Robust quantum control in games: An adversarial learning approach

Xiaozhen Ge,[1] Haijin Ding,[1] Herschel Rabitz,[2,*] and Re-Bing Wu[1,3,†]
[1]*Department of Automation, Tsinghua University, Beijing 100084, China*
[2]*Department of Chemistry, Princeton University, Princeton, New Jersey 08544, USA*
[3]*Center for Quantum Information Science and Technology, BNRist, Beijing 100084, China*

(Received 27 November 2019; accepted 16 April 2020; published 11 May 2020)

High-precision operation of quantum computing systems must be robust to uncertainties and noises in the quantum hardware. We show that through a game played between the uncertainties (or noises) and the controls, adversarial uncertainty samples can be generated to find highly robust controls through the search for Nash equilibria. We propose a broad family of adversarial learning algorithms, namely a-GRAPE algorithms, which includes two effective learning schemes referred to as the best-response approach and the better-response approach within game-theoretic terminology, providing options for learning highly robust controls. Numerical experiments demonstrate that the balance between fidelity and robustness depends on the details of the chosen adversarial learning algorithm, which can effectively lead to a significant enhancement of control robustness while attaining high fidelity.

DOI: 10.1103/PhysRevA.101.052317

I. INTRODUCTION

Recent experimental breakthroughs in quantum computing have signaled its commercialization in the foreseeable future [1–3]. The demands involved require highly precise and stable control techniques for the deterministic implementation of quantum gates [4]. Generally, high precision is relatively easy to achieve if a well-characterized model is available, for example, using the highly efficient GRAPE (gradient ascent pulse engineering) algorithm [5] and methodologies based on reinforcement learning [6]. The real challenge is to maintain high precision in the presence of realistic uncertainties and noises in the model, i.e., finding both high-precision and robust controls. In the literature, a variety of proposals has been put forth for either online or offline searches for robust quantum controls, e.g., stimulated Raman adiabatic passage working to overcome pulse shape errors [7–9], dynamical decoupling to fight against decoherence noises [10–13], and differential evolution [14,15] as well as ensemble-based algorithms [5,16] that aim to deal with a variety of noises and uncertainties.

Recently, a class of robust control design methods have been proposed [17–21]. It was shown that by Monte Carlo sampling of the uncertainties or noises, the GRAPE algorithm can be exploited to effectively improve the robustness against linewidth broadening [17–19], amplitude errors of control fields [18,19], coupling uncertainty [21], and clock noises [22].

Most of the above approaches are based on improving the robustness of a quantum control protocol quantified by the average performance with respect to the system uncertainties or noises to be suppressed. This measure is relatively easy to evaluate and hence to optimize. However, it is not the unique choice nor necessarily the best. For example, fighting against the worst-case performance was widely adopted in the control of classical systems [23–26], which leads to a class of min-max problems. The optimization with respect to such an objective can effectively reduce the risk of failure, which is demanded by fault-tolerant quantum computation. The worst-case optimization for robust quantum controls was first formulated in the control of molecular systems against disturbance [27,28], and a theoretical analysis was later done on the robust performance via the Dyson expansion [29]. However, there are very few algorithms for efficiently solving the associated min-max problem. In [30–32], the so-called H_∞ approach was applied to a class of linear quantum systems in the Heisenberg picture. In [33], sequential convex programming (SCP) was proposed for solving the worst-case robust optimization problem in a single-qubit system, which decomposes the min-max problem into a sequence of convex optimization procedures.

Since the min-max problem can be taken as an adversarial game between the control and the uncertainties (or noises) that attempt to affect the objective of the quantum system (e.g., state or gate infidelity) in opposite directions, one can introduce game-theoretic learning algorithms that seek Nash equilibria (NE) [34]. Such ideas have been successfully used in deep learning. As an example, for the Generative Adversarial Nets (GAN) model, a large number of applications show that the learning model can be trained to be more *generalizable* by actively creating adversarial samples produced by a discriminator neural network [35–38].

In this paper, we will extend game-theoretic ideas to the design of robust quantum controls. Even if the Nash equilibrium does not exist or is hard to reach, the robustness of the control can still be enhanced by the learning process. As the adversarial game has a close relationship with both the fidelity and robustness, the game-based learning process

*hrabitz@princeton.edu
†rbwu@tsinghua.edu.cn

can be adjusted flexibly to make a balance between the two performance measures, forming a broad family of adversarial learning algorithms. The remainder of this paper will be arranged as follows. Section II presents a game-theoretic analysis of the robust control problem. In Sec. III, we introduce several adversarial learning control algorithms based on the dynamic processes in the game theory of learning and evolution. In Sec. IV, the effectiveness of the proposed algorithms is illustrated through simulations of the control design in two representative examples. Finally, concluding remarks are made in Sec. V.

II. THE ZERO-SUM GAME BETWEEN CONTROL AND UNCERTAINTY

Consider an N-dimensional quantum control system whose unitary propagator obeys the following Schrödinger equation:

$$\dot{U}(t;\boldsymbol{u},\boldsymbol{\epsilon}) = -iH[t;\boldsymbol{u},\boldsymbol{\epsilon}]U(t;\boldsymbol{u},\boldsymbol{\epsilon}), \tag{1}$$

where $U(\cdot) \in \mathbb{C}^{N\times N}$ represents the quantum gate operation with the initial value being the identity matrix. The system's Hamiltonian $H[t;\boldsymbol{u},\boldsymbol{\epsilon}]$ is dependent on a vector of control parameters $\boldsymbol{u}$ (e.g., in-phase and quadrature amplitudes, or phases and amplitudes of laser pulses in the frequency domain) and a vector of uncertainty parameters $\boldsymbol{\epsilon}$ (e.g., environmental noises or imprecisely identified parameters).

Let U_f be the target gate operation and

$$L[\boldsymbol{u},\boldsymbol{\epsilon}] = N^{-2}\|U(T;\boldsymbol{u},\boldsymbol{\epsilon}) - U_f\|^2 \tag{2}$$

be the infidelity of the controlled gate under the control $\boldsymbol{u}$ and the uncertainty $\boldsymbol{\epsilon}$, where $\|\cdot\|$ is the Frobenius norm. The robustness objective is to find a control $\boldsymbol{u}$ under which $L[\boldsymbol{u},\boldsymbol{\epsilon}]$ is as small as possible for as many values of the uncertainty $\boldsymbol{\epsilon}$ as possible.

A straightforward approach adopted in most existing studies is to minimize the average infidelity, i.e.,

$$\mathbf{E}_{\boldsymbol{\epsilon}}L[\boldsymbol{u},\boldsymbol{\epsilon}] = \int_{\boldsymbol{\epsilon}} L[\boldsymbol{u},\boldsymbol{\epsilon}]p_0(\boldsymbol{\epsilon})d\boldsymbol{\epsilon}, \tag{3}$$

where $p_0(\boldsymbol{\epsilon})$ is the probability density distribution of $\boldsymbol{\epsilon}$. This objective is relatively easy to estimate and thus to optimize, but the resulting control may not be able to dictate all possible cases of uncertainties due to a lack of control over the variance of the infidelity (i.e., the infidelity can be high over certain small ranges of uncertainties even if the average infidelity is low).

To reduce the risk of encountering high infidelity, we consider the worst-case performance instead of the average performance, which leads to the following min-max problem:

$$\min_{\boldsymbol{u}} \max_{\boldsymbol{\epsilon}} L[\boldsymbol{u},\boldsymbol{\epsilon}], \tag{4}$$

which does not rely on the probability density distribution of $\boldsymbol{\epsilon}$. Once the worst-case infidelity is below the desired threshold value, the risk of high infidelity can be effectively reduced.

From a game-theoretic point of view, the optimization process can be taken as a game between two players, the control $\boldsymbol{u}$ and the uncertainty $\boldsymbol{\epsilon}$, in which $\boldsymbol{u}$ attempts to reduce the gate infidelity while $\boldsymbol{\epsilon}$ tries to increase it. Namely, the payoff functions for the control and the uncertainty are $-L[\boldsymbol{u},\boldsymbol{\epsilon}]$ and $L[\boldsymbol{u},\boldsymbol{\epsilon}]$, respectively. Since the sum of the two players' payoffs is always zero, it is called a zero-sum game. A robust control is naturally associated with the NE point $(\boldsymbol{u}^*,\boldsymbol{\epsilon}^*)$ of the zero-sum game [34], at which each player is unable to go any further by tuning merely $\boldsymbol{u}$ or merely $\boldsymbol{\epsilon}$. That is to say, it holds that

$$\boldsymbol{u}^* = \arg\min_{\boldsymbol{u}} L[\boldsymbol{u},\boldsymbol{\epsilon}^*] \text{ and } \boldsymbol{\epsilon}^* = \arg\max_{\boldsymbol{\epsilon}} L[\boldsymbol{u}^*,\boldsymbol{\epsilon}]. \tag{5}$$

A standard approach to search for such a strategy profile is to alternately minimize (with respect to $\boldsymbol{u}$) and maximize (with respect to $\boldsymbol{\epsilon}$) the infidelity. Even if the NE does not exist or is hard to reach, the robustness of the control may still be enhanced during the process of fighting against the *adversarial* samples of $\boldsymbol{\epsilon}$ that yield the worst performance. In this spirit, a family of learning algorithms can be devised in which the uncertainty parameters play a more active role instead of just being averaged out. Since the algorithms are fundamentally based on the gradient-based GRAPE algorithm [5] applied in the minimization process for updating controls, the adversarial learning algorithms to be presented below will be termed as a-GRAPE, where "a" stands for "adversarial."

III. THE DESIGN OF ADVERSARIAL LEARNING CONTROL ALGORITHMS

In this section, we will propose two types of a-GRAPE algorithms, namely, the best-response and better-response approaches, for training highly robust controls via active selection of adversarial samples.

A. Best-response approach

The simplest NE-seeking approach consists of rounds of alternate optimization with the control and the uncertainty. Suppose that we have obtained an adversarial sample $\boldsymbol{\epsilon}^{(k)}$ subject to the optimal control $\boldsymbol{u}^{(k)}$ in the kth round. In the $(k+1)$-th round, we first minimize $L[\boldsymbol{u},\boldsymbol{\epsilon}^{(k)}]$ with respect to $\boldsymbol{u}$ using the GRAPE algorithm, which updates the control by $\boldsymbol{u}^{(k+1)}$. Then, we update the adversarial sample of uncertainty parameters by $\boldsymbol{\epsilon}^{(k+1)}$ that maximizes $L[\boldsymbol{u}^{(k+1)},\boldsymbol{\epsilon}]$ with respect to $\boldsymbol{\epsilon}$.

Utilizing game-theory terminology, we call such an adversarial learning process a best-response approach because each player chooses its best strategy against the opponent [39]. However, for most robust quantum control problems, the best-response strategy can hardly reach a pure NE. This conclusion follows from the facts that a NE requires $L[\boldsymbol{u}^*,\boldsymbol{\epsilon}] = 0$ for all admissible $\boldsymbol{\epsilon}$, as the control is assumed to have fully adequate resources (e.g., bandwidth) such that $\min_{\boldsymbol{u}} L[\boldsymbol{u},\boldsymbol{\epsilon}] = 0$ for any $\boldsymbol{\epsilon}$, and that this condition is hard to satisfy in practice unless $\boldsymbol{\epsilon}$ is only allowed to vary over a very small domain. Since a min-max problem possesses a NE when the minimization problem is convex and maximization problem is concave, a viable strategy is to seek a mixed NE in a enlarged domain, where the uncertainty $\boldsymbol{\epsilon}$ is allowed to adopt mixed strategies, i.e., instead of picking a single adversarial sample, we generate a distribution of adversarial samples. Mathematically, this leads to the following min-max problem:

$$\min_{\boldsymbol{u}} \max_{p(\boldsymbol{\epsilon})\in\mathcal{P}} \int L[\boldsymbol{u},\boldsymbol{\epsilon}]p(\boldsymbol{\epsilon})d\boldsymbol{\epsilon}, \tag{6}$$

where the maximization is performed over the space of probability density distributions $\mathcal{P}$. The mixed NE $[\boldsymbol{u}^*, p^*(\boldsymbol{\epsilon})]$ satisfies

$$\boldsymbol{u}^* = \arg\min_{\boldsymbol{u}} \int L[\boldsymbol{u}, \boldsymbol{\epsilon}] p^*(\boldsymbol{\epsilon}) d\boldsymbol{\epsilon},$$

$$p^*(\boldsymbol{\epsilon}) = \arg\max_{p(\boldsymbol{\epsilon})} \int L[\boldsymbol{u}^*, \boldsymbol{\epsilon}] p(\boldsymbol{\epsilon}) d\boldsymbol{\epsilon}.$$

The advantage of searching for a mixed NE is that its existence is more likely ensured by the linear dependence in $p(\boldsymbol{\epsilon})$ and hence a mixed NE is easier to find.

However, problem (6) is computationally much more expensive than problem (4) because the search space for the maximization part is much larger. In practice, this issue can be relaxed by approximating the optimal probability distribution in order to simplify the maximization process.

Here, we propose that the optimal distribution $p(\boldsymbol{\epsilon})$ can be approximated by exploiting adversarial samples found in the past rounds. In particular, we perform the maximization process in the same way as the above best-response approach and approximate the optimal probability distribution in the $(k+1)$-th round as

$$p^{(k+1)}(\boldsymbol{\epsilon}) \approx \frac{1}{k} \sum_{j=1}^{k} \delta(\boldsymbol{\epsilon} - \boldsymbol{\epsilon}^{(j)}) \tag{7}$$

using the historic adversarial samples $\boldsymbol{\epsilon}^{(1)}, \boldsymbol{\epsilon}^{(2)}, \ldots, \boldsymbol{\epsilon}^{(k)}$ in the past k rounds. Consequently, in the following minimization process, the objective function based on the distribution (7) can be written as the average infidelity over the k adversarial samples,

$$J[\boldsymbol{u}, B_k] = \frac{1}{k} \sum_{j=1}^{k} L[\boldsymbol{u}, \boldsymbol{\epsilon}^{(j)}], \tag{8}$$

where $B_k = \{\boldsymbol{\epsilon}^{(1)}, \ldots, \boldsymbol{\epsilon}^{(k)}\}$.

In practice, the approximation (7) can be chosen more flexibly. For example, one does not have to use all historic adversarial samples in (8) because it will be too costly when k is large and the early samples are likely less adversarial. Thus, we keep only the latest few samples, i.e., let the algorithm utilize only a finite number, say s, of adversarial samples (see Algorithm 1 for a summary). In this scenario, the originally discussed best-response approach can be taken as a special case with memory size $s = 1$.

B. Better-response approach

In the above best-response approach, the minimization process is usually efficient as long as the control resources (e.g., bandwidth, pulse energy, etc.) are abundant, owing to the underlying nice control landscape topology over which almost all locally optimal controls are actually globally optimal [40,41]. However, the generation of adversarial samples is much harder because the maximization process is usually nonconcave. Here, we relax this problem by choosing strong, but not necessarily the strongest adversarial samples for the training of robust controls. In this regard, we call this method a better-response approach.

Algorithm 1. *best-response* a-GRAPE

Initialize:
an initial control $\boldsymbol{u}^{(0)}$;
an initial uncertainty sample set $B_0 = \{0\}$;
a set memory size s.
Repeat:
(1) Use the GRAPE algorithm to update the control by the optimal solution of minimizing $J[\boldsymbol{u}, B_{k-1}]$, i.e.,

$$\boldsymbol{u}^{(k)} = \arg\min_{\boldsymbol{u}} J[\boldsymbol{u}, B_{k-1}],$$

in which $\boldsymbol{u}^{(k-1)}$ is taken as the initial guess for the GRAPE algorithm. Here, k is an index to the current number of round.
(2) Generate a new adversarial sample by

$$\boldsymbol{\epsilon}^{(k)} = \arg\max_{\boldsymbol{\epsilon}} L[\boldsymbol{u}^{(k)}, \boldsymbol{\epsilon}].$$

(3) Update:
if $|B_k| < s$, **then** $B_k = B_{k-1} \cup \{\boldsymbol{\epsilon}^{(k)}\}$,
else
$B_k = \{\boldsymbol{\epsilon}^{(k)}, \boldsymbol{\epsilon}^{(k-1)}, \ldots, \boldsymbol{\epsilon}^{(k-s+1)}\}$.
End: if the stopping criteria are satisfied.

The simplest way to search for better-response adversarial samples is to randomly choose a batch of uncertainty samples, calculate their corresponding cost, and keep the worst few members among them for the adversarial training in the next round (see Algorithm 2 for a description). The batch size of the samples should be sufficiently large so that the chosen adversarial samples have members close to the worst-case samples, but not too large to maintain computational efficiency. Naturally, the better-response approach relying on random sampling usually takes more rounds of gaming, but each round can be much faster when the batch is not very large. Moreover, the randomness of sample batches in the better-response approach may bring additional benefits for the search to get away from unwanted false worst-case traps.

Algorithm 2. *better-response* a-GRAPE

Initialize:
an initial control $\boldsymbol{u}^{(0)}$;
a set ratio r, $r \in (0, 1)$.
Repeat:
(1) Randomly generate M uncertainty samples, compute the corresponding infidelity, and form an adversarial sample set denoted as B_k by retaining the first rM worst ones, where k denotes the current number of round.
(2) Use the GRAPE algorithm to update the control by

$$\boldsymbol{u}^{(k)} = \arg\min_{\boldsymbol{u}} J[\boldsymbol{u}, B_k],$$

where $\boldsymbol{u}^{(k-1)}$ is taken as the initial guess for the GRAPE algorithm.
End: if the stopping criteria are satisfied.

IV. SIMULATION RESULTS

To illustrate the above game-based adversarial learning strategies for the robust control design, we simulate two quantum gate synthesis examples in this section. As will be seen in the simulations, the best-response approach can effectively suppress the worst-case performance, but does not always lead to good performance in high-precision regimes, where the better-response approach is more effective. That is to say, the best-response and the better-response approaches do not mean the best performance and the better performance, respectively, and which approach performs satisfactorily depends on case-specific requirements.

A. Two-qubit system

We first consider a quantum two-qubit gate control problem in a system of two superconducting transmon qubits. The two qubits are dispersively coupled with strength g, but in practice it may deviate from this value. We also consider the uncertainty of the Rabi driving fields that may shift. The system Hamiltonian is as follows:

$$H(t) = (1+\epsilon_0)g\sigma_{1z}\otimes\sigma_{2z} + \sum_{i=1}^{2}(1+\epsilon_i) \times [u_{ix}(t)\sigma_{ix} + u_{iy}(t)\sigma_{iy}],$$

where $g = 10$ MHz is the identified qubit-qubit coupling strength, with ϵ_0 being the error in the coupling constant; u_{ix} and u_{iy} are the control functions which are delivered to the ith qubit through the same control line; and ϵ_1 and ϵ_2 represent the fluctuations in control fields. The dimensionless three uncertainty parameters are all assumed to be bounded by $|\epsilon_i| \leqslant 0.2$. In the simulation, the time duration of the control pulses is chosen as $T = 300$ ns, which is evenly divided into 100 intervals over which the control fields are piecewise constant. The target U_f is set as the controlled-NOT gate.

In the simulations, we set the initial controls to be $u_{kx}(t) = A_k\cos(\omega_k t + \phi_k)$ and $u_{ky}(t) = A_k\sin(\omega_k t + \phi_k)$, where A_k, ω_k, and ϕ_k are randomly chosen. In the best-response approach, we apply the standard genetic algorithm in MATLAB to seek strongly adversarial samples with up to 10^8 generations and the termination tolerance being 10^{-4} on the function value (TolFun). In each round, we use $s = 10$ historic adversarial samples to train the control function with a quasi-Newton algorithm (MATLAB routine fminunc, TolFun $= 10^{-18}$, MaxIter $= 10^5$). In the better-response approach, we uniformly generate $M = 100$ random uncertainty samples in each round and keep the first 10% (i.e., $r = 0.1$) worst ones as adversarial samples for training the control where the means is the same as above. Figure 1 shows the resulting learning curves, namely, the achieved worst-case infidelity $L_{\max}$ versus the number of rounds, as well as the corresponding minimized average infidelity $J_{\min}$ over the selected adversarial samples versus the number of rounds. The robustness of the controls can be directly seen from the curves of worst-case infidelity, which are all enhanced during the optimization. The $L_{\max}$ curve is initially far from the $J_{\min}$ curve, but the gap is quickly reduced after several rounds of gaming. In the best-response approach, the gap is almost closed, showing that the optimized control and uncertainty samples are likely close to a mixed NE. In the better-response approach, the gap still remains large after 2000 rounds. For both approaches, the control robustness is still enhanced by the game, with the (approximate) worst-case infidelity decreased to the level of 10^{-2}.

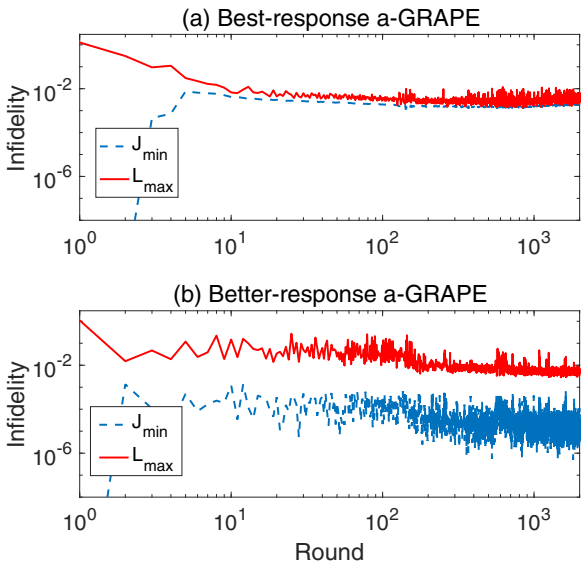

FIG. 1. The learning curves of a-GRAPE for robust controls of the two-qubit system. The red (upper) line corresponds to the worst-case infidelity, while the blue (lower) one corresponds to the minimized average infidelity over selected adversarial samples in cases (a) the best-response approach with memory size $s = 10$ and (b) the better-response approach with $M = 100$ and $r = 0.1$.

As discussed above, neither the average performance nor the worst-case performance is the unique measure for quantifying the control robustness. To better evaluate the overall performance of an optimized control, we calculate and display the cumulative probability distribution function (cdf) $F(l)$ of the gate infidelity, i.e., the probability for the infidelity being smaller than l, in Fig. 2. Here, we also compare the a-GRAPE algorithms with the recently proposed b-GRAPE algorithm [21] (see Appendix for details) for robust control design subject to the average infidelity [i.e., Eq. (3)]. In the simulations, the b-GRAPE algorithm is run by 1 million iterations having the minibatch size $n_{mb} = 1$ and a learning rate $\alpha = 0.002$, while the a-GRAPE algorithms are run by 844 rounds in the best-response approach and 1504 rounds in the better-response approach.

The cdf curve can be used to evaluate the control robustness from two perspectives. First, given a desired value of gate infidelity l_0 (e.g., the threshold error for quantum error correction), the cumulative probability $F(l_0)$ gives the confidence that the control can suppress the error below the value l_0. Second, given an expected confidence F_0 (say, 90%), the cdf can tell us at which threshold value [i.e., l_0 such that $F(l_0) = F_0$] the control can guarantee the confidence. As will be seen below, the robustness performance may vary at different levels of desired infidelity or expected confidence.

Figure 2 clearly shows that the controls optimized by the a-GRAPE algorithms, especially by the better-response approach, are much more robust as almost the entire cdf curves

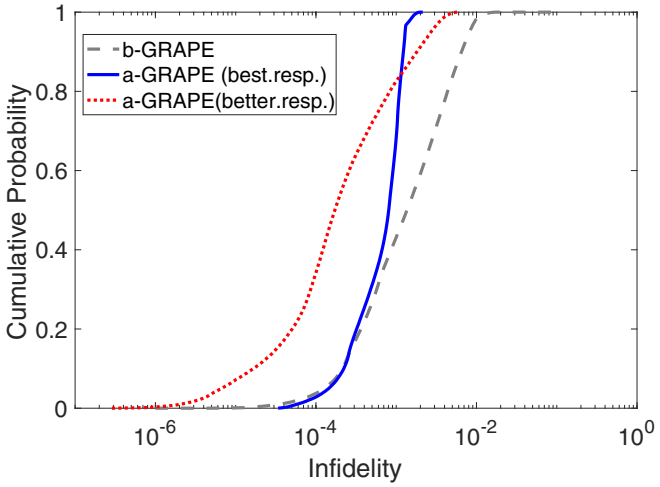

FIG. 2. For the two-qubit system, the cumulative probability functions $F(l)$ of the gate infidelity l under the controls optimized with b-GRAPE (batch size $n_{mb} = 1$), best-response a-GRAPE (memory size $s = 10$), and better-response a-GRAPE ($M = 100$ and $r = 0.1$).

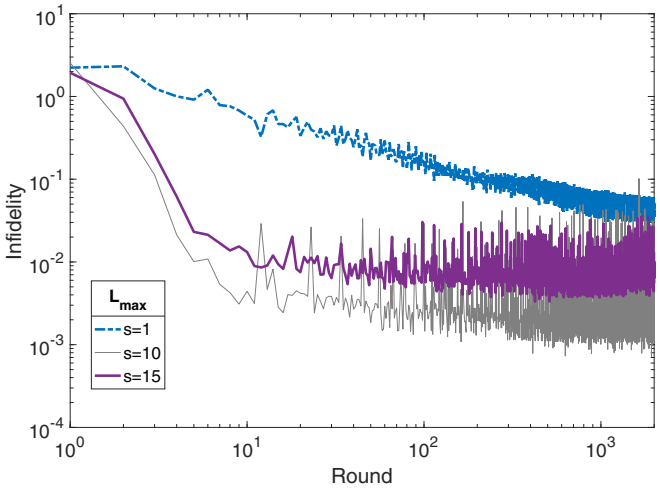

FIG. 3. The worst-case infidelity vs the number of rounds in best-response a-GRAPE optimization of the three-qubit system with different memory sizes.

are higher than that generated with the b-GRAPE algorithm (i.e., with greater confidence at each value of infidelity). For example, as seen in Table I, the control optimized with the better-response a-GRAPE can suppress the gate error below 10^{-3} with a high confidence of 82.5%, while the control optimized with b-GRAPE has only 43.1% confidence. At the higher-precision level (i.e., infidelity lower than 10^{-4}), the better-response approach still maintains 34.2% confidence, while the control optimized with b-GRAPE provides only 3.7% confidence. The performance of the best-response a-GRAPE is only a little poorer than b-GRAPE in the high-precision regime, but much higher in the low-precision to medium-precision regime (i.e., infidelity in $\sim 10^{-3}$–10^{-2}). Additionally, the best-response a-GRAPE achieves a lower worst-case infidelity than that achieved by the better-response a-GRAPE, which is consistent with the finding from the $L_{\max}$ curves in Fig. 1. However, the best-response a-GRAPE has a poorer performance in the high-precision regime than the better-response a-GRAPE, which is also indicated by the $J_{\min}$ curves displayed in Fig. 1. The comparison between different algorithms shows that there is no unique criterion for evaluating the control robustness. An optimized control may achieve satisfactory precision (e.g., infidelity in $\sim 10^{-3}$–10^{-2}) over a large regime of uncertainties, but the highest precision it can achieve may be poor. In practice, one may need a balance between the high precision and the robust regions, especially using limited control resources.

TABLE I. The two rows list the confidence for the gate infidelity to be below 10^{-3} and 10^{-4}.

		a-GRAPE	
	b-GRAPE	Best response	Better response
10^{-3}	43.1%	68.1%	82.5%
10^{-4}	3.7%	2.8%	34.2%

B. Three-qubit system

To see more clearly how the control performance relies on the uncertainties and algorithmic parameters, we simulate a three-qubit system with two uncertainty parameters, whose Hamiltonian is

$$H(t) = J_{12}(1+\epsilon_1)\sigma_{1z}\sigma_{2z} + J_{23}(1+\epsilon_2)\sigma_{2z}\sigma_{3z} + \sum_{k=1}^{3}[u_{kx}(t)\sigma_{kx} + u_{ky}(t)\sigma_{ky}],$$

where the nominal coupling constants are $J_{12} = J_{23} = 10$ MHz. The uncertainty parameters ϵ_1 and ϵ_2 (i.e., identification errors of the coupling constants) are bounded by $|\epsilon_i| \leqslant 0.2$. In the simulation, the target unitary operation is selected as the Toffoli gate. The time period $[0, T]$, where $T = 1\ \mu$s, is evenly divided into 100 intervals, over which the control fields are piecewise constant. In the following simulations, the initial controls are set as specified in the two-qubit example.

The previous illustration with the two-qubit system showed that the controls optimized by the a-GRAPE algorithms could improve the worst-case performance, and here we want to see how the performance depends on the parameters, e.g., s and r. We first compare the best-response a-GRAPE optimization processes with different memory sizes $s = 1,\ 10$, and 15. The learning curves are shown in Fig. 3 in which the minimization curves are not displayed because the worst-case performance is only related with the maximization curve. It can be seen that the algorithm converges faster and finds more robust controls when using more, but not too many, historic adversarial samples. For example, the worst-case infidelity reaches 10^{-2} after only seven rounds when $s = 10$, which converges faster than the case $s = 1$, and the worst infidelity is much lower. However, the case $s = 15$ performs less satisfactorily than the case $s = 10$. This is reasonable because elder historic samples tend to be less adversarial due to the fading-memory effect.

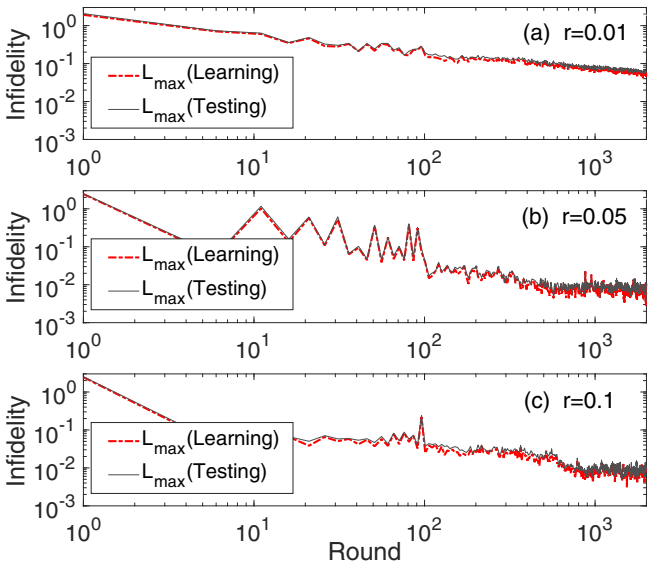

FIG. 4. The worst-case infidelity vs the number of rounds in better-response a-GRAPE optimization of the three-qubit system with $r = 0.01,\ 0.05$, and 0.1.

For the better-response approach, we choose $M = 100$ and compare the performance under $r = 0.01,\ 0.05$, and 0.1, where the true worst-case infidelity in each round is estimated by 2000 independent random samples. Similar to the case of the best-response approach, the simulation results (see the learning curves in Fig. 4) show that the robustness of the optimized controls can be improved by using adequately many adversarial samples, but too many will not bring further improvement.

Since there are only two uncertainty parameters in this example, we can plot a three-dimensional (3D) landscape to show how the infidelity varies with them, from which we can evaluate the overall robustness. In Fig. 5, we display 3D plots under controls optimized with the algorithms b-GRAPE (after 2 million iterations), best-response a-GRAPE (after 629 rounds), and better-response a-GRAPE (after 1986 rounds). The comparison shows that both a-GRAPE algorithms outperform the b-GRAPE algorithm, as most of their landscape surfaces are below that of b-GRAPE. The best-response a-GRAPE achieves the lowest worst-case infidelity and effectively suppresses almost the entire landscape down below the level of $L = 10^{-3}$. However, its overall performance in the higher-precision regime (e.g., $L = 10^{-4}$) is poorer than the better-response a-GRAPE.

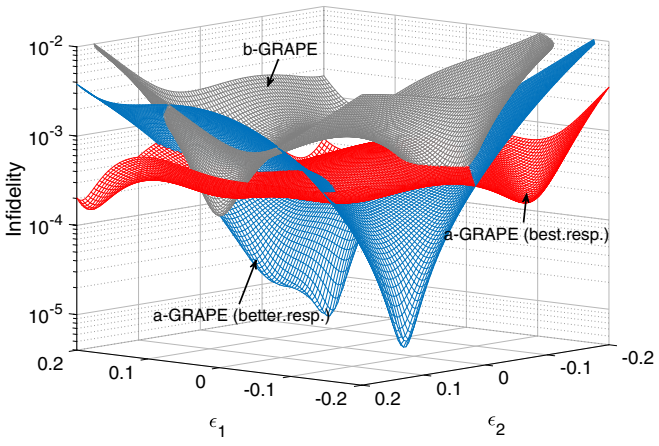

FIG. 5. The infidelity vs uncertainty parameters under controls obtained from b-GRAPE (the gray curve) with $n_{mb} = 1$ and $\alpha = 0.002$, best-response a-GRAPE (the red curve) with memory size $s = 10$, and better-response a-GRAPE (the blue curve) with ratio $r = 0.05$, respectively.

Algorithm 3. relaxed *best*/*better-response* a-GRAPE

Initialize:
an initial control $\boldsymbol{u}$;
a set memory size s or ratio r;
an initial adversarial sample set B.
Repeat:
(1) Randomly generate m uncertainty samples, and select the worst one or the first rm worst ones.
(2) Do the following GRAPE optimization for n iterations:

$$\boldsymbol{u} \leftarrow \boldsymbol{u} - \alpha \cdot \frac{\delta}{\delta \boldsymbol{u}} J[\boldsymbol{u}, B].$$

Here, α represents the learning rate.
(3) Update the adversarial sample set B as described in Algorithm 1 or 2.
End: if the stopping criteria are satisfied.

In addition to the best-response and better-response approaches, the a-GRAPE algorithms can be designed more flexibly such that it still works efficiently when the uncertainty vector $\boldsymbol{\epsilon}$ is of a large scale. For example, we may also perform the minimization process in a relaxed manner. As described in Algorithm 3, we may update the control by only a few gradient-descent iterations (or stop at some prescribed error threshold) without having to reach the ultimate minimum, which responds better but not best to the adversarial samples. The maximization part can be done either with the best-response or better-response approaches. In this regard, the SCP algorithm [33] can be considered as a special case of the relaxed better-response approach with fixed sampled uncertainties and carefully selected learning rates. It is noteworthy that as the minimization and maximization are related to the fidelity and the robustness, respectively, the balance between the two performance indexes can be adjusted flexibly via relaxing the two optimization processes.

To assess the feasibility of this idea, we apply the relaxed best-response (with $s = 5$, $n = 20$, and $m = 20$) and relaxed better-response (with $r = 0.25$, $n = 30$, and $m = 20$) a-GRAPE algorithms to the same three-qubit example. As shown in Fig. 6, where the control robustness is evaluated by the cumulative probability functions, the relaxed a-GRAPE algorithms can also greatly outperform the b-GRAPE algorithm. Compared to their unrelaxed counterparts, the relaxed best-response and relaxed better-response a-GRAPE algorithms are not only faster, but also more robust in the high-precision regime (e.g., near the infidelity level 10^{-4}).

In the comparisons above, although the a-GRAPE approaches outperform the b-GRAPE algorithm, the

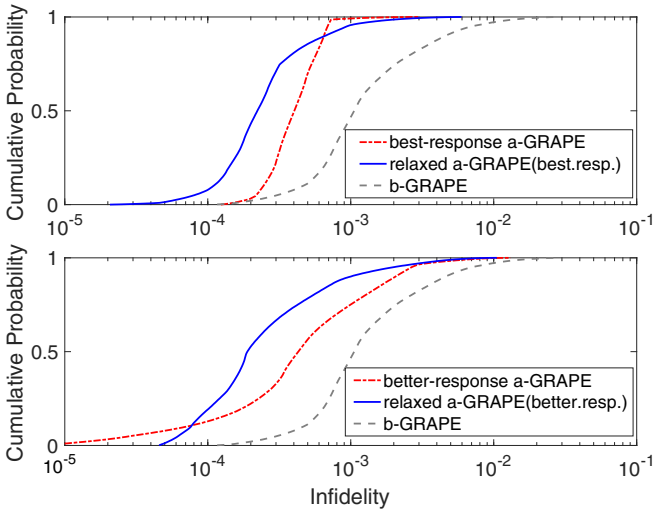

FIG. 6. The cumulative probability vs infidelity under the corresponding controls used in Fig. 5 and the ones obtained from the relaxed best-response approach run by 10 995 rounds and the relaxed better-response algorithm run by 3850 rounds.

corresponding computational overhead is also higher. To better illustrate the advantage of the a-GRAPE algorithms, especially the relaxed a-GRAPE approaches, we compare the algorithms with the same wall time (24 hours) for different initial controls. The results are shown in Fig. 7. The cdf curves show that the two relaxed a-GRAPE approaches (in particular, the relaxed best-response a-GRAPE approach) usually perform much better than the b-GRAPE algorithm. Here, we note that the adopted controls in the relaxed a-GRAPE approaches are not the ones optimized in the last iteration, but the ones that lead to the lowest average infidelity.

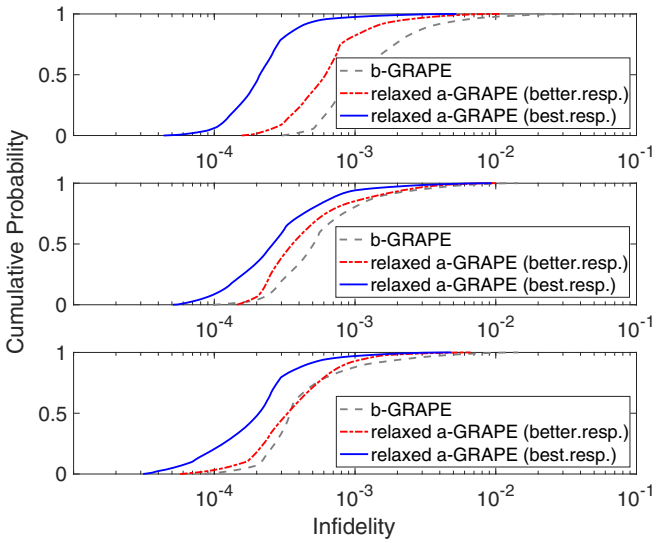

FIG. 7. The cumulative probability vs infidelity under the corresponding controls optimized by the three algorithms for different initial seeds.

Algorithm 4. b-GRAPE

Initialize:
an initial control $\boldsymbol{u}^{(0)}$;
an initial momentum $\boldsymbol{v}^{(0)} = \mathbf{0}$;
a set minibatch size n_{mb}.

Repeat:
(1) Randomly select a subset $\mathcal{S}_k$ of the uncertainty sample with $|\mathcal{S}|_k = n_{mb}$, where k denotes the current number of round.
(2) Update the control by

$$\boldsymbol{u}^{(k)} = \boldsymbol{u}^{(k-1)} + \boldsymbol{v}^{(k)},$$

with

$$\boldsymbol{v}^{(k)} = \lambda \boldsymbol{v}^{(k-1)} - \frac{\alpha}{n_{mb}} \sum_{\boldsymbol{\epsilon}\in\mathcal{S}_k} \frac{\delta}{\delta \boldsymbol{u}} L[\boldsymbol{u}^{(k-1)}, \boldsymbol{\epsilon}].$$

Here, α represents the learning rate and the weight parameter λ is chosen to be 0.9 in this paper.

End: if the stopping criteria are satisfied.

V. CONCLUSION

We have proposed a family of adversarial learning algorithms, including the best-response and better-response approaches, for the design of robust controls for quantum systems. The algorithms are subject to the optimization of worst-case gate infidelity, which can be treated and resolved from a game-theoretic perspective. Numerical simulations show that these a-GRAPE algorithms can achieve high control robustness. In particular, the best-response approach can effectively suppress the error over a larger domain at a satisfactory level of precision, but in the extremely high-precision regime, the better-response approach is superior. Both approaches have their regimes of practical utility depending on the application-specific requirement of the control. We also demonstrate that a family of a-GRAPE algorithms can be expanded by relaxing the maximization and minimization processes.

It should be noted that although a-GRAPE usually outperforms b-GRAPE, the computational burden is also heavier, and the tuning of algorithm parameters (e.g., memory sizes, batch sizes, ratios or learning rates) is application specific. Further studies are needed to deduce how to optimize the choices of these parameters, or even in an adaptive fashion.

As we remarked, we did not require the existence of Nash equilibria (NE) as an appropriate adversarial algorithm can enhance the robustness no matter whether or not the NE exists. In our simulations, it appears that a mixed NE is more likely approached in the best-response a-GRAPE with a larger memory size. From a theoretical perspective, a better understanding of the existence of a NE will be useful to attain. This topic will be explored in future studies.

ACKNOWLEDGMENTS

R.B.W. acknowledges the support of the National Key R&D Program of China (Grants No. 2018YFA0306703 and No. 2017YFA0304304) and NSFC (Grants No. 61833010 and No. 61773232). H.R. acknowledges the support of the National Science Foundation (Grant No. CHE-1763198).

APPENDIX: b-GRAPE ALGORITHM

The b-GRAPE algorithm presented in [21] is a stochastic gradient algorithm. The optimization process follows the gradient evaluated with randomly chosen batches of samples, so that the uncertainties can be effectively used to improve the robustness. Here, "b" stands for the "batch". The b-GRAPE algorithm is described in Algorithm 4.

[1] D. Castelvecchi, Quantum computers ready to leap out of the laboratory in 2017, Nature (London) **541**, 9 (2017).

[2] X. Dai, Y. Deng, X. Peng, and Y. Jin, Quantum-dot light-emitting diodes for large-area displays: Towards the dawn of commercialization, Adv. Mater. **29**, 1607022 (2017).

[3] M. Mohseni, P. Read, H. Neven, S. Boixo, and V. Denchev, Commercialize early quantum technologies, Nature (London) **543**, 171 (2017).

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

[5] 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).

[6] M. Dalgaard, F. Motzoi, J. J. Sorensen, and J. F. Sherson, Global optimization of quantum dynamics with alphazero deep exploration, npj Quantum Inf. **6**, 6 (2020).

[7] K. Bergmann, H. Theuer, and B. W. Shore, Coherent population transfer among quantum states of atoms and molecules, Rev. Mod. Phys. **70**, 1003 (1998).

[8] K. Bergmann, N. V. Vitanov, and B. W. Shore, Perspective: Stimulated Raman adiabatic passage: The status after 25 years, J. Chem. Phys. **142**, 170901 (2015).

[9] N. V. Vitanov, A. A. Rangelov, B. W. Shore, and K. Bergmann, Stimulated Raman adiabatic passage in physics, chemistry, and beyond, Rev. Mod. Phys. **89**, 015006 (2017).

[10] L. Viola, E. Knill, and S. Lloyd, Dynamical Decoupling of Open Quantum Systems, Phys. Rev. Lett. **82**, 2417 (1999).

[11] K. Khodjasteh and L. Viola, Dynamically Error-Corrected Gates for Universal Quantum Computation, Phys. Rev. Lett. **102**, 080501 (2009).

[12] K. Khodjasteh and D. A. Lidar, Performance of deterministic dynamical decoupling schemes: Concatenated and periodic pulse sequences, Phys. Rev. A **75**, 062310 (2007).

[13] K. Khodjasteh and D. A. Lidar, Fault-Tolerant Quantum Dynamical Decoupling, Phys. Rev. Lett. **95**, 180501 (2005).

[14] E. Zahedinejad, J. Ghosh, and B. C. Sanders, Designing High-Fidelity Single-Shot Three-Qubit Gates: A Machine-Learning Approach, Phys. Rev. Appl. **6**, 054005 (2016).

[15] Z. Ehsan, G. Joydip, and B. C. Sanders, High-Fidelity Single-Shot Toffoli Gate via Quantum Control, Phys. Rev. Lett. **114**, 200502 (2015).

[16] J.-S. Li and N. Khaneja, Control of inhomogeneous quantum ensembles, Phys. Rev. A **73**, 030302 (2006).

[17] J. Zhang, L. Greenman, X. Deng, and K. B. Whaley, Robust control pulses design for electron shuttling in solid-state devices, IEEE Trans. Control Syst. Technol. **22**, 2354 (2014).

[18] C. Chen, D. Dong, R. Long, I. R. Petersen, and H. A. Rabitz, Sampling-based learning control of inhomogeneous quantum ensembles, Phys. Rev. A **89**, 023402 (2014).

[19] D. Dong, M. A. Mabrok, I. R. Petersen, Q. Bo, C. Chen, and H. Rabitz, Sampling-based learning control for quantum systems with uncertainties, IEEE Trans. Control Syst. Technol. **23**, 2155 (2013).

[20] C. Wu, B. Qi, C. Chen, and D. Dong, Robust learning control design for quantum unitary transformations, IEEE Trans. Cybernet. **47**, 4405 (2017).

[21] R.-B. Wu, H. Ding, D. Dong, and X. Wang, Learning robust and high-precision quantum controls, Phys. Rev. A **99**, 042327 (2019).

[22] H.-J. Ding and R.-B. Wu, Robust quantum control against clock noises in multiqubit systems, Phys. Rev. A **100**, 022302 (2019).

[23] A. A Stoorvogel, The robust H_2 control problem: A worst-case design, IEEE Trans. Autom. Control **38**, 1358 (1993).

[24] M. G. Yoon, V. A. Ugrinovskii, and I. R. Petersen, On the worst-case disturbance of minimax optimal control, Automatica **41**, 847 (2005).

[25] P. Shi, Gain-scheduled worst-case control on nonlinear stochastic systems subject to actuator saturation and unknown information, J. Optim. Theory Appl. **156**, 844 (2013).

[26] A. S. Poznyak, T. E. Duncan, B. Pasik-Duncan, and V. G. Boltyansky, Robust stochastic maximum principle for multi-model worst-case optimization, Int. J. Control **75**, 1032 (2002).

[27] J. G. B. Beumee and H. Rabitz, Robust optimal control theory for selective vibrational excitation in molecules: A worst case analysis, J. Chem. Phys. **97**, 1353 (1992).

[28] H. Zhang and H. Rabitz, Robust optimal control of quantum molecular systems in the presence of disturbances and uncertainties, Phys. Rev. A **49**, 2241 (1994).

[29] A. Koswara and R. Chakrabarti, Robustness of controlled quantum dynamics, Phys. Rev. A, **90**, 043414 (2014).

[30] M. R. James, H. I. Nurdin, and I. R. Petersen, H_∞ control of linear quantum stochastic systems, IEEE Trans. Autom. Control **53**, 1787 (2007).

[31] A. I. Maalouf and I. R. Petersen, Time-varying H_∞ control for a class of linear quantum systems: A dynamic game approach, Automatica **48**, 2908 (2012).

[32] Aline I. Maalouf and Ian R. Petersen, Finite horizon H_∞ control for a class of sampled-data linear quantum systems, Intl. J. Robust Nonlin. Control **27**, 2292 (2017).

[33] Robert L. Kosut, Matthew D. Grace, and C. Brif, Robust control of quantum gates via sequential convex programming, Phys. Rev. A **88**, 052326 (2013).

[34] J. V. Neumann, O. Morgenstern, and H. W. Kuhn, *Theory of Games and Economic Behavior* (Princeton University Press, Princeton, NJ, 2007).

[35] C. Szegedy, W. Zaremba, I. Sutskever, J. Bruna, D. Erhan, I. Goodfellow, and R. Fergus, Intriguing properties of neural networks, arXiv:1312.6199.

[36] U. Shaham, Y. Yamada, and S. Negahban, Understanding adversarial training: Increasing local stability of supervised

models through robust optimization, Neurocomputing **307**, 195 (2018).

[37] I. J. Goodfellow, J. Shlens, and C. Szegedy, Explaining and harnessing adversarial examples, arXiv:1412.6572.

[38] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, Generative adversarial nets, in *Advances in Neural Information Processing Systems* (Curran Associates, Inc., NY, 2014), pp. 2672–2680.

[39] D. Fudenberg and D. K. Levine, *The Theory of Learning in Games*, Vol. 2 (MIT Press, Cambridge, MA, 1998).

[40] H. Rabitz, M. Hsieh, and C. Rosenthal, Landscape for optimal control of quantum-mechanical unitary transformations, Phys. Rev. A **72**, 052337 (2005).

[41] B. Russell, H. Rabitz, and R.-B. Wu, Control landscapes are almost always trap free: A geometric assessment, J. Phys. A: Math. Theor. **50**, 205302 (2017).