Application of machine learning methods to quantum control problems

Xin (Sunny) Wang

City University of Hong Kong

Department of Physics
物 | 理 | 學 | 系

香港城市大學
City University of Hong Kong
專業 創新 胸懷全球
Professional·Creative
For The World

Our recent works applying machine learning method to quantum control problems

Supervised learning:

- Neural-network-designed pulse sequences for robust control of singlet-triplet qubits,
  X.-C. Yang, M.-H. Yung, and XW,
  Phys. Rev. A **97**, 042324 (2018).
- Spin-qubit noise spectroscopy from randomized benchmarking by supervised learning,
  C. Zhang and XW,
  Phys. Rev. A **99**, 042316 (2019).

Reinforcement learning:

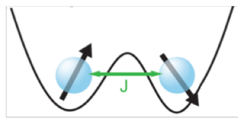
- Automatic spin-chain learning to explore the quantum speed limit, X.-M. Zhang, Z.-W. Cui, XW, and M.-H. Yung, *Phys. Rev. A* **97**, 052333 (2018).
- When reinforcement learning stands out in quantum control? A comparative study on state preparation, X.-M. Zhang, Z. Wei, A. Raza, X.-C. Yang, and XW, arXiv:1902.02157.
- Transferable control for quantum parameter estimation through reinforcement learning, H. Xu, J. Li, L. Liu, Y. Wang, H. Yuan, and XW, arXiv:1904.11298

Outline

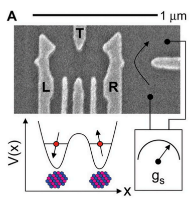
- Background: spin qubit, control, and noises.
- Constructing composite pulses by supervised learning
- Measuring noise spectra by supervised learning
- Quantum state preparation: a comparative study of reinforcement learning and other methods

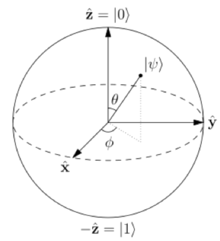
Singlet-triplet (S-T) qubit

- Encoded in $S_z = 0$ singlet and triplet states of two electrons, in a double-quantum-dot system.



$$|0\rangle = |T_0\rangle = \left(|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle\right)/\sqrt{2}$$
$$|1\rangle = |S\rangle = \left(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle\right)/\sqrt{2}$$

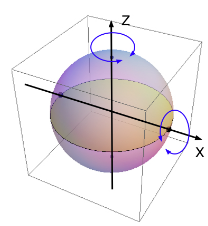




Petta *et al.*, *Science* **309**, 2180 (2005)

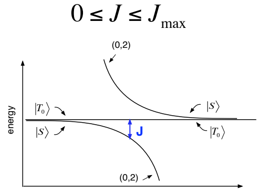
Hamiltonian of an S-T qubit

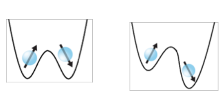
$$H = \frac{J(t)}{2}\sigma_z + \frac{h}{2}\sigma_x$$



z-rotation: exchange interaction J tuned by detuning (gate voltage)

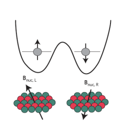
$$0 \le J \le J_{\max}$$





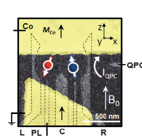
x-rotation: magnetic field gradient (two dots feel different B-field).

Cannot vary dynamically



"Nuclear Polarization"

Foletti et al, Nat. Phys. 2009

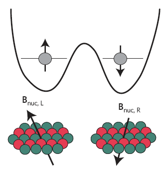


"Micromagnet"

Brunner et al, PRL 2011

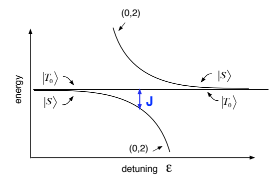
Noises

- **Nuclear (Overhauser) noise**: Fluctuations in the nuclear magnetic field surrounding the quantum dots



$$h \rightarrow h + \delta h$$

- **Charge noise**: Shift of energy levels induced by charged impurities



$$J(t) \rightarrow J(t) + J'(\varepsilon)\delta\varepsilon$$

Composite Pulses

- Expansion of *R(J,ϕ)* to first order in *δJ* and *δh*:

$$R(J,\phi) \equiv \exp\left[-i\left(\frac{h+\delta h}{2}\sigma_x + \frac{J+\delta J}{2}\sigma_z\right)\frac{\phi}{\sqrt{h^2+J^2}}\right] = \exp\left[-i\left(\frac{h}{2}\sigma_x + \frac{J}{2}\sigma_z\right)\frac{\phi}{\sqrt{h^2+J^2}}\right]$$
$$\times\Bigg\{I + i\frac{\delta h}{2(h^2+J^2)^{3/2}}\left[(-h^2\phi - J^2\sin\phi)\sigma_x + 2J\sqrt{h^2+J^2}\sin^2\frac{\phi}{2}\sigma_y + hJ(\sin\phi-\phi)\sigma_z\right] + \mathcal{O}(\delta h^2)$$
$$+ i\frac{\delta J}{2(h^2+J^2)^{3/2}}\left[hJ(\sin\phi-\phi)\sigma_x + 2h\sqrt{h^2+J^2}\sin^2\frac{\phi}{2}\sigma_y + (-J^2\phi - h^2\sin\phi)\sigma_z\right] + \mathcal{O}(\delta J^2)\Bigg\}$$

- **Goal**: Design an imperfect identity operation (being exact identity in absence of fluctuations)

$$\widetilde{I} = I + (a_1\sigma_x + a_2\sigma_y + a_3\sigma_z)\delta h + (b_1\sigma_x + b_2\sigma_y + b_3\sigma_z)\delta J$$

such that when R and $\widetilde{I}$ are applied back-to-back, the errors (at the first order) exactly cancel.

Kestner *et al.*, *PRL* **110**, 140502 (2013)

Composite Pulses

- One can perform a corrected $R(J,\phi)$ rotation as:

$$R(J,\phi)\times$$

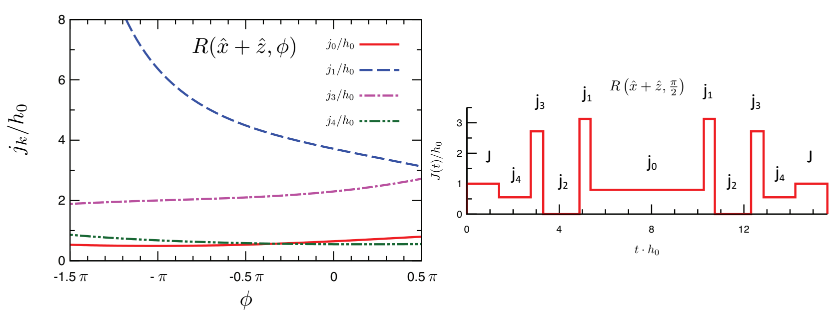
$$\boxed{\begin{aligned}&R(J,\pi-\frac{\phi}{2})R(j_4,\pi)R(j_3,\pi)R(j_2,\pi)R(j_1,\pi)R(j_0,4\pi)\\ &\qquad\times R(j_1,\pi)R(j_2,\pi)R(j_3,\pi)R(j_4,\pi)R(J,\pi+\frac{\phi}{2})\end{aligned}}$$

- Here the identity is formed by “interrupted 2π rotations” around certain axis

Kestner *et al.*, *PRL* **110**, 140502 (2013)

Composite Pulses

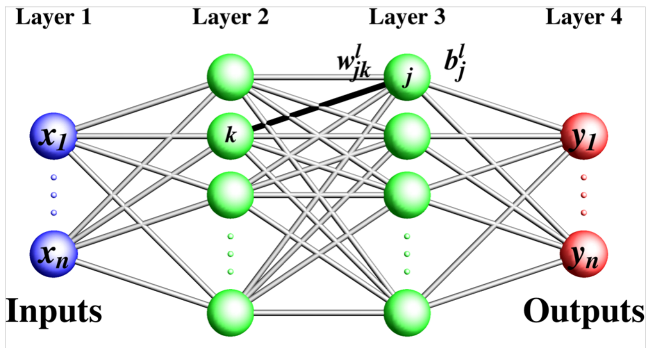
- Results for $J=h$:



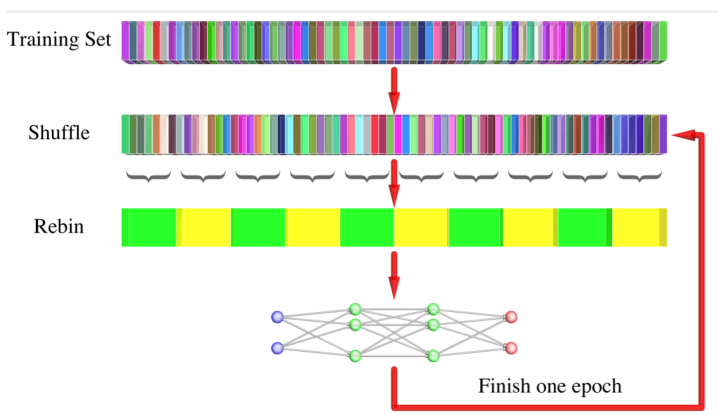
- The pulse sequence cancels both δJ and δh errors.
- Total duration: $\sim$ 14-16π

Kestner *et al.*, *PRL* **110**, 140502 (2013)

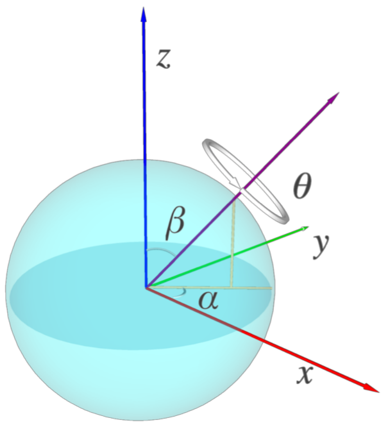
Supervised learning (Neural network)



Supervised learning (Training)



Decomposing Pulses (noiseless)



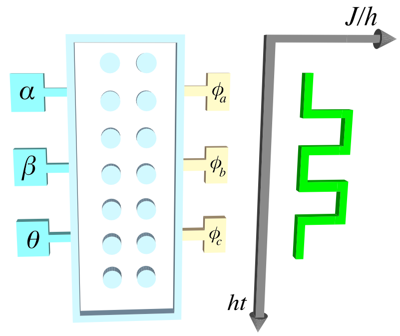


FIG. 1. (a) The Bloch sphere showing α, β, and θ and the Cartesian coordinates. (b) Schematics of the neural network, along with the input $\{\alpha,\beta,\theta\}$, the output $\{\phi_a,\phi_b,\phi_c\}$, and the constructed composite pulse sequence. The neural network contains two layers with N_n neurons each. t is an arbitrary time unit.

Decomposing Pulses (noiseless)

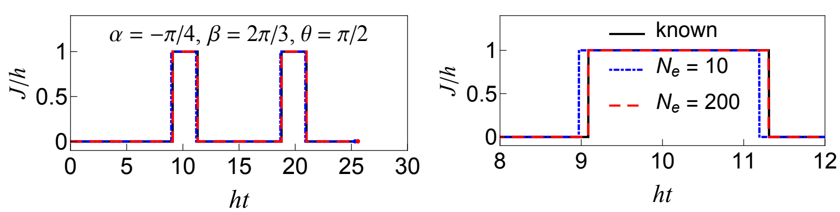
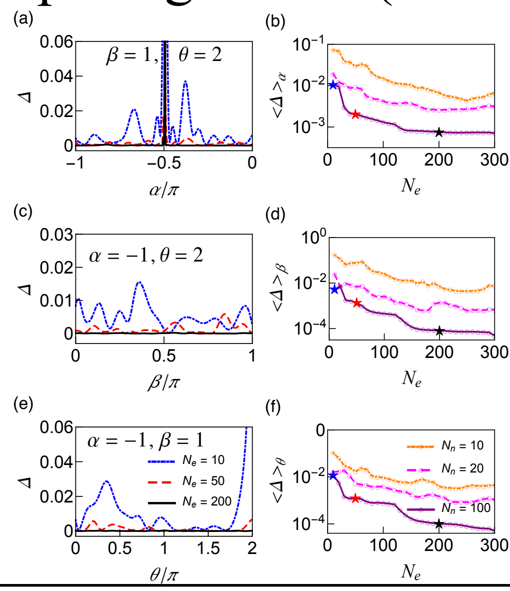
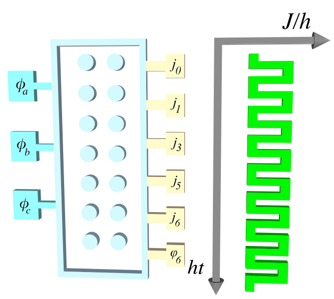


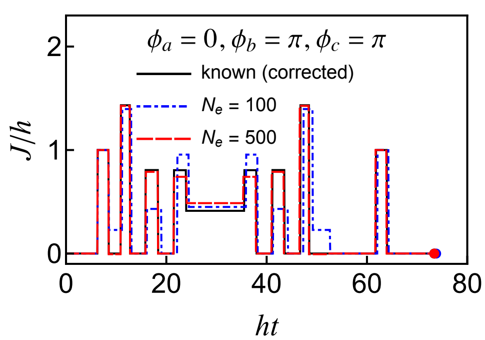
FIG. 2. Comparison of composite pulses predicted by the neural network and the known result for $\{\alpha,\beta,\theta\} = \{-\pi/4, 2\pi/3, \pi/2\}$. (a) The entire pulse sequence and (b) a zoom-in for the range $8 \leqslant ht \leqslant 12$. The black solid lines show the known results, while the predictions from the network after $N_e = 10$ and $N_e = 200$ epochs are shown as blue (gray) dash-dotted lines and red (gray) dashed lines, respectively. $N_n = 100$.

Decomposing Pulses (noiseless)

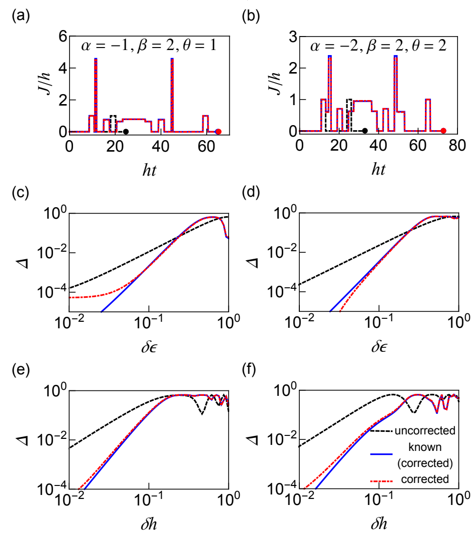


Composite Pulses (robust to noise)





Composite Pulses (robust to noise)



Summary

- A well-trained neural network is capable to generate composite pulse sequences that offer comparable ability in cancelling noise with traditional methods
- Possible to extend to more complicated noises (i.e. optimize for non-static noises etc.)

Measuring a 1/f noise spectrum using supervised learning

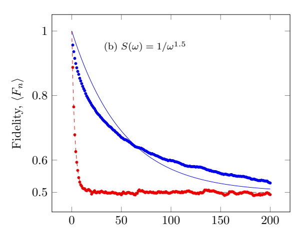
- Quantum gates evolve under noises.
- The spectra of the noises are sometimes important to quantum control.
    - Example: Dynamically corrected gates assume low frequency noises: they will not do well if noises have strong high frequency
- Therefore measuring the spectrum of the noise is important
- *Can we use supervised learning to measure noise spectra?*

Randomized Benchmarking

- Randomized benchmarking: a powerful technique to study the interaction between noises and gates
- Averaged fidelity decays exponentially with number of gates in the sequence as

$$\frac{1}{2}\left[1+(1-d)^n\right]$$

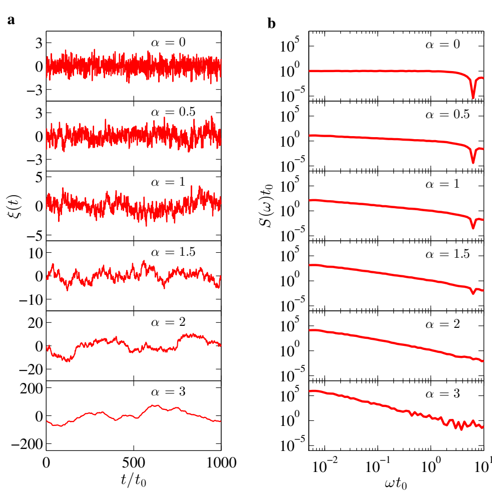
where *d* is the average error per gate



1/f noise

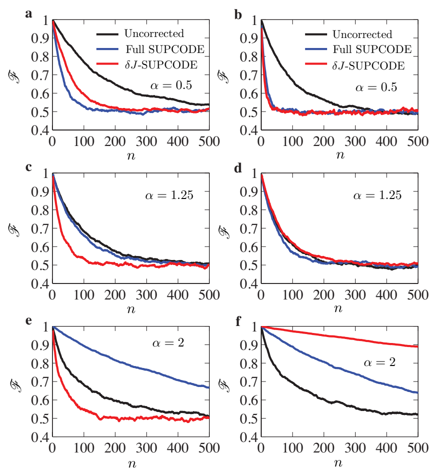
- We would like to study a well-contained problem: assuming the noises are of 1/*f* type, i.e. their spectra are
$$S(\omega) = A/(\omega t_0)^{\alpha}$$
- Experimentally measured value: $\alpha_h \approx 2.6$ (nuclear noise), $\alpha_J \approx 0.7$ (charge noise)
- Noises -> Gate error (straightforward);
  Gate error -> Noises (need help from neural network)
- Amplitude and exponent should be treated separately.

Noises with different exponents



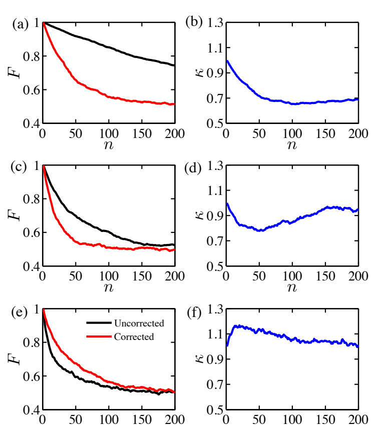
- α small: white-like noise. Noises present in all frequencies.
- α large: more correlation in noise; low frequency

Decay rate v.s. noise



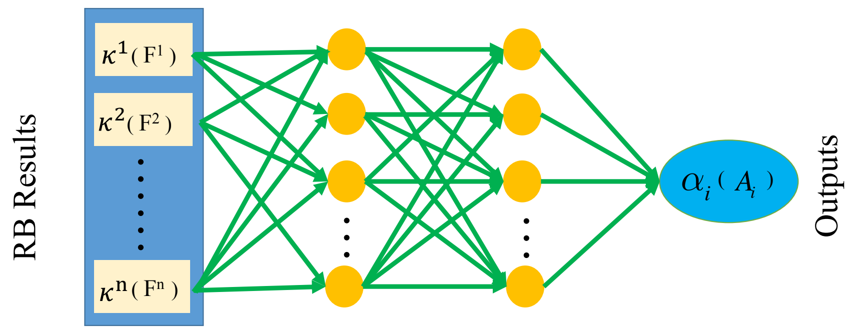
- The decay rates of the gate sequences depend on both the noise amplitude and the exponent.
- Dependence on amplitude: monotonic
- Dependence on exponent: Difference between corrected and uncorrected sequences.

Determining the noise exponent

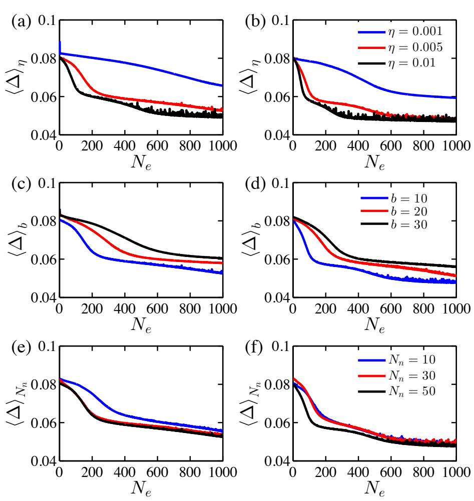


- The exponent is determined first.
- Input: ratio of RB results between corrected and uncorrected sequences.
- Output: α

Supervised learning

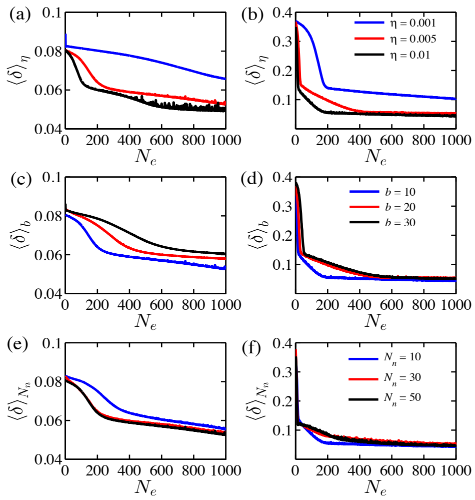


Determining the noise exponent (results)



- α is determined to an error of about ± 0.05. (Comparable to or exceeds other methods, e.g. based on dynamical decoupling)

Determining the amplitude



- After α is known, the amplitude A is then determined using RB results as inputs and A as outputs.
- A can be determined to a relative error of about 5%.

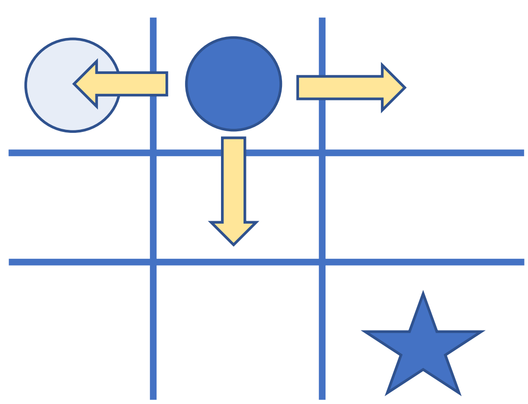
Summary

- The spectrum of 1/f noises can be measured. The two parameters (amplitude and exponent) should be measured separately with two neural networks which are trained with different data
- Amplitude can be measured to a relative error of 5%, while exponent can be measured to an absolute error of 0.05
- *Can we treat more complicated cases (determining a general noise spectrum)?* - More work required.

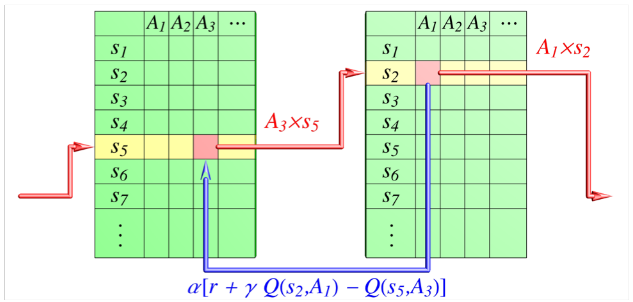
A comparative study of reinforcement learning and other methods

- Often we have the question: why some machine learning techniques perform better than other methods for certain problems?
- In this part, we take a simple problem, i.e. quantum state preparation, as an example, and compare the results from several different methods.

Reinforcement learning (introduction)

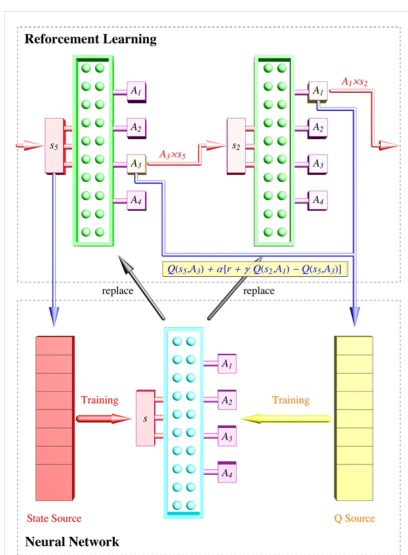


Reinforcement learning (Q-Learning)



- States (s) and actions (A). At each state, the probability to take a certain action is given by the Q-table.
- The values in the Q-table are adjusted throughout the learning process according to fidelity of the resulting state.

Sketch of Deep Q-Learning

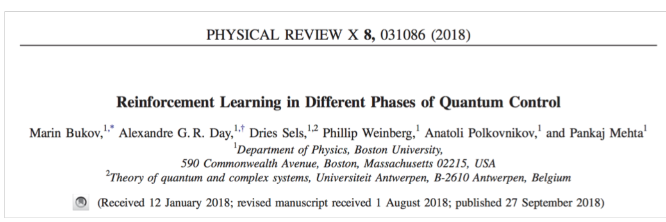


- In the Deep Q-learning, the Q-table is replaced by a neural network.
- Capable to treat more complicated situations.

Application of reinforcement learning to quantum state preparation

- Often we have the question: why some machine learning techniques perform better than other methods for certain problems?
- In this part, we take a simple problem, i.e. quantum state preparation, as an example, and compare the results from several different methods.

Prior work


PHYSICAL REVIEW X **8,** 031086 (2018)

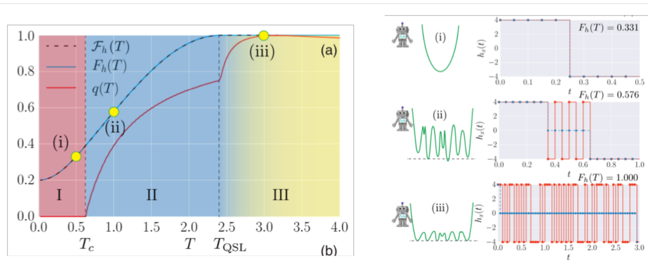
Reinforcement Learning in Different Phases of Quantum Control

Marin Bukov,[1,*] Alexandre G. R. Day,[1,†] Dries Sels,[1,2] Phillip Weinberg,[1] Anatoli Polkovnikov,[1] and Pankaj Mehta[1]

[1]*Department of Physics, Boston University, 590 Commonwealth Avenue, Boston, Massachusetts 02215, USA*

[2]*Theory of quantum and complex systems, Universiteit Antwerpen, B-2610 Antwerpen, Belgium*

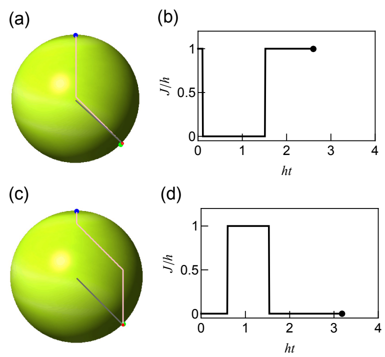
(Received 12 January 2018; revised manuscript received 1 August 2018; published 27 September 2018)



- Q-learning has been applied to the problem of preparing a desired quantum state.
- Nevertheless, the form of control field is complicated.

Prelim. Results on Quantum State Preparation from Deep Q-learning

$$H = \frac{J(t)}{2}\sigma_z + \frac{h}{2}\sigma_x$$

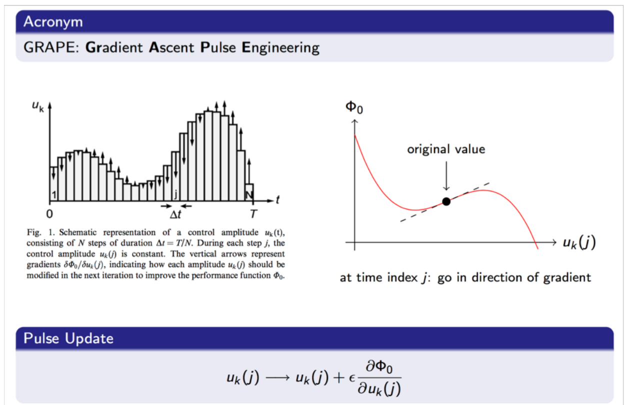


- Deep Q-learning is capable to produce simpler, more manageable pulses.
- *When would Deep Q-learning stand out, as compared to traditional methods?*

When deep Q-Learning (DQL) stands out?

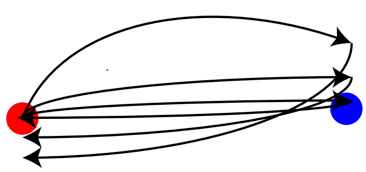
- We would like to study when DQL stands out as compared to other methods. This info will help us understand better the pros and cons of each method, allowing us to choose the most suitable method for a given problem.
- In our work, the following methods are compared:
  - Stochastic Gradient Descendant (SGD)
  - Krotov
  - Q-learning (QL)
  - Deep Q-learning (DQL)

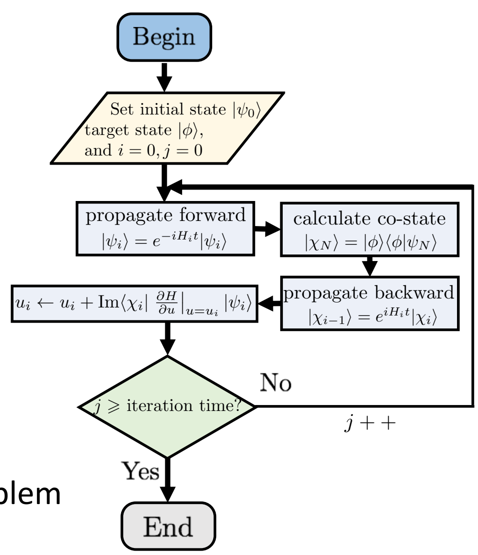
Stochastic Gradient Descendant (similar to GRAPE)



- Gradient-based: Need a well-defined "gradient"
- May not work well if the problem is discrete

Krotov method

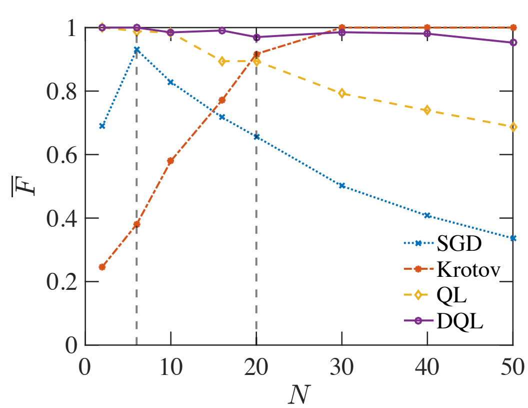




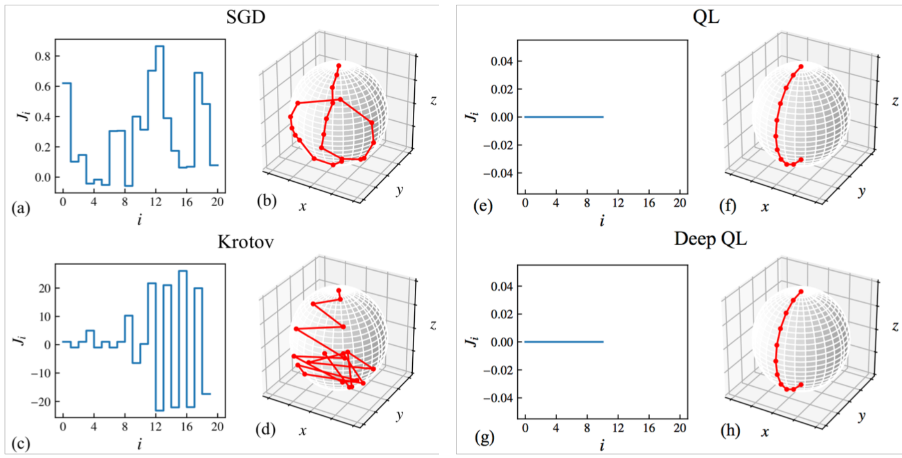
- Continuity of the evolution is critical
- May not work well if the problem is discrete

Control constraints considered

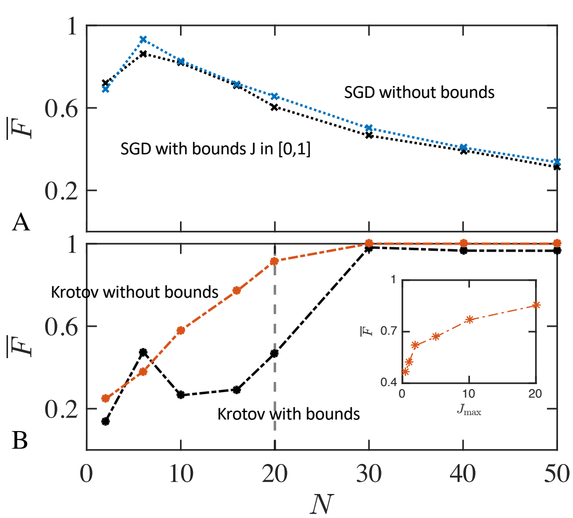
- In our work, we consider three types of control constraints:
  - Total # of time steps to complete the evolution, N
  - Strength of the control field being bounded, i.e. J is in [J_{min}, J_{max}]
  - Control field of J takes discrete values, i.e. J is only allowed to take M values between [J_{min}, J_{max}].



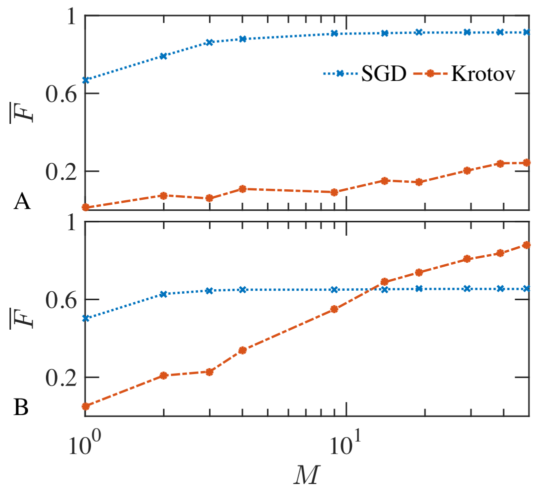
- Comparison of learning outcome vs N, maximum # of time steps in evolution (total time fixed). Target state |1>. # of iterations are fixed at 500 among different methods.
- Krotov favors continuous evolution.
- SGD fidelity low within given #iterations; may improve if more resources are given.
- QL better than SGD, but still not perfect within the #iteration.
- DQL performs well overall.



- Best learning outcome among 20 runs; N=20, target state |1>
- QL & DQL capable of finding the optimal control (in less than 20 steps); SGD and Krotov sticks to N=20 and have not found the optimal ones. (**adaptivity**)



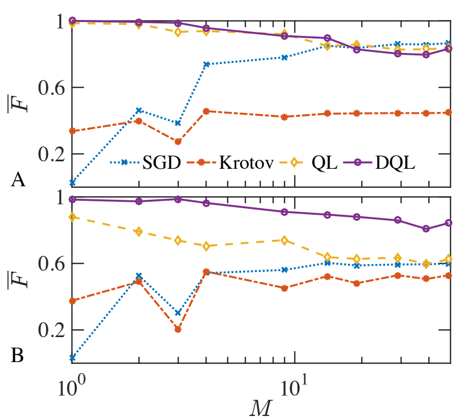
- Effect of bounds in control
- SGD insensitive of bounds in control: its updates are local
- Krotov perform badly with bounds in control: it requires large updates, and having narrow boundaries will compromise its effect.



N=6

N=20

- Effect of discrete control fields for SGD and Krotov methods
- SGD insensitive of M: intrinsically discretized.
- Krotov strongly favors continuous field, both large M and large N



N=6

N=20

- Effect of discrete control fields for SGD and Krotov methods for all four methods, with J bounded in [0,1]
- Krotov fail when J bounded; SGD favors small N and not sensitive to M
- DQL performs best overall.
- QL works better when N is small

Summary for this part

- Key advantage of QL and DQL: adaptivity (can adaptively optimize solutions)
- QL and DQL naturally work on problems with constraints, i.e. control with restricted ranges and discrete values.

Application of reinforcement learning to quantum parameter estimation

- Quantum parameter estimation: measuring the value of a parameter in the Hamiltonian with greater sensitivity, e.g.

$$\hat{H}(t) = \frac{1}{2}\omega_0\hat{\sigma}_3 + \boldsymbol{u}(t)\cdot\boldsymbol{\sigma}$$

- Key quantity: Quantum Fisher Information (QFI)

$$F(t) = \mathrm{Tr}\left[\hat{\rho}(t)\hat{L}_s^2(t)\right] \qquad \partial_\omega\hat{\rho}(t) = \frac{1}{2}\left[\hat{\rho}(t)\hat{L}_s(t) + \hat{L}_s(t)\hat{\rho}(t)\right]$$

- The QFI gives theoretical limit of the sensitivity of measurements, and is **dependent on the detailed evolution → optimal control may improve QFI**

Prior work

PHYSICAL REVIEW A **96**, 012117 (2017)

Quantum parameter estimation with optimal control

Jing Liu and Haidong Yuan*
Department of Mechanical and Automation Engineering, The Chinese University of Hong Kong, Shatin, Hong Kong
(Received 25 April 2016; published 14 July 2017)

A pivotal task in quantum metrology, and quantum parameter estimation in general, is to design schemes that achieve the highest precision with the given resources. Standard models of quantum metrology usually assume that the dynamics is fixed and that the highest precision is achieved by preparing the optimal probe states and performing optimal measurements. However, in many practical experimental settings, additional controls are usually available to alter the dynamics. Here we propose to use optimal control methods for further improvement of the precision limit of quantum parameter estimation. We show that, by exploring the additional degree of freedom offered by the controls, a higher-precision limit can be achieved. In particular, we show that the precision limit under the controlled schemes can go beyond the constraints put by the coherent time, which is in contrast with the standard scheme where the precision limit is always bounded by the coherent time.

Results from GRAPE (Liu and Yuan, PRA 2017)

- Dephasing dynamics:

$$\partial_t \hat{\rho}(t) = -i\left[\hat{H}(t), \hat{\rho}(t)\right] + \frac{\gamma}{2}\left[\hat{\sigma}_{\boldsymbol{n}}\hat{\rho}(t)\hat{\sigma}_{\boldsymbol{n}} - \hat{\rho}(t)\right] \qquad \hat{H}(t) = \frac{1}{2}\omega_0\hat{\sigma}_3 + \boldsymbol{u}(t)\cdot\boldsymbol{\sigma}$$

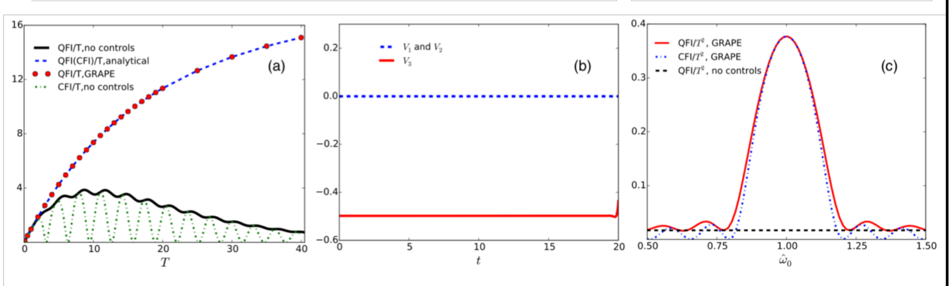
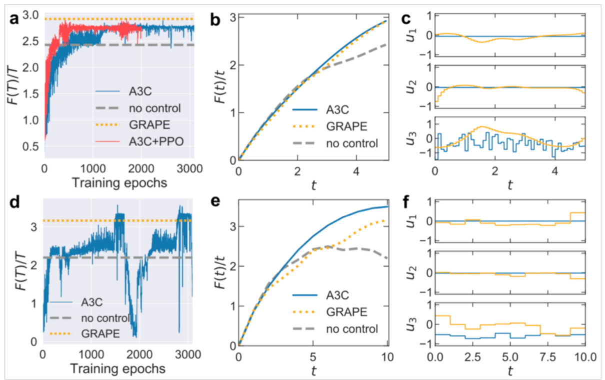


FIG. 4. Transverse dephasing: (a) The evolution of QFI (normalized by T) with and without controls. The red dots and black lines represent the QFI with and without controls, respectively (the wiggles in the black line is not numerical error but caused by some trigonometric functions in the QFI in this case). The dashed blue line is the analytical solution for the QFI (and CFI) under controls. The dash-dotted green line is the CFI without controls. Decay rate $\gamma = 0.1$, and the measurement for CFI is $\{|+\rangle\langle+|, |-\rangle\langle-|\}$. (b) Optimal controls obtained from GRAPE. (c) The enhanced QFI and CFI (normalized by T^2) as a function of $\hat{\omega}_0$. The solid red and dash-dotted blue lines represent the QFI and CFI under controls, respectively. The target time $T = 20$ and $\gamma = 0.2$. The dashed black line represents the value of QFI without controls. The true values of ω_0 in all panels are assumed to be 1.

Our results from RL



- We use a version of RL called Asynchronous Advantage Actor-Critic method (A3C)
- Sometimes, RL gives better QFI than GRAPE

Transferability

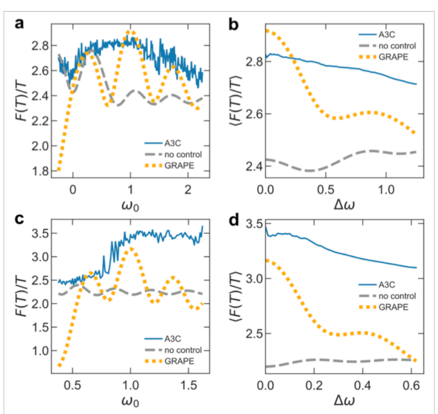


FIG. 3: Transferability of the control under dephasing dynamics. (a), (c): $F(T)/T$ v.s. ω_0 for three different methods. Note that the results from the GRAPE method are obtained using the pulses generated for $\omega_0 = 1$ only, while those from A3C are obtained using a neural network trained at $\omega_0 = 1$. (b), (d): average $F(T)/T$ in a range $[1-\Delta\omega, 1+\Delta\omega]$ corresponding to the results of (a) and (c) respectively. (a), (b): $\Delta T = 0.1$, $T = 5$; (c), (d): $\Delta T = 1$, $T = 10$.

$$\langle F(T)/T\rangle = \frac{1}{2\Delta\omega}\int_{1-\Delta\omega}^{1+\Delta\omega} F(T)/T \, d\omega.$$

- Transferability: the neural network trained at one parameter value can work reasonably well at other values
- No need to re-train the network. Only need to re-generate the control sequences at a different value. O(N)
- Compared to GRAPE: need to re-run GRAPE at each value. O(N^3)
- Especially advantageous when one has to measure an ensemble of systems with parameters distributed in a range

Transferability under other situations

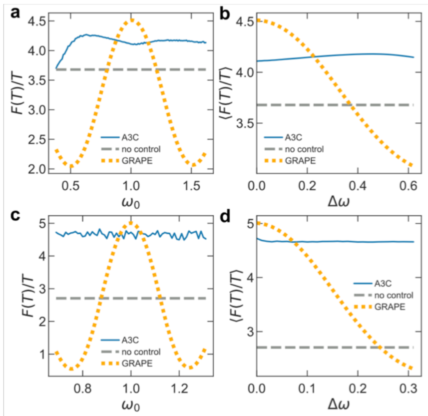


FIG. 5: Transferability of the control under spontaneous emission. (a), (c): $F(T)/T$ v.s. ω_0 for three different methods. Note that the results from the GRAPE method are obtained using the pulses generated for $\omega_0 = 1$ only, while those from A3C are obtained using a neural network trained at $\omega_0 = 1$. (b), (d): average $F(T)/T$ in a range $[1-\Delta\omega, 1+\Delta\omega]$ corresponding to the results of (a) and (c) respectively. (a), (b): $\Delta T = 0.1$, $T = 10$; (c), (d): $\Delta T = 1$, $T = 20$.

Summary for this part

- We have demonstrated that RL-based method can sometimes outperforms GRAPE.
- However, the most important advantage of RL-based methods is that it has the transferability

- Overall, we have shown that reinforcement learning, when applied to quantum control problems, can be more “efficient”:

  Adaptivity

  Transferability