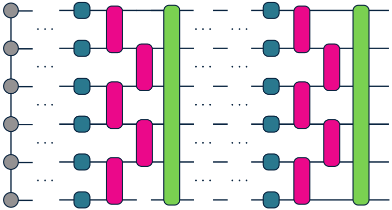
Overparameterization of Realistic Quantum Systems

Matthew Duschenes*, Juan Carrasquilla, Raymond Laflamme
University of Waterloo, Institute for Quantum Computing, & Vector Institute

May 29, 2023

Vector Quantum Workshop

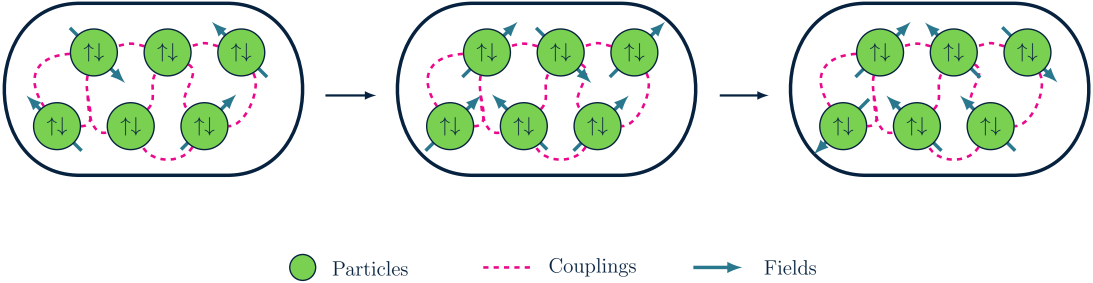


UNIVERSITY OF WATERLOO | IQC Institute for Quantum Computing

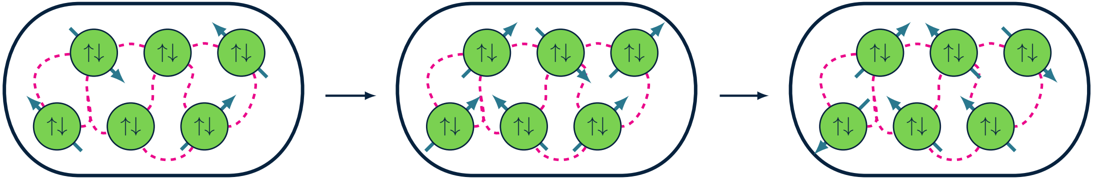
PERIMETER INSTITUTE

VECTOR INSTITUTE

What Are Parameterized Quantum Systems?



What Are Parameterized Quantum Systems?

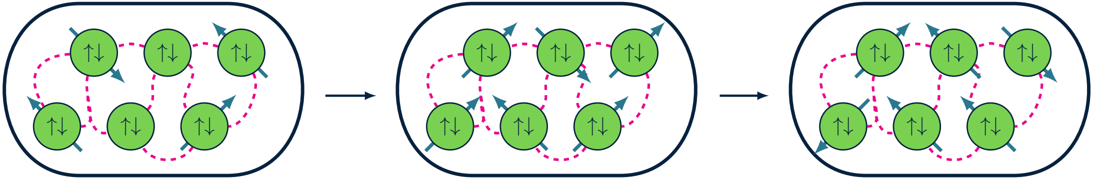


$|\psi\rangle \in \mathbb{C}^{D^N}$ $\qquad$ $U_\theta^{(t)}|\psi\rangle$ $\qquad$ $U_\theta^{(t')}|\psi\rangle$

i.e) $D = 2$ *Qubits*

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle \quad : \quad \langle\psi|\psi\rangle = |\alpha|^2 + |\beta|^2 = 1 \tag{1}$$

What Are Parameterized Quantum Systems?



$|\psi\rangle \in \mathbb{C}^{D^N}$ $\qquad U_\theta^{(t)}|\psi\rangle$ $\qquad U_\theta^{(t')}|\psi\rangle$

i.e) *Pauli* operations on D basis vectors $\sigma \in \{0, 1, \cdots, D-1\}$

$$Z|\sigma\rangle = e^{-i\sigma 2\pi/D}|\sigma\rangle \qquad X|\sigma\rangle = |\sigma \oplus 1\rangle \qquad Y|\sigma\rangle = e^{-i\sigma 2\pi/D}|\sigma \oplus 1\rangle \tag{2}$$

How Do We Evolve Quantum Systems?

Experiments and dynamics are specified by a *Hamiltonian* $H_\theta^{(t)}$ that dictates the instantaneous *energy* of a system, and drives its *unitary* evolution

$$U_\theta^{(t)} = \mathcal{T} e^{-i \int_0^t d\tau \;\; H_\theta^{(\tau)}} \quad : \quad U_\theta^\dagger U_\theta = I \; . \tag{3}$$

How Do We Evolve Quantum Systems?

Experiments and dynamics are specified by a *Hamiltonian* $H_\theta^{(t)}$ that dictates the instantaneous *energy* of a system, and drives its *unitary* evolution

$$U_\theta^{(t)} = \mathcal{T} e^{-i \int_0^t d\tau \;\; H_\theta^{(\tau)}} \quad : \quad U_\theta^\dagger U_\theta = I \;. \tag{3}$$

For classical simulation, evolution is discretized, or *trotterized* into M time steps $\delta = T/M$

$$U_\theta^{(T)} \approx \prod_m^M U_\theta^{(m)} \;+\; O(\delta^2) \tag{4}$$

$$U_\theta^{(m)} = e^{-i\delta H_\theta^{(m)}} \;. \tag{5}$$

How Do We Evolve Quantum Systems?

Experiments and dynamics are specified by a *Hamiltonian* $H_\theta^{(t)}$ that dictates the instantaneous *energy* of a system, and drives its *unitary* evolution

$$U_\theta^{(t)} = \mathcal{T} e^{-i \int_0^t d\tau \ H_\theta^{(\tau)}} \quad : \quad U_\theta^\dagger U_\theta = I \ . \tag{3}$$

For classical simulation, evolution is discretized, or *trotterized* into M time steps $\delta = T/M$

$$U_\theta^{(T)} \approx \prod_m^M U_\theta^{(m)} \ + \ O(\delta^2) \tag{4}$$

$$U_\theta^{(m)} = e^{-i\delta H_\theta^{(m)}} \ . \tag{5}$$

i.e) Nuclear Magnetic Resonance

$$H_\theta^{(t)} = \sum_i \theta_i^{x(t)} X_i \ + \ \sum_i \theta_i^{y(t)} Y_i \ + \ \sum_i h_i Z_i \ + \ \sum_{i<j} J_{ij} Z_i Z_j \tag{6}$$

What Are We Able To Do With Current Quantum Systems?

- Experimental feasibility affects our ability to perform useful *tasks*
  - Unitary *compilation* : $U_\theta \approx U$
  - State *preparation* : $|\psi_\theta\rangle = U_\theta |\phi\rangle \approx |\psi\rangle$
  - Task performance described by *infidelities*: $\mathcal{L}_\theta \sim 1 - \left|\mathrm{tr}\left(U^\dagger U_\theta\right)\right|^2 / D^{2N}$.

What Are We Able To Do With Current Quantum Systems?

- Experimental feasibility affects our ability to perform useful *tasks*
  - Unitary *compilation* : $U_\theta \approx U$
  - State *preparation* : $|\psi_\theta\rangle = U_\theta |\phi\rangle \approx |\psi\rangle$
  - Task performance described by *infidelities*: $\mathcal{L}_\theta \sim 1 - \left|\mathrm{tr}\left(U^\dagger U_\theta\right)\right|^2 / D^{2N}$.

- These systems are also severely *constrained*

  i.e) *Bounds* on the fields, and imposing *Uniformity* or *Boundary-conditions*

What Are We Able To Do With Current Quantum Systems?

- Experimental feasibility affects our ability to perform useful *tasks*
  - Unitary *compilation* : $U_\theta \approx U$
  - State *preparation* : $|\psi_\theta\rangle = U_\theta |\phi\rangle \approx |\psi\rangle$
  - Task performance described by *infidelities*: $\mathcal{L}_\theta \sim 1 - \left|\mathrm{tr}\left(U^\dagger U_\theta\right)\right|^2 / D^{2N}$.

- These systems are also severely *constrained*

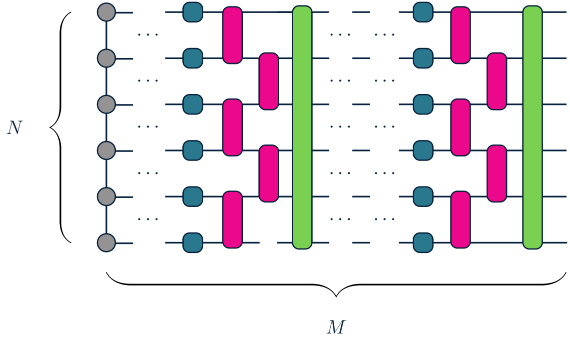
  i.e) *Bounds* on the fields, and imposing *Uniformity* or *Boundary-conditions*

- Systems detrimentally interact with their environment, resulting in *noise* γ

  i.e) *Dephasing* $\mathcal{K}_\gamma = \{\sqrt{1-\gamma}\ I,\ \sqrt{\gamma}\ Z\}$

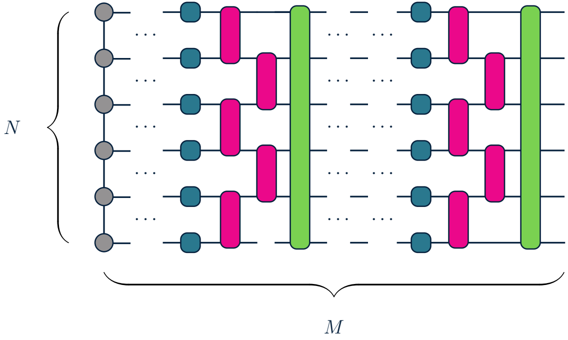
$$|\phi\rangle \ \rightarrow \ \phi = |\phi\rangle\langle\phi| \ \overset{U_\theta, \mathcal{K}_\gamma}{\rightarrow} \ \rho_{\theta\gamma} = \sum_{K_\gamma \, \in \, \mathcal{K}_\gamma} K_\gamma \ U_\theta \ \phi \ U_\theta^\dagger \ K_\gamma^\dagger \tag{7}$$

Learning Optimal Quantum Systems



How does the amount of *noise* γ and the *evolution depth* M
of a *constrained* system
affect its classical simulation and optimization, and resulting infidelities
$\mathcal{L}_{\theta^*\gamma}$?

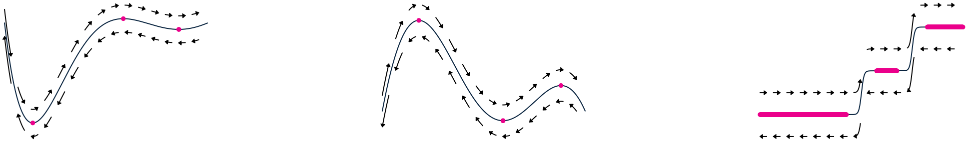
Learning Optimal Quantum Systems



How can we leverage approaches from *quantum optimal control* and *learning theory* to describe these relationships?

Learning Phenomena

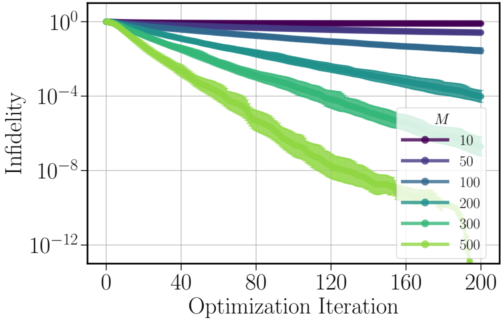
- How do optimization algorithms *learn*, and traverse the *objective landscape*?



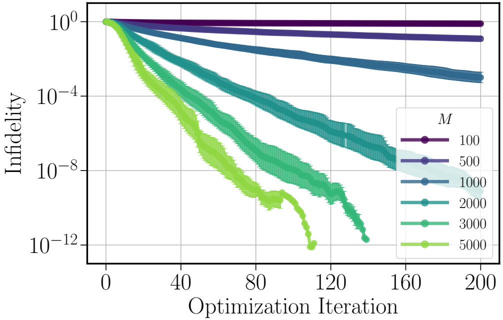
- Learning can converge exponentially quickly in the *overparameterized* regime
- Dimensionality of *dynamical Lie algebra* spanned by Hamiltonian, determines *expressivity* (Larocca *et al.* arXiv:2109.11676 (2021))
- Optimal control pulses must evolve according to a *quantum speed limit* (Deffner *et al.* J. Phys. A, **50** (2017))

Unconstrained vs. Constrained Optimization

- Haar random unitary compilation for $N = 4$ qubits, with *bounded* fields *shared* across all qubits, and Dirichlet *boundary conditions*



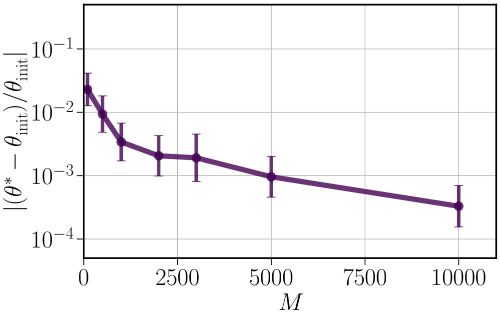
(a) Unconstrained Infidelity



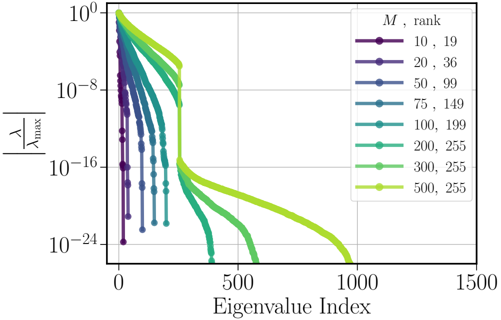
(b) Constrained Infidelity

Overparameterization Phenomena

- *Overparameterized* regime is reached with constraints for sufficient depth $M > O(G)$ (For universal $\mathcal{G}_{\mathrm{NMR}}$, $G = 2^{2N} - 1 = 255$)



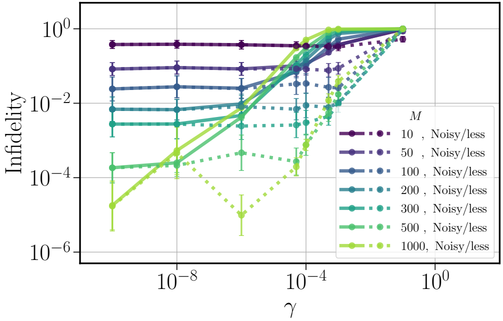
(c) Negligible Relative Change of Parameters from Initialization



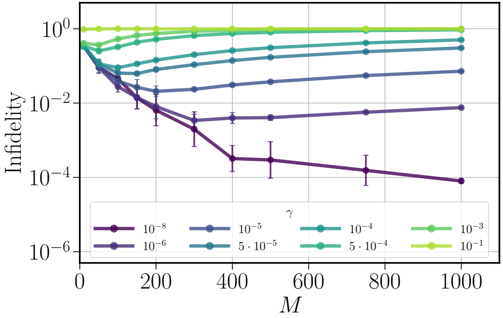
(d) Fisher Information Rank Saturation at G

Noisy Optimization

- Haar random state preparation for $N = 4$ qubits, with independent dephasing



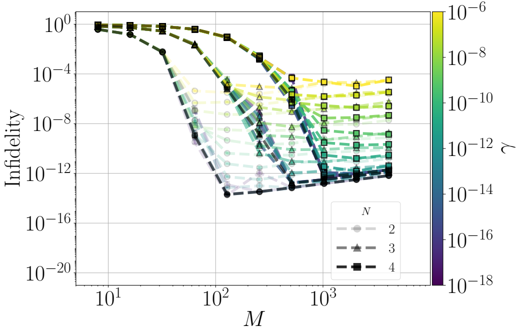
(e) Trained Noisy Infidelity, and Tested Infidelity of Noisy Parameters in Noiseless Ansatz



(f) Critical Depth for Noisy Infidelity

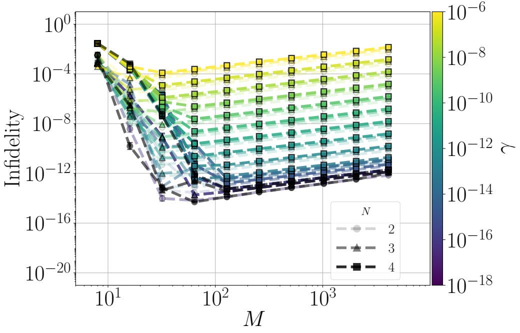
Universal Effects of Noise

- Effects of infidelities on noise for Haar random targets in $n = D^N$ dimensions



(g) Classical floating point noise for unitary compilation

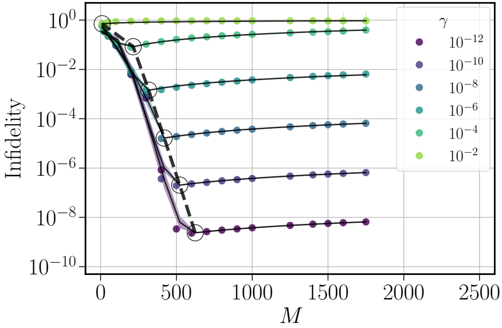
$$|\mathcal{L}_{\theta_\gamma} - \mathcal{L}_\theta| \leq n^{O(NM)}\left|(1+\gamma)^{O(NM)} - 1\right|$$



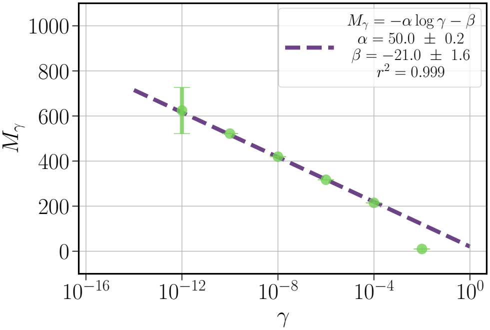
(h) Quantum dephasing noise for state preparation

$$|\mathcal{L}_{\theta_\gamma} - \mathcal{L}_\theta| \leq 2\left|(1-\gamma)^{NM} - 1\right|$$

Noise Induced Critical Depth



(i) Piecewise Fit of Noisy Infidelity



(j) Linear-Log Fit of Critical Depth

Noise Induced Critical Depth

Noise induces a critical depth (Fontana *et al.* PRA **104** (2021))

$$M_{\gamma} \sim \log 1/\gamma \ , \tag{8}$$

meaning the minimum infidelity is *linear-quadratic* ($1 \leq \alpha \leq 2$) in noise

$$\mathcal{L}_{\theta^*\gamma|M_{\gamma}} \sim \gamma^{\alpha} \ , \tag{9}$$

and parameterized noise channels can therefore *mitigate* approximately

$$\bar{M}_{\gamma} \sim \gamma \log 1/\gamma \quad \text{errors} \ . \tag{10}$$

Noise Induced Critical Depth

Noise induces a critical depth (Fontana *et al.* PRA **104** (2021))

$$M_\gamma \sim \log 1/\gamma \ , \tag{8}$$

meaning the minimum infidelity is *linear-quadratic* $(1 \leq \alpha \leq 2)$ in noise

$$\mathcal{L}_{\theta^*\gamma|M_\gamma} \sim \gamma^\alpha \ , \tag{9}$$

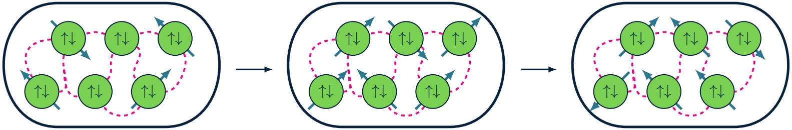
and parameterized noise channels can therefore *mitigate* approximately

$$\bar{M}_\gamma \sim \gamma \log 1/\gamma \quad \text{errors} \ . \tag{10}$$

Is it possible to derive the M, γ scaling of the optimal $\mathcal{L}_{\theta^*\gamma}$ *analytically*?

What Have We Learned About Noisy Overparameterization?

- Overparameterization is *robust* to constraints; requires $\sim O(N)$ greater depth
- Accumulation of noise induces a *critical* depth M_γ that prevents convergence
- Non-trivial compromises between numerical and experimental feasibility
- Channel fidelities, entanglement measures, and mitigation interpretations, will further quantify the abilities of noisy variational ansatz



Appendix

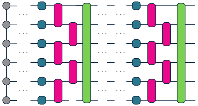
How May We Control Quantum Systems?

- Represented as *channels* $\Lambda_{\theta_\gamma} = \mathcal{N}_\gamma \circ \mathcal{U}_\theta$ with unitary evolution $\mathcal{U}_\theta$, and noise $\mathcal{N}_\gamma$
- Evolution generated by Hamiltonians with localized generators $\{G_\mu\}$

$$H_\theta^{(t)} = \sum_\mu \theta_\mu^{(t)} G_\mu \;\rightarrow\; U_\theta \approx \prod_t^M U_\theta^{(t)} \;:\; U_\theta^{(t)} = e^{-i\delta H_\theta^{(t)}} \approx \prod_\mu e^{-i\delta\theta_\mu^{(t)} G_\mu} \tag{11}$$

i.e) *NMR* with variable transverse fields and constant longitudinal fields (Peterson *et al.* , PRA **13** (2020)) (Coloured in circuit ↘)

$$H_\theta^{(t)} = \sum_i \theta_i^{x\,(t)} X_i + \sum_i \theta_i^{y\,(t)} Y_i + \sum_i h_i Z_i + \sum_{i<j} J_{ij} Z_i Z_j \tag{12}$$



- Noise generated by constant *Kraus* operators $\{\mathcal{K}_{\gamma_\alpha}\}$
  i.e) Dephasing $\{\sqrt{1-\gamma}\; I,\;\; \sqrt{\gamma}\; Z\}$

$$\rho \;\rightarrow\; \rho_{\Lambda_{\theta_\gamma}} = \prod_t^M \mathcal{N}_\gamma \circ \mathcal{U}_\theta^{(t)}(\rho) = \prod_t^M \left[\sum_\alpha \mathcal{K}_{\gamma_\alpha}\, U_\theta^{(t)}\; \rho\; U_\theta^{(t)\,\dagger}\; \mathcal{K}_{\gamma_\alpha}{}^\dagger \right] \tag{13}$$

How Do We Optimize Quantum Systems?

- Systems must be efficiently simulated *classically* i.e) Just-in-time compilation
- Parameters are optimized with *gradient methods* i.e) Automatic differentiation
- Desired tasks are represented as *objectives* to be minimized i.e) (In)Fidelities

$$\mathcal{L}_{\theta\gamma} \sim \mathrm{tr}\left(\rho_{\Lambda_{\theta\gamma}} \rho_U\right) \tag{14}$$

- *Analogous* forms of gradients of objectives in noiseless and noisy system i.e) Exact *parameter-shift* rules, for some generator-dependent angle ζ

$$\partial\mathcal{L}_{\theta\gamma} \sim \mathcal{L}_{\theta+\zeta\ \gamma} - \mathcal{L}_{\theta-\zeta\ \gamma} \tag{15}$$