Letter

Breaking adiabatic quantum control with deep learning

Yongcheng Ding,[1,2,*] Yue Ban,[2,3] José D. Martín-Guerrero,[4,†] Enrique Solano,[1,2,5,6]
Jorge Casanova,[2,5,‡] and Xi Chen[1,2,§]

[1]*International Center of Quantum Artificial Intelligence for Science and Technology (QuArtist) and Department of Physics, Shanghai University, 200444 Shanghai, China*
[2]*Department of Physical Chemistry, University of the Basque Country UPV/EHU, Apartado 644, 48080 Bilbao, Spain*
[3]*College of Materials Science and Engineering, Shanghai University, 200444 Shanghai, China*
[4]*IDAL, Electronic Engineering Department, University of Valencia, Avgda. Universitat s/n, 46100 Burjassot, Valencia, Spain*
[5]*IKERBASQUE, Basque Foundation for Science, Plaza Euskadi 5, 48009 Bilbao, Spain*
[6]*IQM, Nymphenburgerstrasse 86, 80636 Munich, Germany*

(Received 14 September 2020; revised 22 March 2021; accepted 22 March 2021; published 9 April 2021)

In the noisy intermediate-scale quantum era, optimal digitized pulses are requisite for efficient quantum control. This goal is translated into dynamic programming, in which a deep reinforcement learning (DRL) agent is gifted. As a reference, shortcuts to adiabaticity (STA) provide analytical approaches to adiabatic speedup by pulse control. Here, we select the single-component control of qubits, resembling the ubiquitous two-level Landau-Zener problem for gate operation. We aim at obtaining fast and robust digital pulses by combining the STA and DRL algorithm. In particular, we find that DRL leads to robust digital quantum control with the operation time bounded by quantum speed limits dictated by STA. In addition, we demonstrate that robustness against systematic errors can be achieved by DRL without any input from STA. Our results introduce a general framework of digital quantum control, leading to a promising enhancement in quantum information processing.

DOI: 10.1103/PhysRevA.103.L040401

Introduction. For many decades, quantum control has been concerned with the efficient manipulation of physical and chemical processes on the atomic and molecular scale, with various applications ranging from photochemistry to quantum information sciences [1,2]. Specifically, how to implement fast and robust qubit gates with externally controllable parameters is required for realizing universal fault-tolerant quantum computing in physical platforms based on superconducting qubits and trapped ions [3]. Quantum error correction, for instance, has been developed to reduce the noise or imperfection coming from the environment and control parameters themselves, in the implementation of applications for noisy intermediate-scale quantum (NISQ) gate-based computers [4–6].

Two-level systems, hereafter referred to as qubit systems, are the basic units of digital quantum computing. Thus, several studies have been devoted to produce distinct methods for the precise quantum control of qubits with external fields. These include resonant pulses [7], adiabatic passages [8], composite pulses [9–12], pulse-shape engineering [13–15], and further optimizations [2,16–21]. Among these frameworks, "shortcuts to adiabaticity" (STA) [22,23] shares the concept that breaks the adiabatic regime, and it leads to fast nonadiabatic state evolution by combining both merits of resonant pulses and adiabatic passages. Specifically, inverse engineering, as one of the STA techniques, is applied to design the superadiabatic state evolution along the dynamical modes, emanated from the Lewis-Riesenfeld invariant, with appropriate boundary conditions [24]. Thus, the freedom left in the inverse engineering further allows suppressing the inevitable systematic errors such as amplitude noise and dephasing noise, by incorporating other techniques of optimal control [15,25,26], dynamical decoupling [27], and supervised machine learning [28,29].

At the same time, along with the development of deep learning in many areas [30–33], deep neural network and deep reinforcement learning (DRL) has been used for different applications in physics [34–40]. DRL works as a promising method for searching optimal control pulses for fast and robust quantum state preparation [41–43], gate operation [44], and the quantum Szilard engine [45]. More specifically, DRL enhances reinforcement learning, which is a key branch of classical machine learning widely applied to control tasks; this enhancement comes from the use of deep learning in the key aspects of reinforcement learning. Recent works have studied the application of DRL to quantum control [46–53]. Therefore, we find that it is meaningful to compare DRL with STA for a better understanding of both. We believe that this study will lead to more feasible applications in manipulating superconducting transmon qubits [54], Bose-Einstein condensates [55], and quantum dots [56], in which systematic errors, stochastic noise, and experimental constraint are of significance.

*jonzen.ding@gmail.com
†jose.d.martin@uv.es
‡jcasanovamar@gmail.com
§xchen@shu.edu.cn

To this aim, we explore fast and robust quantum control for qubit operation by combining the STA and DRL methods. Strictly, we focus on the single-component control of qubits, being similar to the two-level Landau-Zener (LZ) problem [54–56], with the designed time-dependent frequency sweep. We show the first salient result: Smooth pulses of STA are analytically engineered with a clarified quantum speed limit (QSL), and further optimized with respect to various types of noises, imperfections, and physical constraints, particularly in the feasible experiments when lacking flexibility. More importantly, we look for such quantum control with DRL, benchmarking it by connecting STA. We find that the DRL agent explores digital shortcuts for the same task, resulting in similar characteristics of robustness, when the operation time of STA is used as a hint. As an extension, we train the agent, which is the part of DRL in charge of taking control actions, without any input from STA for the efficient control by suppressing various systematic errors. In our numerical simulations, we further observe that the DRL agent is capable of achieving efficient quantum control with satisfying features. We reckon that one can improve the performance of the framework by fine tuning in an interactive DRL environment with quantum noise, resulting in potential applications in noisy intermediate-scale quantum (NISQ) systems.

Inverse engineering and optimization of STA. Consider the coherent manipulation of a single qubit, whose Hamiltonian reads

$$H(t) = \frac{\hbar}{2}[\Omega\sigma_x + \Delta(t)\sigma_z], \tag{1}$$

where the Rabi frequency Ω is fixed, while the detuning $\Delta(t)$ is time varying. Equation (1) appears in, e.g., the Xmon transmon qubit [54], in Bose-Einstein condensates within accelerated optical lattices [55], and in quantum dot charge qubits [56]. According to the Lewis-Riesenfeld (LR) theory [57], one can construct a dynamical invariant $I(t) = \frac{\hbar}{2}\Omega_0\sum_\pm |\phi_\pm(t)\rangle\langle\phi_\pm(t)|$, where its eigenstates are $|\phi_+(t)\rangle = [\cos(\frac{\theta}{2})e^{-i\frac{\beta}{2}}, \sin(\frac{\theta}{2})e^{i\frac{\beta}{2}}]^{\mathrm{T}}$, and $|\phi_-(t)\rangle = [\sin(\frac{\theta}{2})e^{-i\frac{\beta}{2}}, -\cos(\frac{\theta}{2})e^{i\frac{\beta}{2}}]^{\mathrm{T}}$, with Ω_0 being an arbitrary constant frequency that keeps $I(t)$ in units of energy. Here, the time-dependent angles $\theta \equiv \theta(t)$ and $\beta \equiv \beta(t)$ parametrize the trajectory of an evolving state on the Bloch sphere. The solution of the time-dependent Schrödinger equation is described by the superposition of $|\phi_\pm(t)\rangle$. More specifically, $|\Psi(t)\rangle = \sum_\pm c_\pm \exp(i\gamma_\pm)|\phi_\pm(t)\rangle$, with c_n being constants, and the LR phases $\gamma_\pm$ are calculated as

$$\gamma_\pm(t) = \pm\frac{1}{2}\int_0^t \left(\frac{\dot\theta\cot\beta}{\sin\theta}\right)dt'. \tag{2}$$

The condition $dI(t)/dt \equiv \partial I(t)/\partial t + (1/i\hbar)[I(t), H(t)] = 0$ yields the following auxiliary equations,

$$\dot\theta = -\Omega\sin\beta, \tag{3}$$

$$\dot\beta = -\Omega\cot\theta\cos\beta + \Delta(t), \tag{4}$$

which enables the desired state to evolve along the dynamical mode, $|\phi_\pm(t)\rangle$. In previous works, a STA control framework facilitates optimization over errors and noise under symmetrical constraints [25,26] with two tunable parameters, i.e., the Rabi frequency Ω and detuning Δ, that hold Eqs. (3) and (4). However, in certain quantum platforms tunability on Ω and Δ is not available. For instance, in superconducting Xmon transmon qubits [54] control on the detuning is preferred.

To adapt these requirements, we shall apply the inverse engineering method to design the angle parameter θ for tailoring the time-dependent detuning $\Delta(t)$. Accordingly, we substitute Eq. (4) and its derivative into Eq. (4), leading to the following expression,

$$\Delta(t) = -\frac{\ddot\theta}{\Omega\sqrt{1-\left(\frac{\dot\theta}{\Omega}\right)^2}} + \Omega\cot\theta\sqrt{1-\left(\frac{\dot\theta}{\Omega}\right)^2}, \tag{5}$$

with constant Rabi frequency Ω. This allows us to drive the state evolution along one of the dynamical mode, $|\phi_+(t)\rangle$, by a single component within a finite short time T, constrained by QSL. First of all, the following boundary conditions are imposed,

$$\theta(0) = 0, \quad \theta(T) = \pi. \tag{6}$$

These determine a qubit flip from $|0\rangle$ to $|1\rangle$. Second, the protocol can be optimized for canceling systematic errors, as environmental fluctuations and deviations on the control parameters are unavoidable in an experimental scenario. To this end, we consider the errors in Rabi frequency and detuning, i.e., $\Omega \to \Omega(1+\delta_\Omega)$ and $\Delta(t) \to \Delta(t) + \delta_\Delta$, and write down the transition probability, keeping the first-order term in the time-dependent perturbation theory [58],

$$P = \frac{\hbar^2}{4}\left|\int_0^T \langle\Psi_-(t)|(\delta_\Omega\Omega\sigma_x + \delta_\Delta\sigma_z)|\Psi_+(t)\rangle\right|^2, \tag{7}$$

with $|\Psi_\pm(t)\rangle = e^{i\gamma_\pm(t)}|\phi_\pm(t)\rangle$ being the two orthogonal dynamical modes of the invariant. Plugging Eqs. (3) and (4) into Eq. (7), we obtain the following condition for error cancellation,

$$\left|\int_0^T dt e^{i\eta(t)}(\delta_\Delta\sin\theta - i2\delta_\Omega\dot\theta\sin^2\theta)\right| = 0, \tag{8}$$

where $\eta(t) = 2\gamma_+(t)$, yielding $\dot\eta = \dot\theta\cot\beta/\sin\theta$ by combining with Eq. (2).

Inspired by Ref. [15], a global phase $\eta(t) = 2\gamma_+(t)$ in the integral (8) is expanded as

$$\eta(t) = 2\theta + \alpha_1\sin(2\theta) + \alpha_2\sin(4\theta) + \cdots + \alpha_n\sin(2n\theta), \tag{9}$$

such that we get $\sin\beta = -1/\sqrt{1+4M^2\sin^2\theta}$, with $M = 1 + \sum_n n\alpha_n\cos(2n\theta)$. As a result, by solving Eq. (3) with the given coefficients α_n and the initial condition $\theta(0) = 0$, one obtains the corresponding θ that evolves to $\theta(T) = \pi$, with T bounded by QSL time as (see Refs. [18,21])

$$\Omega T = \int_0^\pi d\theta\sqrt{1+4M^2\sin^2\theta} \geqslant \pi. \tag{10}$$

This protocol allows robust qubit flipping from $|0\rangle$ to $|1\rangle$ with arbitrary series coefficients for pulse engineering. In principle, by introducing the free parameters α_n in Eq. (9), one can nullify the above integral (8), such that the errors in both σ_x and σ_z terms can be simultaneously suppressed. For simplicity, here we may set $\delta_\Omega = 0$ or $\delta_\Delta = 0$, to independently discuss each

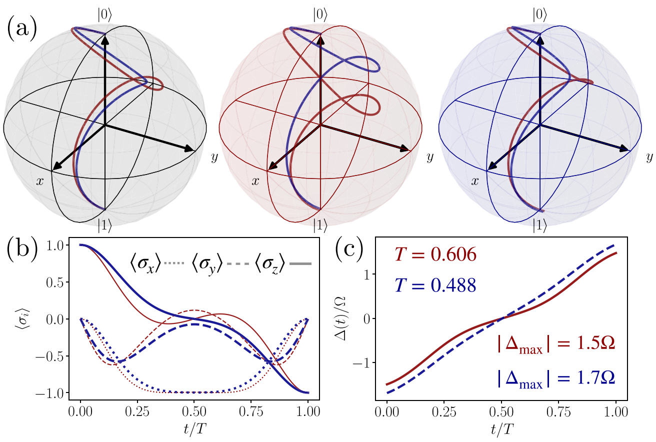

FIG. 1. (a) STA of qubits on Bloch spheres, where the continuous pulses of LZ type are designed from Eq. (5) by using the ansatz of $\theta(t)$ (11). With $a = 0.604$ (red curve) and 0.728 (blue curve), the qubit without any errors (gray sphere) is flipped from $|0\rangle$ to $|1\rangle$, being robust against Δ and Ω errors, respectively, which is proved by numerical simulations under the amplitudes of $\delta_\Delta = \pm 0.1\Delta_{\max}$ (red and blue curves on the red sphere) and $\delta_\Omega = \pm 0.1$ (red and blue curves on the blue sphere). (b) Evolution of expectations on different directions, corresponding to the trajectory on the gray sphere, by minimizing the Δ error (red thin curve) or the Ω error (blue thick curve). We fix the Rabi frequency $\Omega = 20 \times 2\pi$ MHz, resulting in an operation time $T = 60.6$ ns and $T = 48.8$ ns for STA protocols canceling previous errors. (c) The corresponding smooth pulses are depicted as well.

error source, without presuming the ratio of error amplitudes. We find out that the Ω error and Δ error can be eliminated with only a first-order expansion of $\eta(t)$, resulting in $\alpha_1 = -1$ and $\alpha_1 = -1.74$, respectively.

While robust quantum control can be achieved within the preceding framework, we notice that one can hardly predict the shape of the detuning as well as its adjustable range. For example, the detuning against the Δ error has a maximum amplitude of more than 3 Ω with abrupt changes at the edges of the operation (see Supplemental Material [58]). For more feasible implementations, we prefer smooth controls with a detuning $\Delta(t)$ that does not oscillate drastically. Thus, we propose the following ansatz for θ,

$$\theta(t) = \frac{\Omega T}{a}\left[as - \frac{\pi^2}{2}(1-s)^2 + \frac{\pi^2}{3}(1-s)^3 + \cos(\pi s) + A\right], \tag{11}$$

where $s = t/T$, $A = \pi^2/6 - 1$, and $T = -\pi a/[(2 - a - \pi^2/6)\Omega]$ are found by using the boundary conditions (6), while $a > 2 - \pi^2/6$ being a free parameter. This ansatz leads to a detuning $\Delta(t)$ that grows almost linearly during the operation time, resembling the original LZ scheme, but with finite values at $t = 0$ and $t = T$ (see Fig. 1 and Supplemental Material [58]). Accordingly, this protocol provides with $a = 0.604$ and 0.728 for robust qubit flipping against the Δ and Ω errors, respectively, resulting in an operation time $T = 60.6$ ns and $T = 48.8$ ns, of the same order of the characteristic gate time for a superconducting Xmon transmon qubit [54].

Deep reinforcement learning. Even though STA has enabled fast and robust control, we consider other numerical methods for more complicated cases, e.g., if we are only allowed to drive the quantum states with a given number N of detuned pulses within a fixed time. This task is indeed combinational optimization (i.e., maximizing robustness with an optimal configuration of discretized pulses) which is equivalent to dynamic programming, namely, a decision problem of multiple steps. Although the complexity of dynamic programming grows exponentially with the number of steps, one can still approximately solve it, e.g., with an artificial neural network (ANN) approach, which naturally leads us to the concept of DRL. Actually, when one talks about DRL two main approaches arise. The first one is based on the use of deep learning to approximate the dynamic programming solution. The second approach deals with the so-called deep policy networks, i.e., the ability to test many different control systems in parallel. We focus on the former approach. In this framework, the assumption of DRL is that there exists an optimal policy π, giving an action $\mathbf{a}(t_i)$ for any observable state $\mathbf{s}(t_i)$ to complete a certain task in a system. This state-action relation can be characterized by a function $\pi(\mathbf{a}|\mathbf{s})$, which can be approximated by a deep ANN. State $\mathbf{s}(t_i)$ is encoded in input variables within zero to one, being observed by the agent ANN. After propagations between the layers and nonlinear activations of the ANN nodes, the output layer gives an action $\mathbf{a}(t_i)$, which evolves the system to the next state $\mathbf{s}(t_{i+1})$ within one time step.

An environment consists of these equations, governing the evolution, as well as rewarding the agent. Optimizers tune the parameters of the ANN according to rewards, leading to a well-trained agent to provide optimized actions for completing the task. Thus, we notice that the concept of inverse engineering from STA, i.e., choosing an ansatz to procure the protocol via auxiliary equations, is similar to DRL. One should design a reward function $r(\mathbf{s}, t_i)$ to educate the agent. An adequate reward function accelerates the convergence of the DRL algorithm, preventing the agent from getting stuck in trivial solutions or cheating by repetitive actions.

In our practice of DRL, we renormalize the tunable detuning range $[-\Delta_{\max}, \Delta_{\max}]$ into $\tilde{\Delta} \in [0, 1]$, which is the encoded action at time step t_i: $\tilde{\Delta}(t_i) = [\Delta(t_i) + \Delta_{\max}]/2\Delta_{\max}$. The state consists of the element of the density matrix $|\rho_{22}(t_i)|$, renormalized detuning $\tilde{\Delta}(t_{i-1})$ as the action of last time step, and the current system time i/N. Here, we use the Liouville–von Neumann equation instead of the Schrödinger equation for further generalization to the Lindblad master equation with the quantum noise involved elsewhere.

The training strategy of the agent is shown in Fig. 2(a), where extensive details regarding this strategy can be found in the Supplemental Material [58]. We find that DRL agents converge to (sub)optimal solutions by approximating the policy that maximizes the artificial reward. This enables DRL to explore various types of quantum control, requiring an adequate reward function for our task. For example, we notice that proximal policy optimization (PPO) [59], instead of trust region policy optimization (TRPO) [60,61], rapidly learns the time-optimal solution, i.e., a resonant flat π pulse, with a trivial reward function $r(t_i) = |\rho_{22}(t_i)| - 1$. However, for LZ-type control, we pretrain the agent with $r(t_i) = -|\tilde{\Delta}(t_i) - \frac{i-1}{N-1}|$,

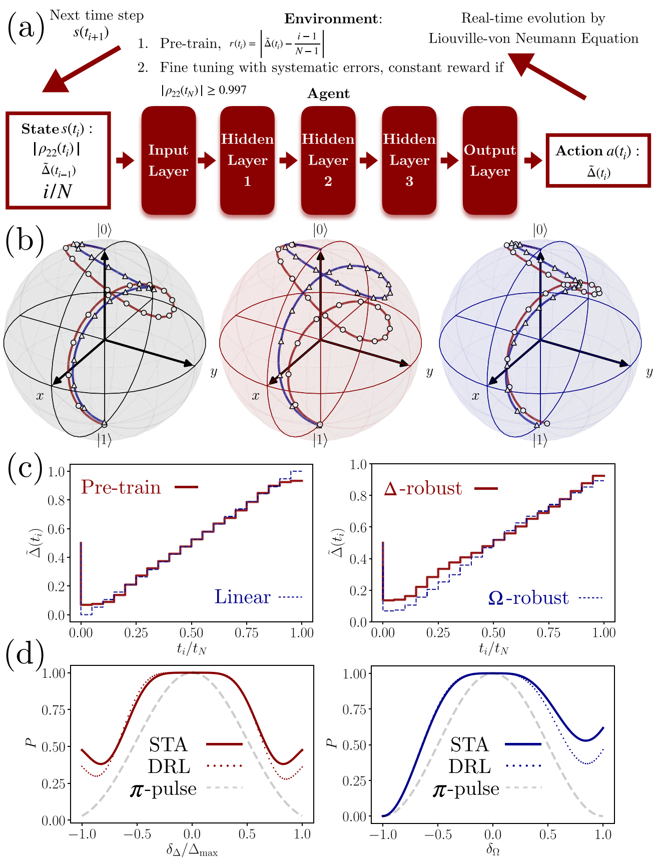

FIG. 2. (a) Scheme of DRL approach to quantum control with the LZ scheme for one time step in training. An ANN (agent) of three hidden layers observes a state, which encodes the physical information of the qubit. An action is outputted, evolving the system for one time step, resulting in its state of the next time step. The environment rewards or punishes the agent by an artificially designed reward function, enabling the agent to learn by an accumulation of them. (b) State evolution of the qubit on Bloch spheres, driven by digital pulses from DRL within 20 time steps, where the parameters are the same as those in Fig. 1. (c) Renormalized detuning pulses after pretraining for control of LZ type, and fine tuning according to systematic errors and populations. Following Fig. 1(c), $\Delta_{\max}$ are set to 1.5 and 1.7 Ω for the Δ and Ω errors, respectively. (d) Final population of state $|1\rangle$ vs relative systematic errors. Protocols designed by STA and obtained from DRL are both robust against systematic errors with similar features, comparing to the resonant flat π pulse as a time-optimal solution.

rewarding the linear growth of detuning. Pretraining can filter other strategies and accelerate the convergence as well. Later, we reward the agent by a constant for fine tuning under random systematic errors characterized by a Gaussian distribution, if $|\rho_{22}| > 0.997$ at the final time step.

By using TENSORFORCE [62] and QUTIP [63], we first investigate if the DRL agent learns digital quantum control resembling STA [see Fig. 1(c)], with an operation time $T =$ 60.6 and 48.8 ns split equally by using 20 pulses as the only hint. The control calculated by STA eliminates the error transition, which bounds the upper limit of robustness. Thus, the performance of DRL can be easily benchmarked. We find out that the DRL agent manages to flip the qubit against systematic errors by digital pulses [see Figs. 2(b) and 2(c)], which are not the coarse-grained analog controls.

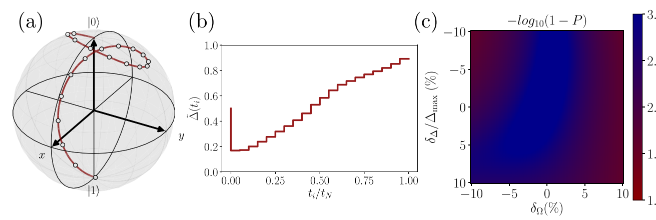

FIG. 3. (a) State evolution during an arbitrary time $t_N = 55$ ns in an ideal system. (b) Digital pulses given by DRL, where $\Delta_{\max}$ is chosen as 1.6 Ω without knowledge of STA. (c) Deviation of population under both Ω and Δ errors.

In Fig. 2(d), we compare the robustness of STA, DRL, and the flat π pulse against Δ/Ω errors. The agent discovers digital quantum control with the same feature of STA, which is quite satisfying for approaching the theoretical maximum of robustness. Inspired by this preliminary result, we further employ DRL for the sake of searching for robust digital control against both Ω and Δ errors. In this scenario, DRL is more straightforward since the inverse engineering from STA does not work perfectly even with more free parameters, depending on a certain proportion of δ_Δ and δ_Ω. With the same training strategy, we educate the DRL agent for qubit flipping against both types of errors, which is shown in Fig. 3. It is worthwhile to mention that we set the operation time and tunable range of detuning without any knowledge from STA. The agent learns its goal, resulting in populations exceeding 0.99 within $\delta_\Delta/\Delta_{\max}, \delta_\Omega \in [-0.1, 0.1]$.

Discussion. Although one may argue that other numerical algorithms such as GRAPE and CRAB are also capable of completing similar tasks, in our practice, we perceive that these gradient algorithms have their limitations. Some of them easily get stuck to local minima, being far from optimal solutions because of initial configurations. For the specific problem presented in this Letter, one might not meet the local minimum problem but unwanted control pulses instead, even if the trial control is initialized according to the goal of realizing LZ-type control (see Supplemental Material [58]). By contrast, we obtain the control pulses of LZ type with the appropriate reward function design. DRL algorithms tune the ANN for maximizing the accumulative reward, achieving an almost optimal result with a similar robustness feature of analytical STA, which bounds the upper limit by the second-order perturbation theory. Moreover, it is proved that the global minimum (maximum) can be found in ANN with a gradient descent [64,65]. Studies also verify that any complex ANN can be reduced to one with much smaller sizes without loss of performance [66], massively saving training time. This theoretical research from the computer science community will continuously improve DRL, widening its application in quantum control. We also emphasize the critical issue that we are far from exploiting the power of DRL in this Letter because of physical constraints. In other scenarios of applied DRLs, states for agents are easily observed. For example, states of Go, real-time strategy (RTS) games, and automatic driving are already digitized during information collection processes.

However, the precise observation of states in quantum control destroys the system immediately, requiring enormous copies for learning by real devices or controlling an unknown system by trained agents. Therefore, we use digital pulses given by the agent in an ideal environment for all systematic errors. In other words, we forbid the agent to observe any state during evaluation, which is also for a fair comparison between fixed STA and the DRL. In spite of the constraint we imposed, the cost of observing the state for the DRL agent in quantum control can be reduced by experimental strategies inspired by machine learning [67–69]. If an agent is allowed to observe the state before each time step, it should dynamically change the action according to optimized policies as it does for other tasks. This improvement also requires a precise measurement of the quantum state. Otherwise, the performance can be even worse than fixed quantum control. An experiment in a trapped-ion system was recently carried out [70], showing perfect agreement with our theoretical prediction. Although we do not introduce superoperators in the Lindblad master equation, our protocols are still robust in the real experimental environment by choosing a two-level system in which energy levels are sensitive to external noises.

Conclusion. By comparison, we figure out that STA provides a well-optimized analytical method for designing fast and robust quantum control, in terms of inverse engineering with parameter variations. Beyond that, it enhances the performance of DRL, providing knowledge of the optimal evolving time, accelerating its convergence as well. On the other hand, DRL also shows its capability of learning physics with artificially designed reward functions. The agent obtains digital pulses with complex constraints, which prominently eliminate systematic errors. Pulses designed for more complicated cases are also satisfying even without field knowledge. Last but not least, our DRL framework can be extended to multiqubit systems [71,72] and quantum noise without many efforts, for the application of robust control optimization in quantum algorithms [73,74] based on NISQ devices. Moreover, agent training in an interactive environment with real quantum devices as fine tuning can lead to a significant enhancement, if the experimental time cost is acceptable.

Acknowledgments. This work is partially supported from NSFC (12075145), STCSM (2019SHZDZX01-ZX04, 18010500400, and 18ZR1415500), Program for Eastern Scholar, HiQ funding for developing STA&QAOA (YBN2019115204), QMiCS (820505) and OpenSuperQ (820363) of the EU Flagship on Quantum Technologies, Spanish Government PGC2018-095113-B-I00 (MCIU/AEI/FEDER, UE), Basque Government IT986-16, EU FET Open Grant Quromorphic (828826), as well as EPIQUS (899368). X.C. acknowledges Ramn y Cajal program (RYC-2017-22482). J.C. acknowledges the Ramón y Cajal program (RYC2018- 025197-I) and the EUR2020-112117 project of the Spanish MICINN, as well as support from the UPV/EHU through the grant EHUrOPE.

[1] D. D'Alessandro, *Introduction to Quantum Control and Dynamics* (Chapman & Hall/CRC, Boca Raton, FL, 2007).
[2] S. J. Glaser, U. Boscain, T. Calarco, C. P. Koch, W. Köckenberger, R. Kosloff, I. Kuprov, B. Luy, S. Schirmer, T. Schulte-Herbrüggen, D. Sugny, and F. K. Wilhelm, Eur. Phys. J. D **69**, 279 (2015).
[3] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, U.K., 2010).
[4] J. Kelly, R. Barends, A. G. Fowler, A. Megrant, E. Jeffrey, T. C. White, D. Sank, J. Y. Mutus, B. Campbell, Y. Chen, Z. Chen, B. Chiaro, A. Dunsworth, I.-C. Hoi, C. Neill, P. J. J. O'Malley, C. Quintana, P. Roushan, A. Vainsencher, J. Wenner *et al.*, Nature (London) **519**, 66 (2015).
[5] M. Takita, A. W. Cross, A. D. Córcoles, J. M. Chow, and J. M. Gambetta, Phys. Rev. Lett. **119**, 180501 (2017).
[6] S. Rosenblum, P. Reinhold, M. Mirrahimi, L. Jiang, L. Frunzio, and R. J. Schoelkopf, Science **361**, 266 (2018).
[7] L. Allen and J. H. Eberly, *Optical Resonance and Two-Level Atoms* (Wiley, New York, 1975).
[8] P. Král, I. Thanopulos, and M. Shapiro, Rev. Mod. Phys. **79**, 53 (2007).
[9] M. H. Levitt, Prog. Nucl. Magn. Reson. Spectrosc. **18**, 61 (1986).
[10] K. R. Brown, A. W. Harrow, and I. L. Chuang, Phys. Rev. A **70**, 052318 (2004).
[11] B. T. Torosov, S. Guérin, and N. V. Vitanov, Phys. Rev. Lett. **106**, 233001 (2011).
[12] X. Rong, J. Geng, F. Shi, Y. Liu, K. Xu, W. Ma, F. Kong, Z. Jiang, Y. Wu, and J. Du, Nat. Commun. **6**, 8748 (2015).
[13] M. Steffen and R. H. Koch, Phys. Rev. A **75**, 062326 (2007)
[14] E. Barnes and S. Das Sarma, Phys. Rev. Lett. **109**, 060401 (2012).
[15] D. Daems, A. Ruschhaupt, D. Sugny, and S. Guérin, Phys. Rev. Lett. **111**, 050404 (2013).
[16] T. Caneva, M. Murphy, T. Calarco, R. Fazio, S. Montangero, V. Giovannetti, and G. E. Santoro, Phys. Rev. Lett. **103**, 240501 (2009).
[17] S. Guérin, V. Hakobyan, and H. R. Jauslin, Phys. Rev. A **84**, 013423 (2011).
[18] G. C. Hegerfeldt, Phys. Rev. Lett. **111**, 260501 (2013).
[19] A. Garon, S. J. Glaser, and D. Sugny, Phys. Rev. A **88**, 043422 (2013).
[20] L. Van Damme, Q. Ansel, S. J. Glaser, and D. Sugny, Phys. Rev. A **95**, 063403 (2017).
[21] C. Arenz, B. Russell, D. Burgarth, and H. Rabitz, New J. Phys. **19**, 103015 (2017).
[22] D. Guéry-Odelin, A. Ruschhaupt, A. Kiely, E. Torrontegui, S. Martínez-Garaot, and J. G. Muga, Rev. Mod. Phys. **91**, 045001 (2019).
[23] E. Torrontegui, S. Ibánez, S. Martínez-Garaot, M. Modugno, A. del Campo, D. Guéry-Odelin, A. Ruschhaupt, X. Chen, and J. G. Muga, Adv. At., Mol., Opt. Phys. **62**, 117 (2013).
[24] X. Chen, A. Ruschhaupt, S. Schmidt, A. del Campo, D. Guéry-Odelin, and J. G. Muga, Phys. Rev. Lett. **104**, 063002 (2010).
[25] A. Ruschhaupt, X. Chen, D. Alonso, and J. G. Muga, New J. Phys. **14**, 093040 (2012).

[26] X.-J. Lu, X. Chen, A. Ruschhaupt, D. Alonso, S. Guérin, and J. G. Muga, Phys. Rev. A **88**, 033406 (2013).
[27] C. Munuera-Javaloy, Y. Ban, X. Chen, and J. Casanova, Phys. Rev. Applied **14**, 054054 (2020).
[28] E. Zahedinejad, J. Ghosh, and B. C. Sanders, Phys. Rev. Applied **6**, 054005 (2016).
[29] B.-J. Liu, X.-K. Song, Z.-Y. Xue, X. Wang, and M.-H. Yung, Phys. Rev. Lett. **123**, 100501 (2019).
[30] V. Mnih, K. Kavukcuoglu, D. Silver, A. A. Rusu, J. Veness, M. G. Bellemare, A. Graves, M. Riedmiller, A. K. Fidjeland, G. Ostrovski, S. Petersen, C. Beattie, A. Sadik, I. Antonoglou, H. King, D. Kumaran, D. Wierstra, S. Legg, and D. Hassabis, Nature (London) **518**, 529 (2015).
[31] V. Mnih, K. Kavukcuoglu, D. Silver, A. Graves, I. Antonoglou, D. Wierstra, and M. Riedmiller, arXiv:1312.5602.
[32] D. Silver, A. Huang, C. J. Maddison, A. Guez, L. Sifre, G. Driessche, J. Schrittwieser, I. Antonoglou, V. Panneershelvam, M. Lanctot, S. Dieleman, D. Grewe, J. Nham, N. Kalchbrenner, I. Sutskever, T. Lillicrap, M. Leach, K. Kavukcuoglu, T. Graepel, and D. Hassabis, Nature (London) **529**, 484 (2016).
[33] D. Silver, T. Hubert, J. Schrittwieser, I. Antonoglou, M. Lai, A. Guez, M. Lanctot, L. Sifre, D. Kumaran, T. Graepel, T. Lillicrap, K. Simonyan, and D. Hassabis, arXiv:1712.01815.
[34] G. Carleo and M. Troyer, Science **355**, 602 (2017).
[35] T. Fösel, P. Tighineanu, T. Weiss, and F. Marquardt, Phys. Rev. X **8**, 031084 (2018).
[36] N. Yoshioka and R. Hamazaki, Phys. Rev. B **99**, 214306 (2019).
[37] A. Nagy and V. Savona, Phys. Rev. Lett. **122**, 250501 (2019).
[38] M. J. Hartmann and G. Carleo, Phys. Rev. Lett. **122**, 250502 (2019).
[39] F. Vicentini, A. Biella, N. Regnault, and C. Ciuti, Phys. Rev. Lett. **122**, 250503 (2019).
[40] R. Iten, T. Metger, H. Wilming, L. del Rio, and R. Renner, Phys. Rev. Lett. **124**, 010508 (2020).
[41] B. M. Henson, D. K. Shin, K. F. Thomas, J. A. Ross, M. R. Hush, S. S. Hodgman, and A. G. Truscott, Proc. Natl. Acad. Sci. USA **115**, 13216 (2018).
[42] X.-M. Zhang, Z. Wei, R. Asad, X.-C. Yang, and X. Wang, npj Quantum Inf. **5**, 85 (2019).
[43] T. Haug, W.-K. Mok, J.-B. You, W. Zhang, C. E. Png, and L.-C. Kwek, Mach. Learn.: Sci. Technol. **2**, 01LT02 (2021).
[44] Z. An and D. L. Zhou, Europhys. Lett. **126**, 60002 (2019).
[45] V. B. Sørdal and J. Bergli, Phys. Rev. A **100**, 042314 (2019).
[46] M. Bukov, A. G. R. Day, D. Sels, P. Weinberg, A. Polkovnikov, and P. Mehta, Phys. Rev. X **8**, 031086 (2018).
[47] R. Porotti, D. Tamascelli, M. Restelli, and E. Prati, Commun. Phys. **2**, 1 (2019).
[48] I. Paparelle, L. Moro, and E. Prati, Phys. Lett. A **384**, 126266 (2020).
[49] M. Y. Niu, S. Boixo, V. N. Smelyanskiy, and H. Neven, npj Quantum Inf. **5**, 33 (2019).
[50] M. Dalgaard, F. Motzoi, J. J. Sørensen, and J. Sherson, npj Quantum Inf. **6**, 6 (2020).
[51] X.-M. Zhang, Z.-W. Cui, X. Wang, and M.-H. Yung, Phys. Rev. A **97**, 052333 (2018).
[52] R.-B. Wu, H. Ding, D. Dong, and X. Wang, Phys. Rev. A **99**, 042327 (2019).
[53] M. Ostaszewski, J. A. Miszczak, L. Banchi, and P. Sadowski, Quantum Inf. Process. **18**, 126 (2019).
[54] J. M. Martinis and M. R. Geller, Phys. Rev. A **90**, 022307 (2014).
[55] M. G. Bason, M. Viteau, N. Malossi, P. Huillery, E. Arimondo, D. Ciampini, R. Fazio, V. Giovannetti, R. Mannella, and O. Morsch, Nat. Phys. **8**, 147 (2012).
[56] X. Wang, L. S. Bishop, J. P. Kestner, E. Barnes, K. Sun, and S. D. Sarma, Nat. Commun. **3**, 997 (2012).
[57] H. R. Lewis and W. B. Riesenfeld, J. Math. Phys. **10**, 1458 (1969).
[58] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevA.103.L040401 for further explanations and details of the calculation.
[59] J. Schulman, F. Wolski, P. Dhariwal, and A. Radford, arXiv:1707.06347.
[60] J. Schulman, S. Levine, P. Moritz, M. Jordan, and P. Abbeel, in *Proceedings of the 32nd International Conference on Machine Learning*, edited by F. Bach and D. Blei, Proceedings of Machine Learning Research Vol. 37 (PMLR, 2015), pp. 1889–1897.
[61] L. Engstrom, A. Ilyas, S. Santurkar, D. Tsipras, F. Janoos, L. Rudolph, and A. Madry, Implementation matters in deep RL: A case study on PPO and TRPO, in *International Conference on Learning Representations*, 2019 (unpublished).
[62] M. Schaarschmidt, A. Kuhnle, and K. Fricke, TENSORFORCE: A TENSORFLOW library for applied reinforcement learning, https://github.com/tensorforce/tensorforce.
[63] J. R. Johansson, P. D. Nation, and F. Nori, Comput. Phys. Commun. **183**, 1760 (2012).
[64] S. S. Du, J. D. Lee, H. Li, L. Wang, and X. Zhai, arXiv:1811.03804.
[65] Z. A. Zhu, Y. Li, and Z. Song, arXiv:1811.03962.
[66] H. Yu, S. Edunov, Y. Tian, and A. S. Morcos, arXiv:1906.02768.
[67] M. Krenn, M. Malik, R. Fickler, R. Lapkiewicz, and A. Zeilinger, Phys. Rev. Lett. **116**, 090405 (2016).
[68] M. Krenn, M. Erhard, and A. Zeilinger, Nat. Rev. Phys. **2**, 649 (2020).
[69] Y. Ding, J. D. Martín-Guerrero, M. Sanz, R. Magdalena-Benedicto, X. Chen, and E. Solano, Phys. Rev. Lett. **124**, 140504 (2020).
[70] M.-Z. Ai, Y. Ding, Y. Ban, J. D. Martín-Guerrero, J. Casanova, J.-M. Cui, Y.-F. Huang, X. Chen, C.-F. Li, and G.-C. Guo, arXiv:2101.09020
[71] Y. Chen, C. Neill, P. Roushan, N. Leung, M. Fang, R. Barends, J. Kelly, B. Campbell, Z. Chen, B. Chiaro, A. Dunsworth, E. Jeffrey, A. Megrant, J. Y. Mutus, P. J. J. O'Malley, C. M. Quintana, D. Sank, A. Vainsencher, J. Wenner, T. C. White *et al.*, Phys. Rev. Lett. **113**, 220502 (2014).
[72] J. Ghosh, A. Galiautdinov, Z. Zhou, A. N. Korotkov, J. M. Martinis, and M. R. Geller, Phys. Rev. A **87**, 022309 (2013).
[73] Z.-C. Yang, A. Rahmani, A. Shabani, H. Neven, and C. Chamon, Phys. Rev. X **7**, 021027 (2017).
[74] Y. Dong, X. Meng, L. Lin, R. Kosut, and K. B. Whaley, arXiv:1911.00789.