control* (Routledge, Abingdon-on-Thames, England, 2018).

[3] J. Werschnik and E. K. U. Gross, Quantum optimal control theory, J. Phys. B: At., Mol. Opt. Phys. **40**, R175 (2007).

[4] S. J. Glaser, U. Boscain, T. Calarco, C. P. Koch, W. Köckenberger, R. Kosloff, I. Kuprov, B. Luy, S. Schirmer, T. Schulte-Herbrüggen *et al.*, Training schrödinger's cat: Quantum optimal control, Eur. Phys. J. D **69**, 279 (2015).

[5] S. Lloyd and S. Montangero, Information theoretical analysis of quantum optimal control, Phys. Rev. Lett. **113**, 010502 (2014).

[6] S. Machnes, E. Assémat, D. J. Tannor, and F. K. Wilhelm, Tunable, flexible, and efficient optimization of control pulses for practical qubits, Phys. Rev. Lett. **120**, 150401 (2018).

[7] P. J. Liebermann and F. K. Wilhelm, Optimal qubit control using single-flux quantum pulses, Phys. Rev. Appl. **6**, 024022 (2016).

[8] D. J. Egger and F. K. Wilhelm, Adaptive hybrid optimal quantum control for imprecisely characterized systems, Phys. Rev. Lett. **112**, 240503 (2014).

[9] F. Dolde, V. Bergholm, Y. Wang, I. Jakobi, B. Naydenov, S. Pezzagna, J. Meijer, F. Jelezko, P. Neumann, T. Schulte-Herbrüggen *et al.*, High-fidelity spin entanglement using optimal control, Nat. Commun. **5**, 3371 (2014).

[10] F. Platzer, F. Mintert, and A. Buchleitner, Optimal dynamical control of many-body entanglement, Phys. Rev. Lett. **105**, 020501 (2010).

[11] M. H. Goerz, G. Gualdi, D. M. Reich, C. P. Koch, F. Motzoi, K. B. Whaley, J. Vala, M. M. Müller, S. Montangero, and T. Calarco, Optimizing for an arbitrary perfect entangler. II. Application, Phys. Rev. A **91**, 062307 (2015).

[12] P. Watts, J. Vala, M. M. Müller, T. Calarco, K. B. Whaley, D. M. Reich, M. H. Goerz, and C. P. Koch, Optimizing for an arbitrary perfect entangler. i. functionals, Phys. Rev. A **91**, 062306 (2015).

[13] N. Khaneja, T. Reiss, C. Kehlet, T. Schulte-Herbrüggen, and S. J. Glaser, Optimal control of coupled spin dynamics: Design of NMR pulse sequences by gradient ascent algorithms, J. Magn. Reson. **172**, 296 (2005).

[14] C. Ferrie, Self-guided quantum tomography, Phys. Rev. Lett. **113**, 190404 (2014).

[15] G. Turinici, Stochastic learning control of inhomogeneous quantum ensembles, Phys. Rev. A **100**, 053403 (2019).

[16] V. Krotov, Global methods in optimal control theory, in *Advances in Nonlinear Dynamics and Control: A Report from Russia* (Springer, New York, 1993), pp. 74–121.

[17] M. Ozawa, Soundness and completeness of quantum root-mean-square errors, npj Quantum Inf. **5**, 1 (2019).

[18] M. Y. Niu, S. Boixo, V. N. Smelyanskiy, and H. Neven, Universal quantum control through deep reinforcement learning, npj Quantum Inf. **5**, 33 (2019).

[19] X.-M. Zhang, Z.-W. Cui, X. Wang, and M.-H. Yung, Automatic spin-chain learning to explore the quantum speed limit, Phys. Rev. A **97**, 052333 (2018).

[20] F. Albarrán-Arriagada, J. C. Retamal, E. Solano, and L. Lamata, Measurement-based adaptation protocol with quantum reinforcement learning, Phys. Rev. A **98**, 042315 (2018).

[21] M. August and J. M. Hernández-Lobato, Taking gradients through experiments: Lstms and memory proximal policy optimization for black-box quantum control, in *International Conference on High Performance Computing* (Springer, New York, 2018), pp. 591–613.

[22] M. Bukov, A. G. R. Day, D. Sels, P. Weinberg, A. Polkovnikov, and P. Mehta, Reinforcement learning in different phases of quantum control, Phys. Rev. X **8**, 031086 (2018).

[23] J.-J. Chen and M. Xue, Manipulation of spin dynamics by deep reinforcement learning agent, arXiv:1901.08748.

[24] P. Doria, T. Calarco, and S. Montangero, Optimal control technique for many-body quantum dynamics, Phys. Rev. Lett. **106**, 190501 (2011).

[25] N. Rach, M. M. Müller, T. Calarco, and S. Montangero, Dressing the chopped-random-basis optimization: A bandwidth-limited access to the trap-free landscape, Phys. Rev. A **92**, 062343 (2015).

[26] J. Kelly, R. Barends, B. Campbell, Y. Chen, Z. Chen, B. Chiaro, A. Dunsworth, A. G. Fowler, I.-C. Hoi, E. Jeffrey *et al.*, Optimal quantum control using randomized benchmarking, Phys. Rev. Lett. **112**, 240504 (2014).

[27] N. Leung, M. Abdelhafez, J. Koch, and D. Schuster, Speedup for quantum optimal control from automatic differentiation based on graphics processing units, Phys. Rev. A **95**, 042318 (2017).

[28] X. Ge and R.-B. Wu, Risk-sensitive optimization for robust quantum controls, Phys. Rev. A **104**, 012422 (2021).

[29] X. Ge, H. Ding, H. Rabitz, and R.-B. Wu, Robust quantum control in games: An adversarial learning approach, Phys. Rev. A **101**, 052317 (2020).

[30] X.-d. Yang, C. Arenz, I. Pelczer, Q.-M. Chen, R.-B. Wu, X. Peng, and H. Rabitz, Assessing three closed-loop learning algorithms by searching for high-quality quantum control pulses, Phys. Rev. A **102**, 062605 (2020).

[31] M. Dalgaard, F. Motzoi, J. H. M. Jensen, and J. Sherson, Hessian-based optimization of constrained quantum control, Phys. Rev. A **102**, 042612 (2020).

[32] J. Saywell, M. Carey, M. Belal, I. Kuprov, and T. Freegarde, Optimal control of raman pulse sequences for atom interferometry, J. Phys. B: At., Mol. Opt. Phys. **53**, 085006 (2020).

[33] S. Boutin, C. K. Andersen, J. Venkatraman, A. J. Ferris, and A. Blais, Resonator reset in circuit QED by optimal control for large open quantum systems, Phys. Rev. A **96**, 042315 (2017).

[34] S. Kwon, A. Tomonaga, G. Lakshmi Bhai, S. J. Devitt, and J.-S. Tsai, Gate-based superconducting quantum computing, J. Appl. Phys. **129**, 041102 (2021).

[35] O. Katz, M. Cetina, and C. Monroe, Programmable N-body interactions with trapped ions, PRX Quantum **4**, 030311 (2023).

[36] X. Rong, J. Geng, F. Shi, Y. Liu, K. Xu, W. Ma, F. Kong, Z. Jiang, Y. Wu, and J. Du, Experimental fault-tolerant universal quantum gates with solid-state spins under ambient conditions, Nat. Commun. **6**, 8748 (2015).

[37] M. Gong, M.-C. Chen, Y. Zheng, S. Wang, C. Zha, H. Deng, Z. Yan, H. Rong, Y. Wu, S. Li *et al.*, Genuine 12-qubit entanglement on a superconducting quantum processor, Phys. Rev. Lett. **122**, 110501 (2019).

[38] R. H. Byrd, P. Lu, J. Nocedal, and C. Zhu, A limited memory algorithm for bound constrained optimization, SIAM J. Sci. Comput. **16**, 1190 (1995).

[39] E. Knill, R. Laflamme, R. Martinez, and C.-H. Tseng, An algorithmic benchmark for quantum information processing, Nature (London) **404**, 368 (2000).

[40] J. Li, R. Fan, H. Wang, B. Ye, B. Zeng, H. Zhai, X. Peng, and J. Du, Measuring out-of-time-order correlators on a nuclear magnetic resonance quantum simulator, Phys. Rev. X **7**, 031011 (2017).

[41] G. W. K. Moore, S. E. L. Howell, M. Brady, X. Xu, and K. McNeil, Anomalous collapses of Nares Strait ice arches leads to enhanced export of Arctic sea ice, Nat. Commun. **12**, 1 (2021).

[42] X. Peng, X. Zhu, X. Fang, M. Feng, K. Gao, X. Yang, and M. Liu, Preparation of pseudo-pure states by line-selective pulses in nuclear magnetic resonance, Chem. Phys. Lett. **340**, 509 (2001).

[43] S. Ryu and T. Takayanagi, Aspects of holographic entanglement entropy, J. High Energy Phys. 08 (2006) 045.

[44] M. A. Nielsen and I. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, England, 2002).