Title: Overparameterization of Realistic Quantum Systems

Speakers: Matt Duschenes
Date: November 28, 2022 - 2:00 PM

URL: https://pirsa.org/22110060

Abstract: In order for quantum computing devices to accomplish preparation of quantum states, or simulation of other quantum systems, exceptional control of experimental parameters is required. The optimal parameters, such as time dependent magnetic fields for nuclear magnetic resonance, are found via classical simulation and optimization. Such idealized parameterized quantum systems have been shown to exhibit different phases of learning during optimization, such as overparameterization and lazy training, where global optima may potentially be reached exponentially quickly, while parameters negligibly change when the system is evolved for sufficient time (Larocca et al., arXiv:2109.11676, 2021). Here, we study the effects of imposing constraints related to experimental feasibility on the controls, such as bounding or sharing parameters across operators, and relevant noise channels are added after each time step. We observe overparameterization being robust to parameter constraints, however fidelities converge to zero past a critical simulation duration, due to catastrophic accumulation of noise. Compromises arise between numerical and experimental feasibility, suggesting limitations of variational ansatz to account for noise.

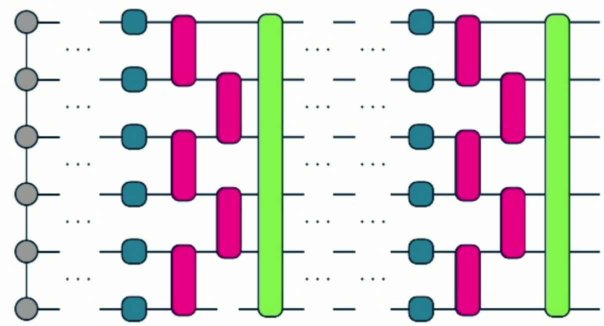
Zoom link: https://pitp.zoom.us/j/98649931693?pwd=Z2s1MlZvSmFVNEFqdjk2dlZNRm9PQT09

Overparameterization of Realistic Quantum Systems

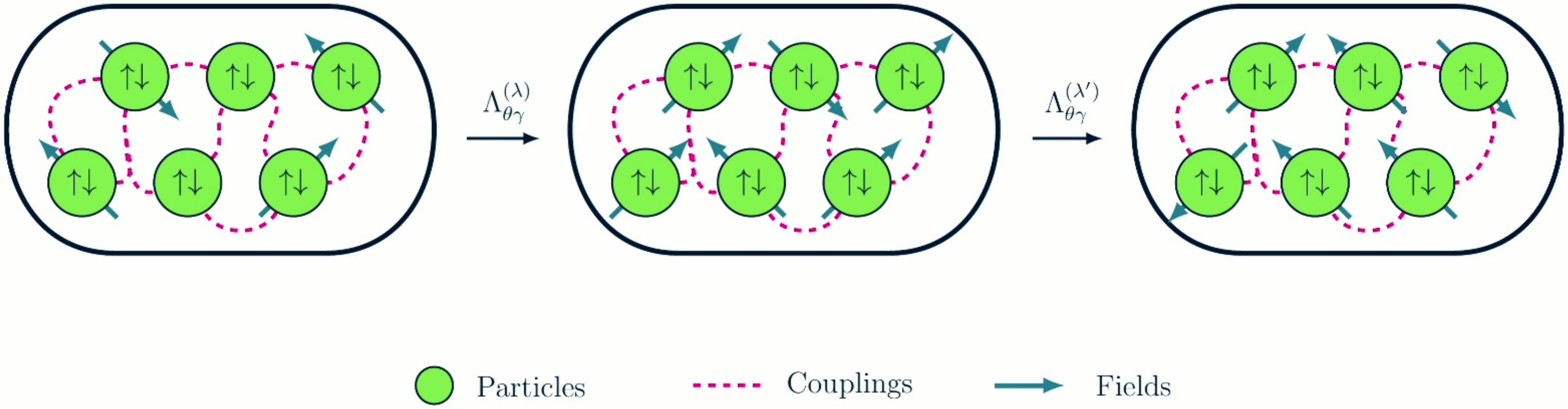
Matthew Duschenes*, Juan Carrasquilla, Raymond Laflamme
University of Waterloo, Institute for Quantum Computing, & Vector Institute

November 28, 2022

PI Student Seminars



Parameterized Quantum Systems



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

What Are We Able To Do With Current Quantum Systems?

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

  i.e) Magnetic field pulses, Coherent laser sources, Optical tweezers

What Are We Able To Do With Current Quantum Systems?

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

  i.e) Magnetic field pulses, Coherent laser sources, Optical tweezers

- These systems are also severely affected by *experimental feasibility*

  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, Depolarizing, Amplitude damping noise

- These constraints hinder our abilities to perform useful *tasks*

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

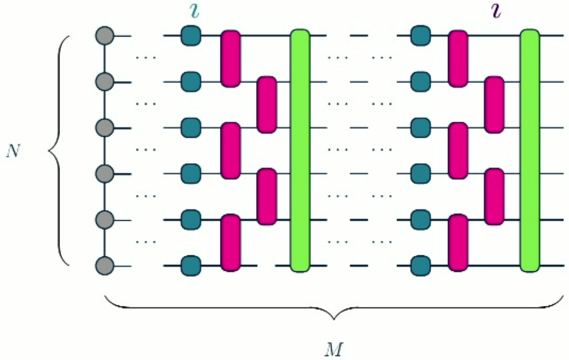
How May We Control Quantum Systems?

- 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 $\{G_\mu\}$

$$H_\theta^{(\lambda)} = \sum_\mu \theta_\mu^{(\lambda)} G_\mu \;\rightarrow\; U_\theta \approx \prod_\lambda^M U_\theta^{(\lambda)} \;:\; U_\theta^{(\lambda)} = e^{-i\delta H_\theta^{(\lambda)}} \approx \prod_\mu e^{-i\delta\theta_\mu^{(\lambda)} G_\mu} \tag{1}$$

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

$$H_\theta^{(\lambda)} = \sum_i \theta_i^{x(\lambda)} X_i \;+\; \sum_i \theta_i^{y(\lambda)} Y_i \;+\; \sum_i h_i Z_i \;+\; \sum_{i<j} J_{ij} Z_i Z_j \tag{2}$$



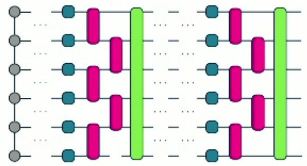
How May We Control Quantum Systems?

- 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 $\{G_\mu\}$

$$H_\theta^{(\lambda)} = \sum_\mu \theta_\mu^{(\lambda)} G_\mu \ \rightarrow \ U_\theta \approx \prod_\lambda^M U_\theta^{(\lambda)} \ : \ U_\theta^{(\lambda)} = e^{-i\delta H_\theta^{(\lambda)}} \approx \prod_\mu e^{-i\delta\theta_\mu^{(\lambda)} G_\mu} \qquad (1)$$

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

$$H_\theta^{(\lambda)} = \sum_i \theta_i^{x(\lambda)} X_i \ + \sum_i \theta_i^{y(\lambda)} Y_i \ + \sum_i h_i Z_i \ + \sum_{i<j} J_{ij} Z_i Z_j \qquad (2)$$



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

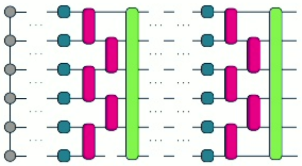
How May We Control Quantum Systems?

- 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 $\{G_\mu\}$

$$H_\theta^{(\lambda)} = \sum_\mu \theta_\mu^{(\lambda)} G_\mu \;\rightarrow\; U_\theta \approx \prod_\lambda^M U_\theta^{(\lambda)} \;:\; U_\theta^{(\lambda)} = e^{-i\delta H_\theta^{(\lambda)}} \approx \prod_\mu e^{-i\delta\theta_\mu^{(\lambda)} G_\mu} \tag{1}$$

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

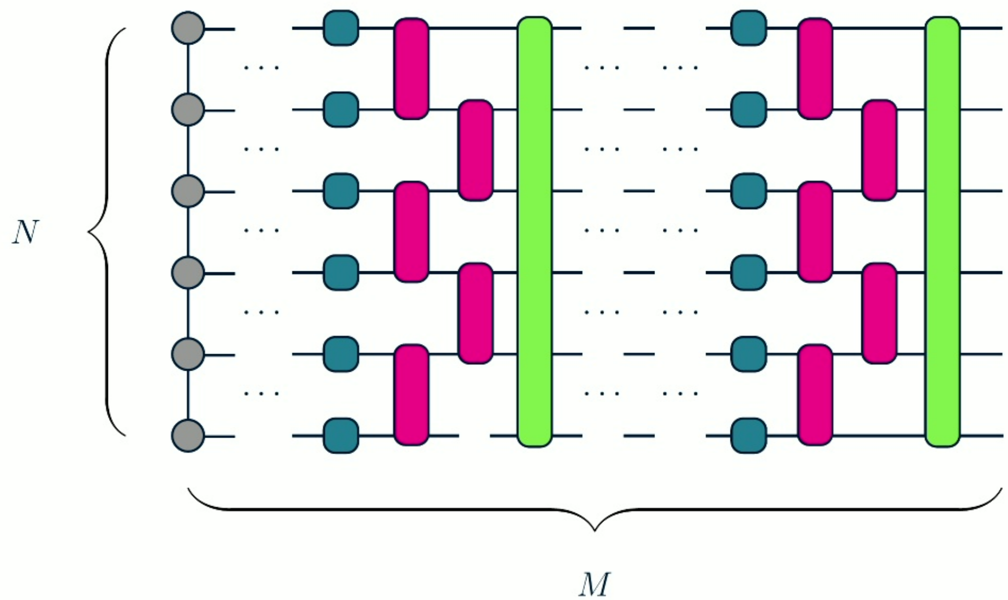
$$H_\theta^{(\lambda)} = \sum_i \theta_i^{x(\lambda)} X_i + \sum_i \theta_i^{y(\lambda)} Y_i + \sum_i h_i Z_i + \sum_{i<j} J_{ij} Z_i Z_j \tag{2}$$



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

$$\rho \;\rightarrow\; \rho_{\Lambda_{\theta_\gamma}} = \prod_\lambda^M \mathcal{N}_\gamma \circ \mathcal{U}_\theta^{(\lambda)}(\rho) = \prod_\lambda^M \left[\sum_\alpha \mathcal{K}_{\gamma_\alpha}\, U_\theta^{(\lambda)}\, \rho\, U_\theta^{(\lambda)^\dagger}\, \mathcal{K}_{\gamma_\alpha}{}^\dagger\right] \tag{3}$$

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 optimization and resulting parameters θ?

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

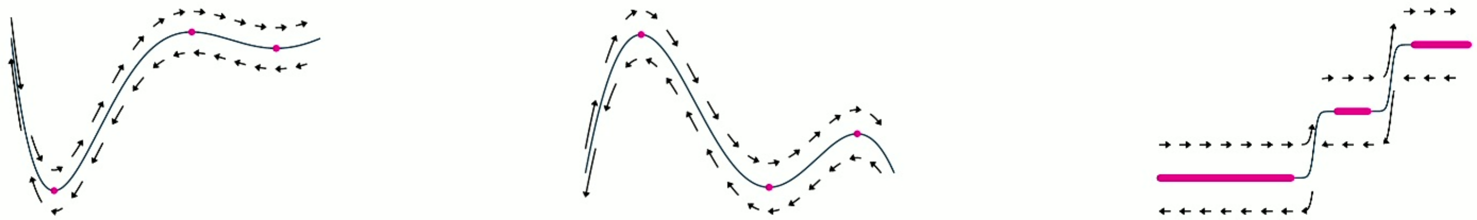
$$\mathcal{L}_{\theta_\gamma} \sim \operatorname{tr}\left(\rho_{\Lambda_{\theta_\gamma}} \rho_U\right) \tag{4}$$

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

$$\partial \mathcal{L}_{\theta_\gamma} \sim \mathcal{L}_{\theta+\zeta\ \gamma} - \mathcal{L}_{\theta-\zeta\ \gamma} \tag{5}$$

Learning Phenomena

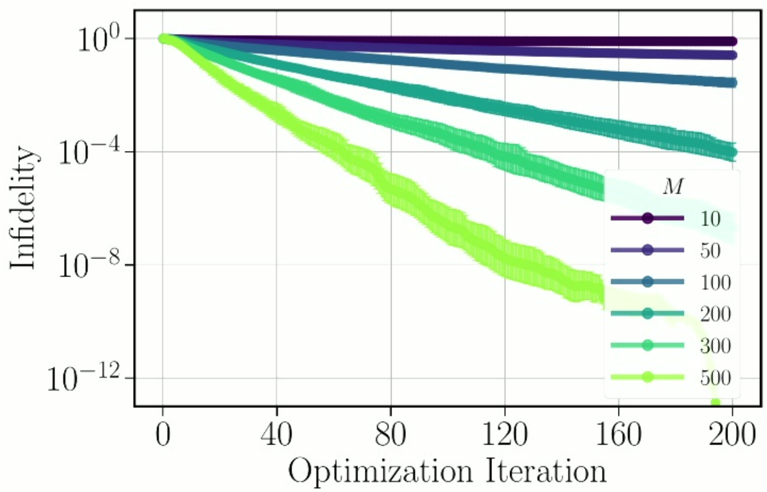
- Essential to characterize how optimization and learning algorithms *learn*, and converge towards an optimal solution, as they traverse the *objective landscape*



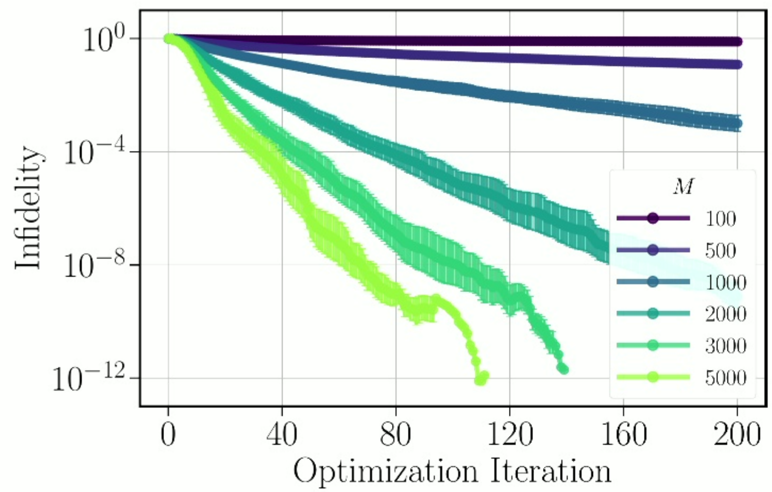
- Several interesting phenomena observed to occur in conventional learning
  - *Overparameterization*: Learning can converge *exponentially* quickly
  - *Lazy training*: Parameters may change *negligibly* from their initial values
- Ansatz generators $\{G_\mu\}$ form a *dynamical Lie algebra* $\mathcal{G}$, with dimensionality $G = |\mathcal{G}|$, that determines the *expressivity* of an ansatz, depending if the circuit *depth* $M \lessgtr O(G)$ (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*



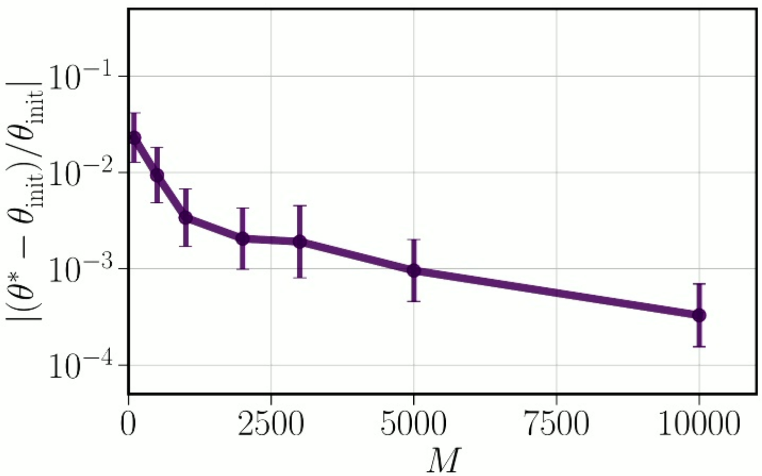
(a) Unconstrained Infidelity



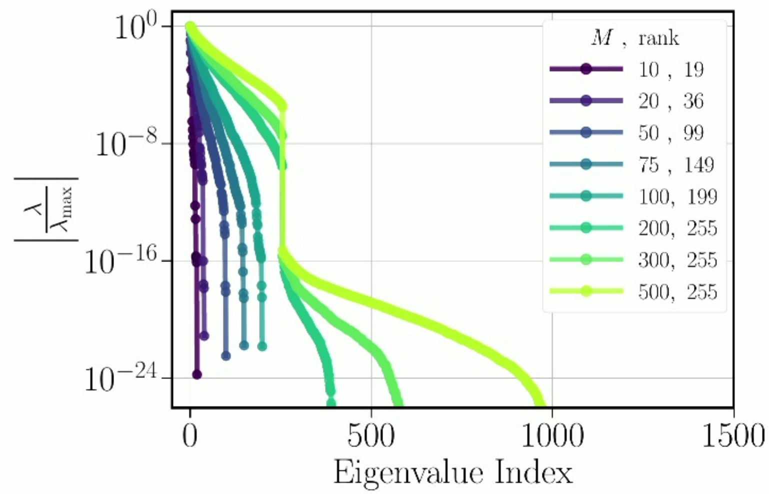
(b) Constrained Infidelity

Overparameterization Phenomena

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



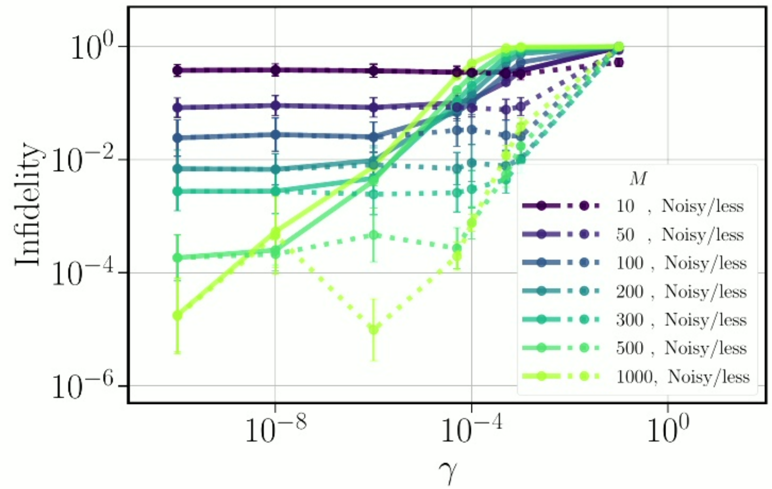
(c) Negligible Relative Change of Parameters from Initialization



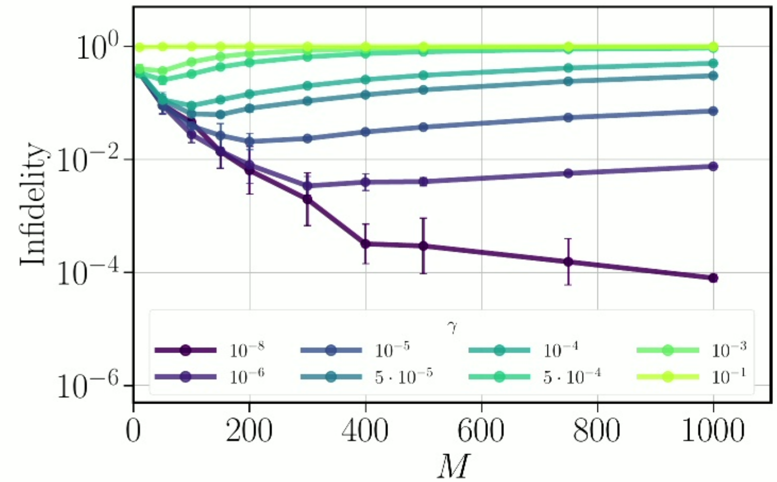
(d) Fisher Information Rank Saturation at G

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

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