Noise-Robust End-to-End Quantum Control using Deep Autoregressive Policy Networks (RL-QAOA)

Jiahao Yao, Paul Köttering, Hans Gundlach, Lin Lin and Marin Bukov

University of California, Berkeley

Aug 2021

MSML 2021

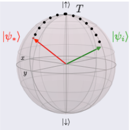
ArXiv: 2012.06701

Berkeley
UNIVERSITY OF CALIFORNIA

Background: Quantum Approximate Optimization Algorithm (QAOA)

- QAOA was first proposed by Farhi et al.[1]
  - Inspired by Quantum Adiabatic Algorithm (QAA)
  - originally, approximate solutions to combinatorial optimization problems, e.g. max-cut problem
  - the extension of QAOA: Quantum Alternating Operator Ansatz
- the idea of using the alternating gate sequences is quite generic, one simple kind of parameterized quantum circuits (PQCs), or variational circuit ansatz (VQCs).
- focus on the problem of state transfer (or state preparation) problem:

Given initial quantum state $|\psi_i\rangle \in \mathcal{H}$,
Find the control to reach target state $|\psi_*\rangle$



[1] Edward Farhi, Jeffrey Goldstone, and Sam Gutmann. "A Quantum Approximate Optimization Algorithm". In: *arXiv preprint arXiv:1411.4028* (2014).

state preparation problem

- Ising models

$$H = H_1 + H_2, \qquad H_1 = \sum_{i=1}^{N} J S^z_{i+1} S^z_i + h_z S^z_i, \quad H_2 = \sum_{i=1}^{N} h_x S^x_i,$$

- Heisenberg model

$$H = H_1 + H_2, \qquad H_1 = J \sum_{j=1}^{N} (S^x_{j+1} S^x_j + S^y_{j+1} S^y_j), \quad H_2 = \Delta \sum_{j=1}^{N} S^z_{j+1} S^z_j.$$

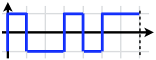
- intial state $|\psi_i\rangle = |\uparrow \cdots \uparrow\rangle$, and target state is the ground state of H
- **QAOA Ansatz**:

$$|\psi\rangle = U(\{\alpha_i, \beta_i\}_{i=1}^{p}) |\psi_i\rangle = e^{-iH_2\beta_p} e^{-iH_1\alpha_p} \cdots e^{-iH_2\beta_1} e^{-iH_1\alpha_1} |\psi_i\rangle .$$

- **Energy (Fidelity)**:

$$\mathcal{E}(\{\alpha_j, \beta_j\}_{j=1}^{p}) = \langle\psi_i| U^\dagger H U |\psi_i\rangle / N. \left(F(\{\alpha_i, \beta_i\}_{i=1}^{p}) = |\langle\psi_*| U |\psi_i\rangle|^2 \right)$$

- **Optimization**:



$$\{\alpha_i^*, \beta_i^*\}_{i=1}^{p} = \underset{\{\alpha_i, \beta_i\}_{i=1}^{p}}{\arg\min}\ \mathcal{E}(\{\alpha_j, \beta_j\}_{j=1}^{p}), \left(\underset{\{\alpha_i, \beta_i\}_{i=1}^{p}}{\operatorname{argmax}}\ F(\{\alpha_i, \beta_i\}_{i=1}^{p}) \right)$$

RL approach

RL approach was (first) introduced in literature[2] (tabular Q learning).

- RL is model-free/universal, adaptive, autonomous
- recent work using the gradient[3,4]: expand the gradient in terms of linear comb. of Pauli & importance sampling
- in the following discussion, we focus on the derivative free methods.
- **reward** ($R_t = 0, t < T$)

$$R_T = F(\{\alpha_i, \beta_i\}_{i=1}^p) = |\langle\psi_*|U(\{\alpha_i, \beta_i\}_{i=1}^p)|\psi_i\rangle|^2 .$$

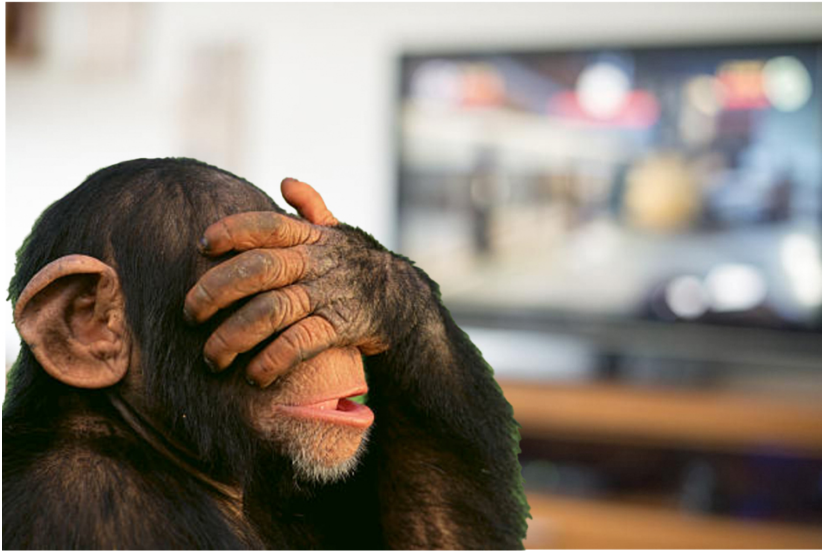
[2] Marin Bukov et al. "Reinforcement learning in different phases of quantum control". In: *Physical Review X* 8.3 (2018), p. 031086.

[3] Ryan Sweke et al. "Stochastic gradient descent for hybrid quantum-classical optimization". In: *Quantum* 4 (2020), p. 314.

[4] Aram Harrow and John Napp. "Low-depth gradient measurements can improve convergence in variational hybrid quantum-classical algorithms". In: *arXiv preprint arXiv:1901.05374v1* (2019).

Reinforcement Learning

RL4QAOA



CD driving method

Goal: introduces an auxiliary counter-diabatic (CD) Hamiltonian drive on top of a target Hamiltonian to suppress all transitions between eigenstates.

Consider a state $|\psi\rangle$, evolving under a time dependent Hamiltonian $H_0(\lambda(t))$

$$i\hbar\partial_t |\psi\rangle = H_0(\lambda(t)) |\psi\rangle , |\psi_i\rangle = |\psi_{\mathrm{GS}}(\lambda = 0)\rangle , |\psi_*\rangle = |\psi_{\mathrm{GS}}(\lambda = 1)\rangle \quad (1)$$

Go to the rotating frame: (Hamiltonian remains stationary)
unitary transformation $U(\lambda(t))$ — instantaneous eigenbasis of Hamiltonian
wave function ($|\tilde{\psi}\rangle = U(\lambda)|\psi\rangle$) satisfies effective Schrödinger equation:

$$i\hbar\partial_t|\tilde{\psi}\rangle = \left(\tilde{H}_0(\lambda(t)) - \dot{\lambda}\tilde{\mathcal{A}}_\lambda\right) |\tilde{\psi}\rangle \quad (2)$$

$$\tilde{H}_0(\lambda(t)) = U^\dagger H_0(\lambda(t))\, U, \tilde{\mathcal{A}}_\lambda = iU^\dagger \partial_\lambda U$$

CD driving method (Cont'd)

Specifically, the Hamiltonian picks up an extra contribution and becomes

$$H_0^{\text{eff}} = \tilde{H}_0 - \dot{\lambda}\tilde{\mathcal{A}}_\lambda, \tag{3}$$

The idea of the CD driving is to evolve the system with the Hamiltonian

$$H_{\text{CD}}(t) = H_0 + \dot{\lambda}\mathcal{A}_\lambda$$

In the moving frame $H_{\text{CD}}^{\text{eff}}(t) = \tilde{H}_0$ is stationary and no transitions occur.

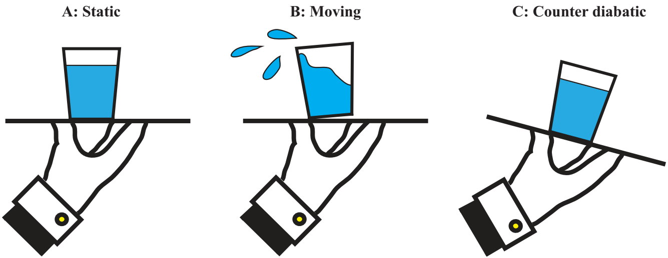
Transitionless Driving: counter-diabatic driving[5][6]

$$H_{\text{CD}}(t) = H_0(\lambda(t)) + \dot{\lambda}\mathcal{A}(\lambda(t)),\ \dot{\lambda}(0) = 0 \quad \dot{\lambda}(T) = 0$$

[5] Dries Sels and Anatoli Polkovnikov. "Minimizing irreversible losses in quantum systems by local counterdiabatic driving". In: *Proceedings of the National Academy of Sciences* 114.20 (2017), E3909–E3916.

[6] Narendra N. Hegade et al. "Shortcuts to Adiabaticity in Digitized Adiabatic Quantum Computing". In: *Physical Review Applied* 15.2 (2021).

CD driving (illustration of water delivery)



- **Goal**: to deliver the water to the table with a high fidelity, i.e. without spilling or splashing it.
- the glass should be **vertical** in the beginning of the process, i.e. when the waiter is leaving the bar and at the end of the process, when the waiter reaches the table.
- (A) — **adiabatic**: slowly moves along the shortest path (geodesic) and keeps the tray vertically at all times. slow
- (C) — **counter-diabatic**: tilting the tray by applying an opposite force against the pseudo-force. faster

CD inspired generalized ansatz

- Examples of gauge potentials in counter-diabatic driving[7][8]:
  $H_{\mathrm{CD}}(t) = H(\lambda(t)) + \dot{\lambda}\mathcal{A}(\lambda(t)),\ \dot{\lambda}(0) = 0 \quad \dot{\lambda}(T) = 0$

[7] Sels and Polkovnikov, "Minimizing irreversible losses in quantum systems by local counterdiabatic driving".

[8] Hegade et al., "Shortcuts to Adiabaticity in Digitized Adiabatic Quantum Computing".

CD inspired generalized ansatz

- Examples of gauge potentials in counter-diabatic driving[7][8]:
  $H_{\mathrm{CD}}(t) = H(\lambda(t)) + \dot{\lambda}\mathcal{A}(\lambda(t)),\ \dot{\lambda}(0) = 0 \quad \dot{\lambda}(T) = 0$
  - single qubit: $H(\lambda) = \sigma^z + \lambda\sigma^x$, gauge potential: $\mathcal{A}(\lambda) \sim \sigma^y$

[7] Sels and Polkovnikov, "Minimizing irreversible losses in quantum systems by local counterdiabatic driving".

[8] Hegade et al., "Shortcuts to Adiabaticity in Digitized Adiabatic Quantum Computing".

CD inspired generalized ansatz

- Examples of gauge potentials in counter-diabatic driving[7][8]:

$$H_{\mathrm{CD}}(t) = H(\lambda(t)) + \dot{\lambda}\mathcal{A}(\lambda(t)), \;\; \dot{\lambda}(0) = 0 \quad \dot{\lambda}(T) = 0$$

  - single qubit: $H(\lambda) = \sigma^z + \lambda\sigma^x$, gauge potential: $\mathcal{A}(\lambda) \sim \sigma^y$
  - spin chain: $H(\lambda) = \sum_{j=1}^{N} S_{j+1}^z S_j^z + S_j^z + \lambda S_j^x$, asymptotic expansion:
    $\mathcal{A}(\lambda) = \sum_j \alpha(\lambda) S_j^y + \beta(\lambda) \left(S_{j+1}^x S_j^y + S_{j+1}^y S_j^x \right) + \gamma(\lambda) \left(S_{j+1}^z S_j^y + S_{j+1}^y S_j^z \right) + \cdots$

[7] Sels and Polkovnikov, "Minimizing irreversible losses in quantum systems by local counterdiabatic driving".

[8] Hegade et al., "Shortcuts to Adiabaticity in Digitized Adiabatic Quantum Computing".

CD inspired generalized ansatz

- Examples of gauge potentials in counter-diabatic driving[7,8]:
  $H_{\mathrm{CD}}(t) = H(\lambda(t)) + \dot{\lambda}\mathcal{A}(\lambda(t)), \ \dot{\lambda}(0) = 0 \quad \dot{\lambda}(T) = 0$
  - single qubit: $H(\lambda) = \sigma^z + \lambda\sigma^x$, gauge potential: $\mathcal{A}(\lambda) \sim \sigma^y$
  - spin chain: $H(\lambda) = \sum_{j=1}^{N} S_{j+1}^z S_j^z + S_j^z + \lambda S_j^x$, asymptotic expansion:
    $\mathcal{A}(\lambda) = \sum_j \alpha(\lambda) S_j^y + \beta(\lambda)\left(S_{j+1}^x S_j^y + S_{j+1}^y S_j^x\right) + \gamma(\lambda)\left(S_{j+1}^z S_j^y + S_{j+1}^y S_j^z\right) + \cdots$
- 5-mode "bang-bang" protocols

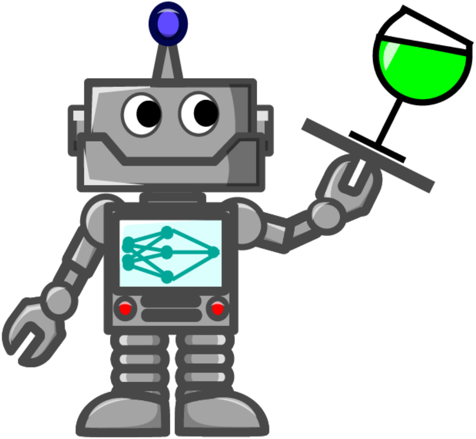
$$
\begin{aligned}
U_0(\alpha_0) &= \exp(-i\alpha_0 H_0), H_0 = \sum_j S_{j+1}^z S_j^z + S_j^z \\
U_1(\alpha_1) &= \exp(-i\alpha_1 H_1), H_1 = \sum_j S_j^x \\
U_2(\alpha_2) &= \exp(-i\alpha_2 H_2), H_2 = \sum_j S_j^y \\
U_3(\alpha_3) &= \exp(-i\alpha_3 H_3), H_3 = \sum_j S_{j+1}^y S_j^x + S_{j+1}^x S_j^y \\
U_4(\alpha_4) &= \exp(-i\alpha_4 H_4), H_4 = \sum_j S_{j+1}^y S_j^z + S_{j+1}^z S_j^y
\end{aligned}
$$

- the number of gates in the sequence: $|A|(|A|-1)^{q-1}$

[7] Sels and Polkovnikov, "Minimizing irreversible losses in quantum systems by local counterdiabatic driving".

[8] Hegade et al., "Shortcuts to Adiabaticity in Digitized Adiabatic Quantum Computing".

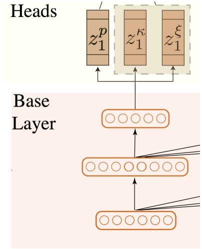
When *CD* driving meets **RL**



RL-QAOA (ancestral sampling)

- **Robust** autoregressive hybrid policy: $\pi_{\boldsymbol{\theta}}(\tau) = \pi^{\mathrm{c}}_{\boldsymbol{\theta}}(\tau^{\mathrm{c}})\, \pi^{\mathrm{d}}_{\boldsymbol{\theta}}\left(\tau^{\mathrm{d}}\right)$
- Hybrid policy gradient:

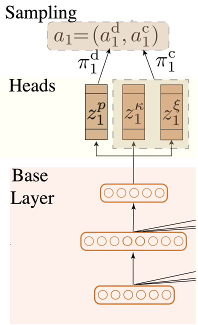
$$\mathcal{J}(\boldsymbol{\theta}) = \mathbb{E}_{\tau}\left[\mathcal{G}^{\mathrm{d}}(\tau^{\mathrm{d}}; \boldsymbol{\theta}, \epsilon^{\mathrm{d}}) + \mathcal{G}^{\mathrm{c}}(\tau^{\mathrm{c}}; \boldsymbol{\theta}, \epsilon^{\mathrm{c}})\right] + \beta_S^{-1}(\mathcal{S}^{\mathrm{d}} + \mathcal{S}^{\mathrm{c}})$$



RL-QAOA (ancestral sampling)

- **Robust** autoregressive hybrid policy: $\pi_{\boldsymbol{\theta}}(\tau) = \pi^{\mathrm{c}}_{\boldsymbol{\theta}}(\tau^{\mathrm{c}})\,\pi^{\mathrm{d}}_{\boldsymbol{\theta}}\left(\tau^{\mathrm{d}}\right)$
- Hybrid policy gradient:

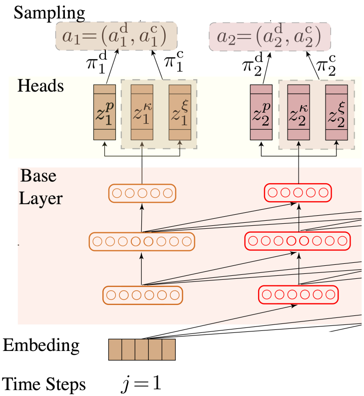
$$\mathcal{J}(\boldsymbol{\theta}) = \mathbb{E}_{\tau}\left[\mathcal{G}^{\mathrm{d}}(\tau^{\mathrm{d}};\boldsymbol{\theta},\epsilon^{\mathrm{d}}) + \mathcal{G}^{\mathrm{c}}(\tau^{\mathrm{c}};\boldsymbol{\theta},\epsilon^{\mathrm{c}})\right] + \beta_S^{-1}(\mathcal{S}^{\mathrm{d}} + \mathcal{S}^{\mathrm{c}})$$



RL-QAOA (ancestral sampling)

- **Robust** autoregressive hybrid policy: $\pi_{\boldsymbol{\theta}}(\tau) = \pi^{\mathrm{c}}_{\boldsymbol{\theta}}\left(\tau^{\mathrm{c}}\right)\pi^{\mathrm{d}}_{\boldsymbol{\theta}}\left(\tau^{\mathrm{d}}\right)$
- Hybrid policy gradient:

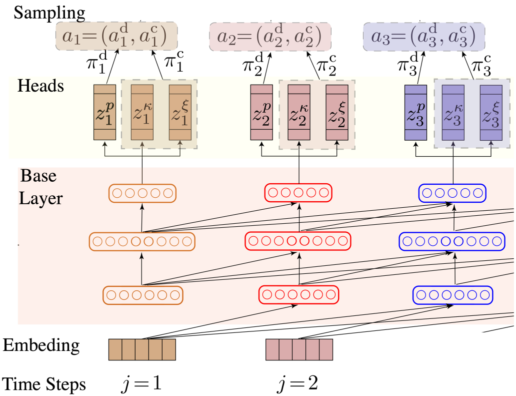
$$\mathcal{J}(\boldsymbol{\theta}) = \mathbb{E}_{\tau}\left[\mathcal{G}^{\mathrm{d}}(\tau^{\mathrm{d}};\boldsymbol{\theta},\epsilon^{\mathrm{d}}) + \mathcal{G}^{\mathrm{c}}(\tau^{\mathrm{c}};\boldsymbol{\theta},\epsilon^{\mathrm{c}})\right] + \beta_S^{-1}(\mathcal{S}^{\mathrm{d}} + \mathcal{S}^{\mathrm{c}})$$



RL-QAOA (ancestral sampling)

- **Robust** autoregressive hybrid policy: $\pi_{\boldsymbol{\theta}}(\tau) = \pi^{\mathrm{c}}_{\boldsymbol{\theta}}\left(\tau^{\mathrm{c}}\right)\pi^{\mathrm{d}}_{\boldsymbol{\theta}}\left(\tau^{\mathrm{d}}\right)$
- Hybrid policy gradient:

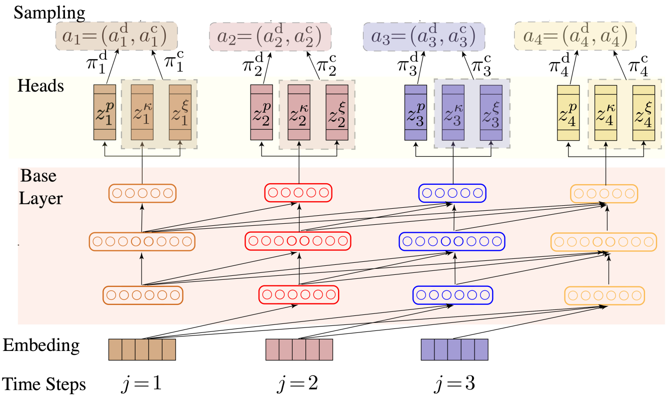
$$\mathcal{J}(\boldsymbol{\theta}) = \mathbb{E}_{\tau}\left[\mathcal{G}^{\mathrm{d}}(\tau^{\mathrm{d}};\boldsymbol{\theta},\epsilon^{\mathrm{d}}) + \mathcal{G}^{\mathrm{c}}(\tau^{\mathrm{c}};\boldsymbol{\theta},\epsilon^{\mathrm{c}})\right] + \beta_S^{-1}(\mathcal{S}^{\mathrm{d}} + \mathcal{S}^{\mathrm{c}})$$



RL-QAOA (ancestral sampling)

- **Robust** autoregressive hybrid policy: $\pi_{\boldsymbol{\theta}}(\tau) = \pi^{\mathrm{c}}_{\boldsymbol{\theta}}\left(\tau^{\mathrm{c}}\right)\pi^{\mathrm{d}}_{\boldsymbol{\theta}}\left(\tau^{\mathrm{d}}\right)$
- Hybrid policy gradient:

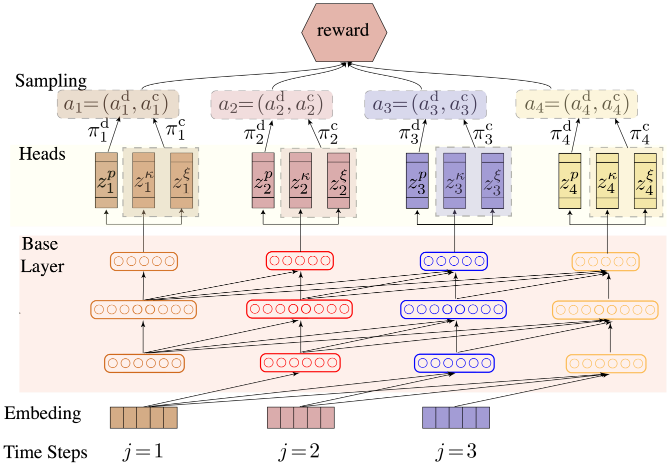
$$\mathcal{J}(\boldsymbol{\theta}) = \mathbb{E}_{\tau}\left[\mathcal{G}^{\mathrm{d}}(\tau^{\mathrm{d}};\boldsymbol{\theta},\epsilon^{\mathrm{d}}) + \mathcal{G}^{\mathrm{c}}(\tau^{\mathrm{c}};\boldsymbol{\theta},\epsilon^{\mathrm{c}})\right] + \beta_S^{-1}(\mathcal{S}^{\mathrm{d}} + \mathcal{S}^{\mathrm{c}})$$



RL-QAOA (reward evaluation)

- **Robust** autoregressive hybrid policy: $\pi_{\boldsymbol{\theta}}(\tau) = \pi_{\boldsymbol{\theta}}^{\mathrm{c}}(\tau^{\mathrm{c}})\, \pi_{\boldsymbol{\theta}}^{\mathrm{d}}(\tau^{\mathrm{d}})$
- Hybrid policy gradient:

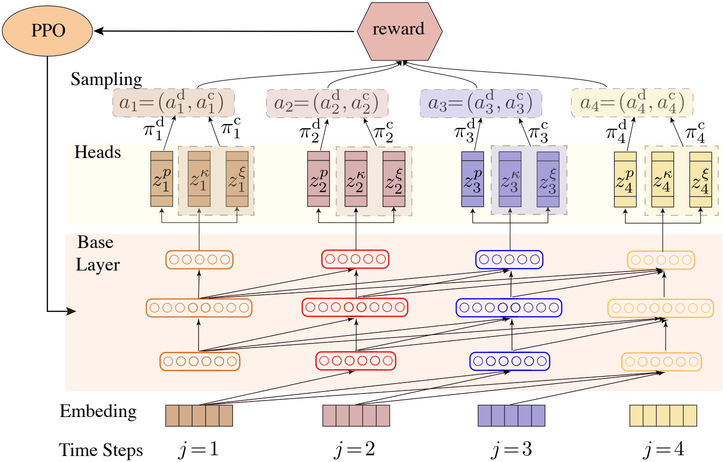
$$\mathcal{J}(\boldsymbol{\theta}) = \mathbb{E}_{\tau}\left[\mathcal{G}^{\mathrm{d}}(\tau^{\mathrm{d}}; \boldsymbol{\theta}, \epsilon^{\mathrm{d}}) + \mathcal{G}^{\mathrm{c}}(\tau^{\mathrm{c}}; \boldsymbol{\theta}, \epsilon^{\mathrm{c}})\right] + \beta_S^{-1}(\mathcal{S}^{\mathrm{d}} + \mathcal{S}^{\mathrm{c}})$$



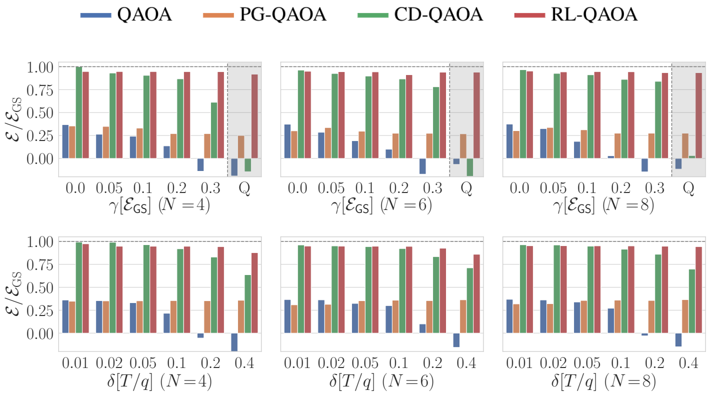
RL-QAOA (proximal policy gradient)

- **Robust** autoregressive hybrid policy: $\pi_{\boldsymbol{\theta}}(\tau) = \pi^{\mathrm{c}}_{\boldsymbol{\theta}}(\tau^{\mathrm{c}})\, \pi^{\mathrm{d}}_{\boldsymbol{\theta}}(\tau^{\mathrm{d}})$
- Hybrid policy gradient:

$$\mathcal{J}(\boldsymbol{\theta}) = \mathbb{E}_{\tau}\left[\mathcal{G}^{\mathrm{d}}(\tau^{\mathrm{d}}; \boldsymbol{\theta}, \epsilon^{\mathrm{d}}) + \mathcal{G}^{\mathrm{c}}(\tau^{\mathrm{c}}; \boldsymbol{\theta}, \epsilon^{\mathrm{c}})\right] + \beta_S^{-1}(\mathcal{S}^{\mathrm{d}} + \mathcal{S}^{\mathrm{c}})$$

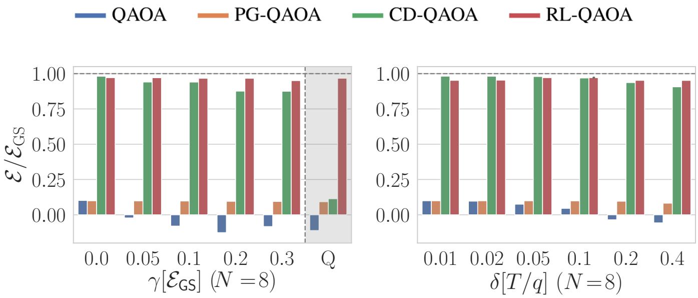


RL-QAOA numerical results



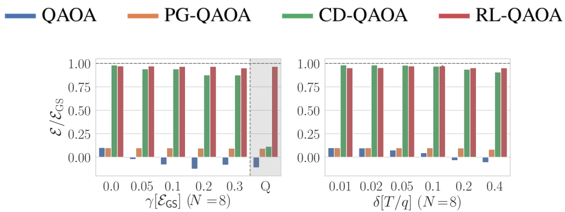
- energy minimization: Ising spin-$1/2$ model
- 1st row: classical noise, quantum measurement noise; 2nd row: gate noise
- $|\psi_i\rangle = |\uparrow \cdots \uparrow\rangle$ and $|\psi_*\rangle = |\psi_{\rm GS}(H)\rangle$
- $\mathcal{A}_{QAOA} = \{H_1, H_2\}$; $\mathcal{A}_{CDQAOA} = \{H_1, H_2; Y, X|Y, Y|Z\}$.
- continuous policies are Sigmoid-Gaussian distributions.

RL-QAOA for Heisenberg spin-1 model



- energy minimization: Heisenberg spin-1 model
- classical noise & gate noise
- $|\psi_i\rangle = |\uparrow\downarrow\cdots\uparrow\downarrow\rangle$ and $|\psi_*\rangle = |\psi_{\mathrm{GS}}(H)\rangle$
- $\mathcal{A}_{QAOA} = \{H_1, H_2\}$;
  $\mathcal{A}_{CDQAOA} = \{H_1, H_2, Z, X|X; Y, XY, YZ, X|Y-XY, Y|Z-YZ\}$.
- continuous policies are Beta distributions.

RL-QAOA for Heisenberg spin-1 model



- $\mathcal{A}_{QAOA} = \{H_1, H_2\}$;
  $\mathcal{A}_{CDQAOA} = \{H_1, H_2, Z, X|X; Y, XY, YZ, X|Y - XY, Y|Z - YZ\}$.

1. hybrid continuous-discrete policies
   $\pi_{\boldsymbol{\theta}}(a_1, a_2, \cdots, a_q) = \prod_{j=1}^{q} \pi_{\boldsymbol{\theta}^{\rm d}}\left(a_j^{\rm d} | s_j\right) \pi_{\boldsymbol{\theta}}^{\rm c}\left(a_j^{\rm c} \mid s_j, a_j^{\rm d}\right)$.
2. agnostic to the physical source of noise
3. currently fixed sequence/protocol length, but versatile enough to accommodate a variable length by adding a “stop” action

Conclusions and Future work

- Discrete-continuous hybrid optimization designed for reinforcement learning displays versatility for such tasks and resilience in the presence of noise.
- Greater performance compared with other gradient-based algorithms and commonly used blackbox optimization, like PGQAOA and CDQAOA.
- Generalized QAOA ansatz is more expressive and goes beyond the framework of QAOA.
- not just selected from a pool of operators, but those are based on principles of variational counter-diabatic driving

Conclusions and Future work

- Discrete-continuous hybrid optimization designed for reinforcement learning displays versatility for such tasks and resilience in the presence of noise.
- Greater performance compared with other gradient-based algorithms and commonly used blackbox optimization, like PGQAOA and CDQAOA.
- Generalized QAOA ansatz is more expressive and goes beyond the framework of QAOA.
- not just selected from a pool of operators, but those are based on principles of variational counter-diabatic driving

Method	QAOA	PG-QAOA	CD-QAOA	RL-QAOA
protocol sequence optimization (discrete)	✗	✗	∇-free	∇-free
gate durations optimization (continuous)	∇-free	∇-free	∇-free	∇-free
RL optimization	✗	continuous	discrete	continuous & discrete
noise-robust	✗	✓	✗	✓
autoregressive	✗	✗	✓	✓

Conclusions and Future work

- Discrete-continuous hybrid optimization designed for reinforcement learning displays versatility for such tasks and resilience in the presence of noise.
- Greater performance compared with other gradient-based algorithms and commonly used blackbox optimization.
- Generalized QAOA ansatz is more expressive and goes beyond the framework of QAOA.
- not just selected from a pool of operators, but those are based on principles of variational counter-diabatic driving

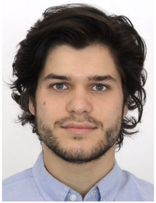
Future directions we're currently working on:

- a fair comparison against gradient-based method, i.e. query complexity.
- extension to more advanced physics systems or optimization algorithms.
- more realistic noise model & implementation on real quantum devices.
- much more powerful neural network ansatz like recurrent networks, self-attention layers, transformer networks

Thank you for your Attention!

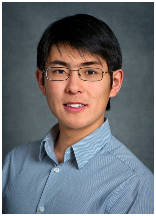
More: CD-QAOA (PRX): **2010.03655** , PG-QAOA (MSML20): **2002.01068**

Hans Gundlach



Paul Köttering

Marin Bukov



Lin Lin

Challenge Institute for Quantum Computation

U.S. DEPARTMENT OF ENERGY Office of Science

NSF

GORDON AND BETTY MOORE FOUNDATION

QAT4Chem

Google Quantum AI

python package for ED & many-body dynamics

QuSpin: http://weinbe58.github.io/QuSpin

ArXiv:2002.01068, 2005.11011, 2010.03655, 2012.06701