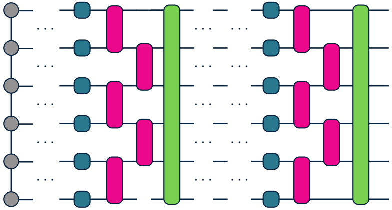
Overparameterization of Realistic Quantum Systems

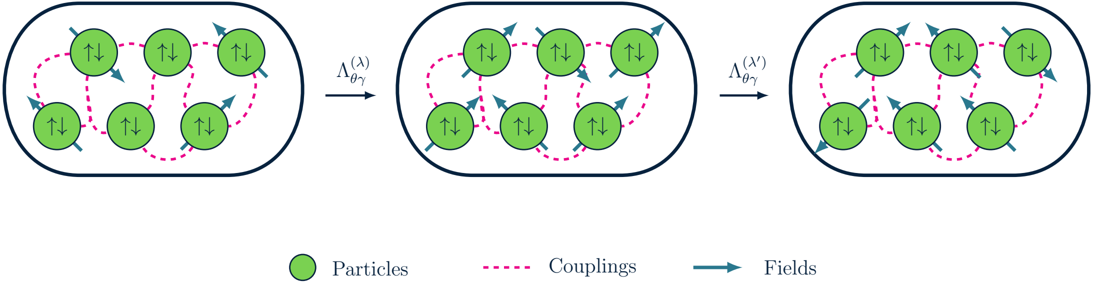
Matthew Duschenes*, Juan Carrasquilla, Raymond Laflamme
University of Waterloo, Institute for Quantum Computing, & Vector Institute

January 26, 2023

Canadian Graduate Quantum Conference



Parameterized Quantum Systems



Evolution via $\Lambda_{\theta\gamma}^{(\lambda)}$ at time λ

What Are We Able To Do With Current Quantum Systems?

- Experimental feasibility affects our ability to perform useful *tasks*

  i.e) Unitary *compilation* $\Lambda_{\theta_\gamma} \approx U$, State *preparation* $\rho_{\Lambda_{\theta_\gamma}} \approx \rho_U$

What Are We Able To Do With Current Quantum Systems?

- Experimental feasibility affects our ability to perform useful *tasks*

  i.e) Unitary *compilation* $\Lambda_{\theta_\gamma} \approx U$, State *preparation* $\rho_{\Lambda_{\theta_\gamma}} \approx \rho_U$

- These systems require incredibly *precise* control of parameters $\theta^{(\lambda)}$

$$H_\theta^{(\lambda)} = \sum_i \theta_i^{x(\lambda)} X_i + \sum_i \theta_i^{y(\lambda)} Y_i + \sum_i h_i Z_i + \sum_{i<j} J_{ij} Z_i Z_j \tag{1}$$

What Are We Able To Do With Current Quantum Systems?

- Experimental feasibility affects our ability to perform useful *tasks*

  i.e) Unitary *compilation* $\Lambda_{\theta_\gamma} \approx U$, State *preparation* $\rho_{\Lambda_{\theta_\gamma}} \approx \rho_U$

- These systems require incredibly *precise* control of parameters $\theta^{(\lambda)}$

$$H_\theta^{(\lambda)} = \sum_i \theta_i^{x(\lambda)} X_i + \sum_i \theta_i^{y(\lambda)} Y_i + \sum_i h_i Z_i + \sum_{i<j} J_{ij} Z_i Z_j \tag{1}$$

- 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*

  i.e) Unitary *compilation* $\Lambda_{\theta_\gamma} \approx U$, State *preparation* $\rho_{\Lambda_{\theta_\gamma}} \approx \rho_U$

- These systems require incredibly *precise* control of parameters $\theta^{(\lambda)}$

$$H_\theta^{(\lambda)} = \sum_i \theta_i^{x(\lambda)} X_i + \sum_i \theta_i^{y(\lambda)} Y_i + \sum_i h_i Z_i + \sum_{i<j} J_{ij} Z_i Z_j \tag{1}$$

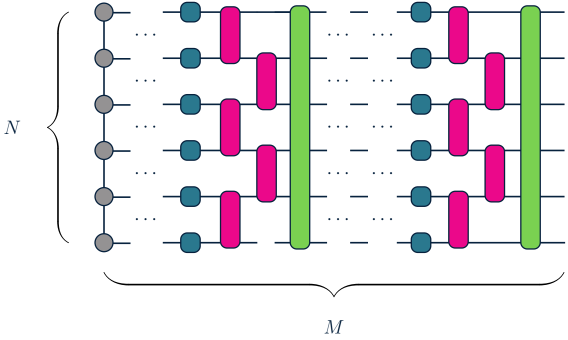
- These systems are also severely *constrained*

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

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

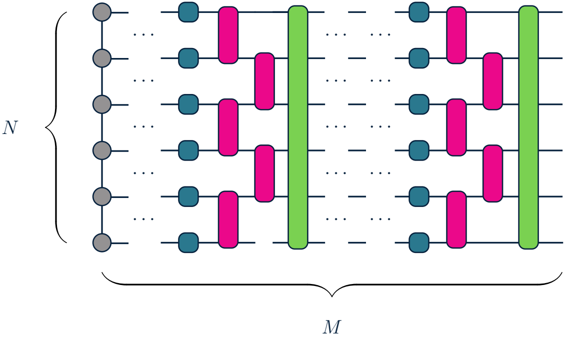
  i.e) *Dephasing* $\mathcal{K}_\gamma = \{\sqrt{1-\gamma}\ I,\ \sqrt{\gamma}\ Z\}$

Learning Optimal Quantum Systems



How does the amount of *noise* γ and the *evolution depth* M
of a *constrained* system $\Lambda_{\theta\gamma}$
affect its *classical simulation and optimization*, and resulting parameters θ?

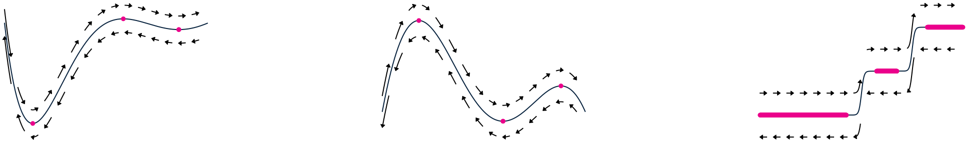
Learning Optimal Quantum Systems



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

Learning Phenomena

- Essential to characterize how optimization and learning algorithms *learn*, and converge towards an optimal solution, as they traverse the *objective landscape*

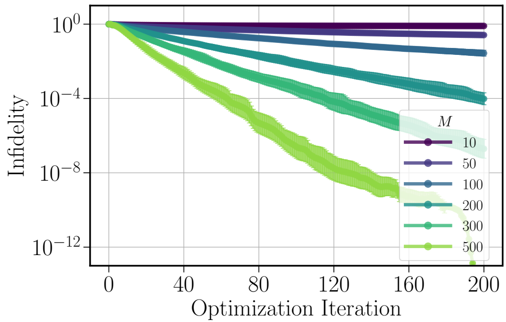


- Several interesting phenomena occur in conventional learning
  - *Overparameterization*: Learning can converge *exponentially* quickly
  - *Lazy training*: Parameters may change *negligibly* from their initial values

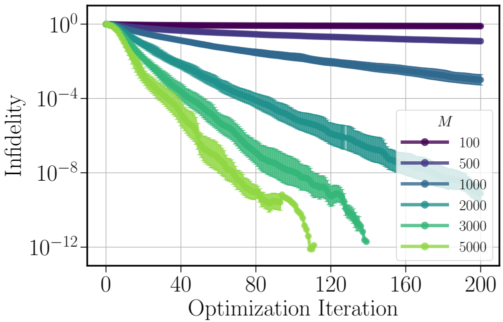
- Ansatz generators $\{G_\mu\}$ form a *dynamical Lie algebra* $\mathcal{G}$, with dimensionality $G = |\mathcal{G}|$, that determines the *expressivity* of an ansatz, depending if the circuit *depth* $M \leqslant O(G)$ (Larocca *et al.* arXiv:2109.11676 (2021))

Unconstrained vs. Constrained Optimization

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



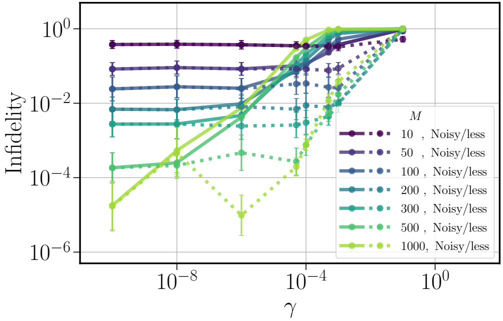
(a) Unconstrained Infidelity



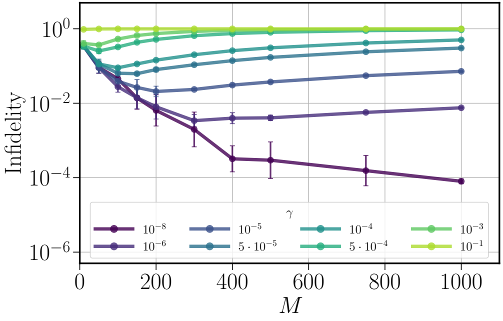
(b) Constrained Infidelity

Noisy Optimization

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



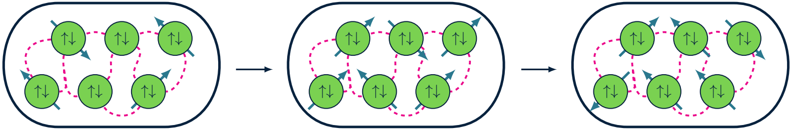
(c) Trained Noisy Infidelity, and Tested Infidelity of Noisy Parameters in Noiseless Ansatz



(d) Critical Depth for Noisy Infidelity

What Have We Learned About Learning?

- 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 and entanglement measures will further quantify effects of noise on the abilities of variational ansatz to learn tasks



Appendix

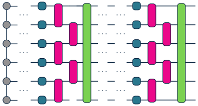
How May We Control Quantum Systems?

- 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 $\{G_\mu\}$

$$H_\theta^{(\lambda)} = \sum_\mu \theta_\mu^{(\lambda)} G_\mu \;\rightarrow\; U_\theta \approx \prod_\lambda^M U_\theta^{(\lambda)} \;:\; U_\theta^{(\lambda)} = e^{-i\delta H_\theta^{(\lambda)}} \approx \prod_\mu e^{-i\delta\theta_\mu^{(\lambda)} G_\mu} \tag{2}$$

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

$$H_\theta^{(\lambda)} = \sum_i \theta_i^{x(\lambda)} X_i + \sum_i \theta_i^{y(\lambda)} Y_i + \sum_i h_i Z_i + \sum_{i<j} J_{ij} Z_i Z_j \tag{3}$$



- Noise generated by constant *Kraus* operators $\{\mathcal{K}_{\gamma_\alpha}\}$

i.e) Dephasing $\{\sqrt{1-\gamma}\; I,\; \sqrt{\gamma}\; Z\}$

$$\rho \;\rightarrow\; \rho_{\Lambda_{\theta_\gamma}} = \prod_\lambda^M \mathcal{N}_\gamma \circ \mathcal{U}_\theta^{(\lambda)}(\rho) = \prod_\lambda^M \left[\sum_\alpha \mathcal{K}_{\gamma_\alpha}\, U_\theta^{(\lambda)}\, \rho\, {U_\theta^{(\lambda)}}^\dagger\, {\mathcal{K}_{\gamma_\alpha}}^\dagger\right] \tag{4}$$

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

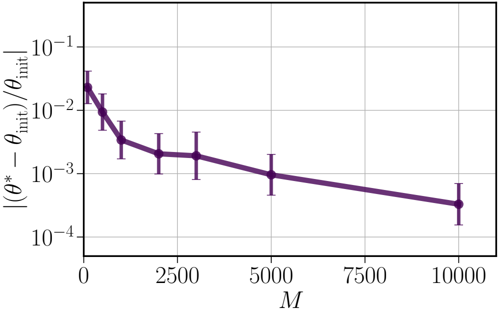
$$\mathcal{L}_{\theta\gamma} \sim \mathrm{tr}\left(\rho_{\Lambda_{\theta\gamma}} \rho_U\right) \tag{5}$$

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

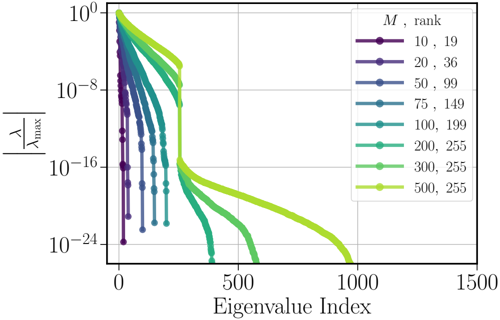
$$\partial\mathcal{L}_{\theta\gamma} \sim \mathcal{L}_{\theta+\zeta\ \gamma} - \mathcal{L}_{\theta-\zeta\ \gamma} \tag{6}$$

Overparameterization Phenomena

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



(e) Negligible Relative Change of Parameters from Initialization



(f) Fisher Information Rank Saturation at G