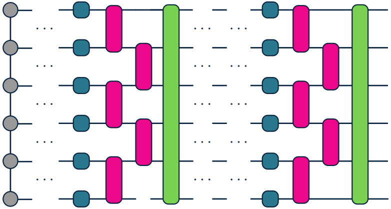
Overparameterization of Realistic Quantum Systems

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

November 29, 2023

PI/IVADO/Courtois/MILA Quantum and AI Workshop

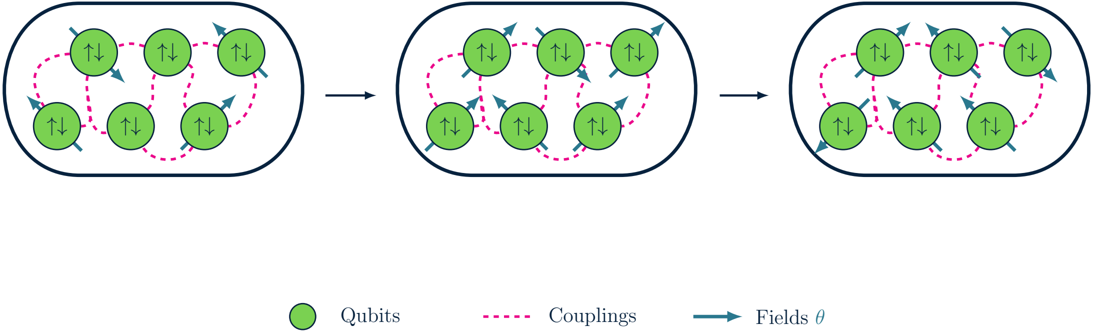


UNIVERSITY OF WATERLOO | IQC Institute for Quantum Computing

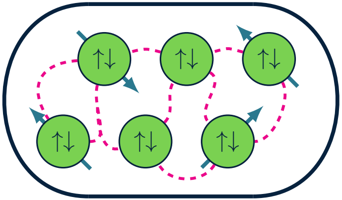
PERIMETER INSTITUTE

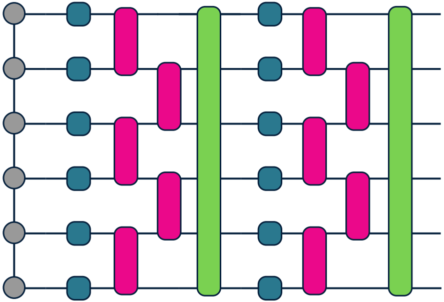
VECTOR INSTITUTE

What Are Parameterized Quantum Systems?



What Are Parameterized Quantum Systems?





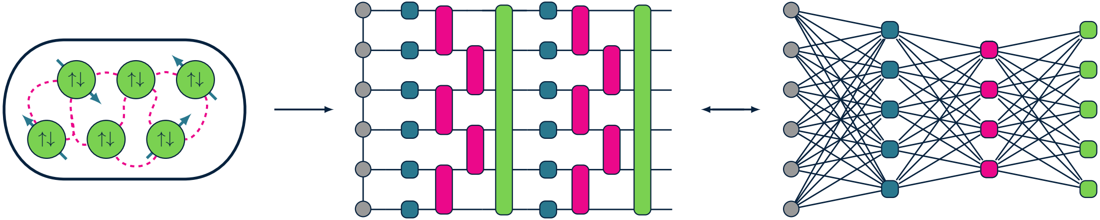
Quantum System

$H_\theta = \sum_\mu \theta_\mu G_\mu$

Quantum Circuit

$U_\theta = \prod_\mu U_\theta^\mu$

What Are Parameterized Quantum Systems?

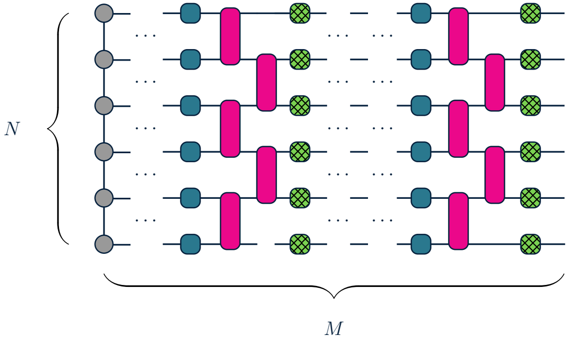


Quantum System | Quantum Circuit | Classical Algorithm

$H_\theta = \sum_\mu \theta_\mu G_\mu$ | $U_\theta = \prod_\mu U_\theta^\mu$ | $f_\theta = \circ_\mu f_\theta^\mu$

Tasks of Interest: Unitary Compilation, State Preparation

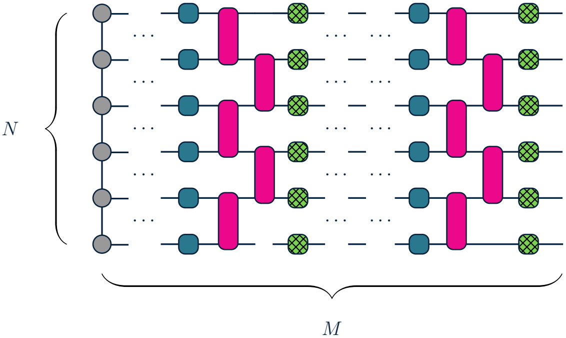
Learning Phenomena of Quantum Systems



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

$$\mathcal{L}_{\theta^*\gamma} \;:\; U_{\theta\gamma} \approx U \;,\; \rho_{\theta\gamma} \approx \rho \;?$$

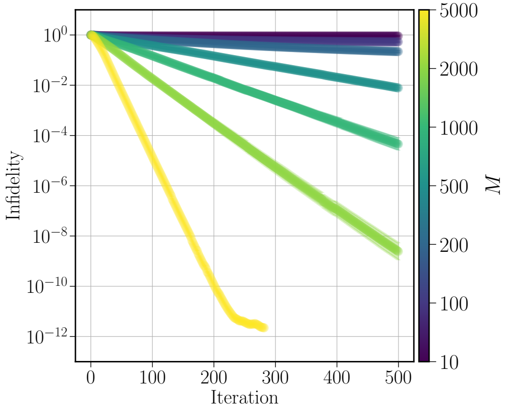
Learning Phenomena of Quantum Systems



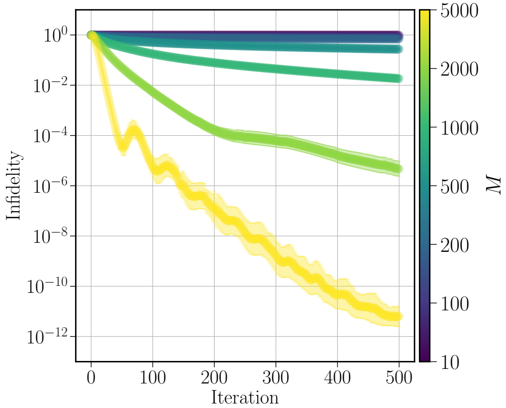
How can we leverage approaches from
quantum optimal control and *learning theory* to describe these relationships?

$$\textit{Infidelity}\text{: } 1 - \operatorname{tr}\left(\rho\rho_{\theta\gamma}\right)\ ,\ \textit{Impurity}\text{: } 1 - \operatorname{tr}\left(\rho_{\theta\gamma}^{2}\right)\ ,\ \textit{Entropy}\text{: } -\operatorname{tr}\left(\rho_{\theta\gamma}\log\rho_{\theta\gamma}\right)$$

Unconstrained vs. Constrained Optimization

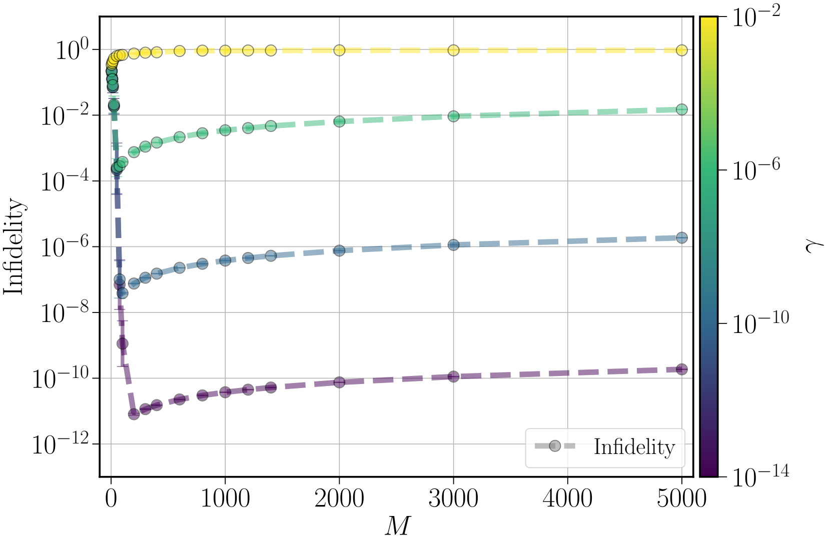


(a) Unconstrained Unitary Compilation



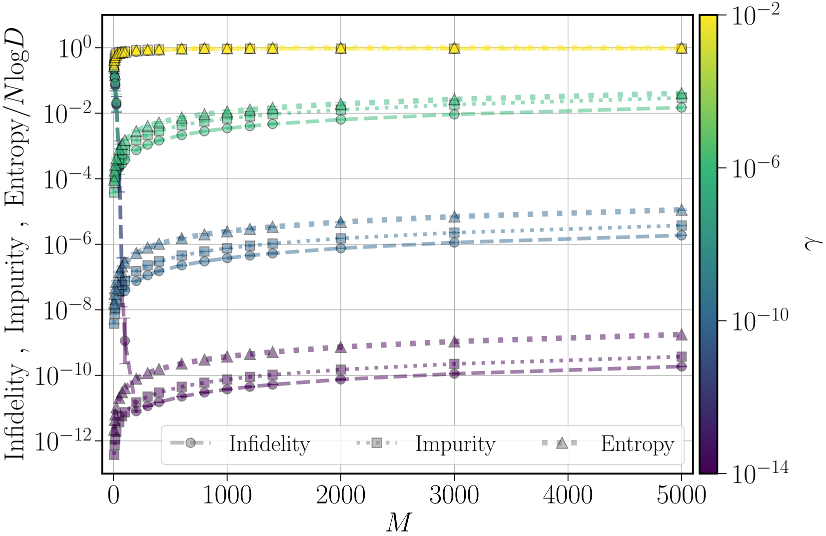
(b) Constrained Unitary Compilation

Regimes of Noisy Optimization



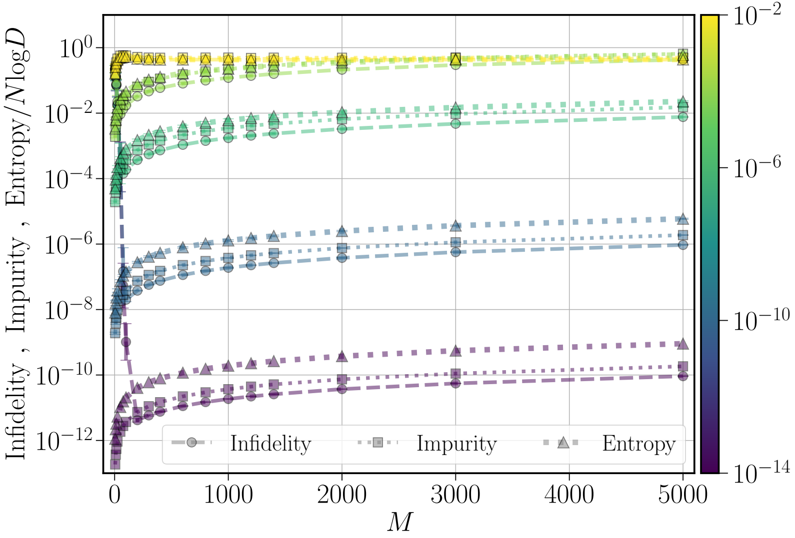
(c) Unital Dephasing for State Preparation

Regimes of Noisy Optimization



(d) Unital Dephasing for State Preparation

Regimes of Noisy Optimization



(e) Non-Unital Amplitude Damping for State Preparation

Noise Induced Critical Depth

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

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

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

$$\mathcal{L}_{\theta^*\gamma|M_{\gamma}} \sim \gamma^{\alpha} \ , \tag{2}$$

and parameterized noise channels can therefore *mitigate* approximately

$$\bar{M}_{\gamma} \sim \gamma \log 1/\gamma \quad \text{errors} \ . \tag{3}$$

Noise Induced Critical Depth

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

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

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

$$\mathcal{L}_{\theta^*\gamma|M_\gamma} \sim \gamma^\alpha \ , \tag{2}$$

and parameterized noise channels can therefore *mitigate* approximately

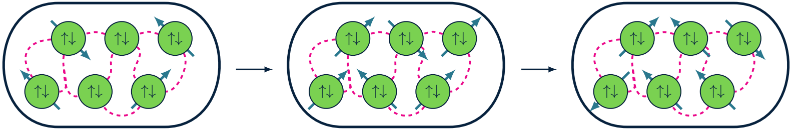
$$\bar{M}_\gamma \sim \gamma \log 1/\gamma \quad \text{errors} \ . \tag{3}$$

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

$$|\mathcal{L}_{\theta_\gamma} - \mathcal{L}_\theta| \leq 2\Big|(1-\gamma)^{NM} - 1\Big| \quad , \quad \mathcal{F}^\rho_{\theta_\gamma} \sim O\left(NM\gamma(1 - \langle\rho, \rho_{\theta_\gamma} - \rho\rangle)\right) \tag{4}$$

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
- Fidelities, purities, entropies highly correlated in $\gamma \ll 1, M \gg 1$ regime
- How can parameterized systems be applied to entropy mitigation?



Appendix

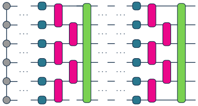
How May We Control Quantum Systems?

- 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 $\{G_\mu\}$

$$H_\theta^{(t)} = \sum_\mu \theta_\mu^{(t)} G_\mu \;\rightarrow\; U_\theta \approx \prod_t^M U_\theta^{(t)} \;:\; U_\theta^{(t)} = e^{-i\delta H_\theta^{(t)}} \approx \prod_\mu e^{-i\delta\theta_\mu^{(t)} G_\mu} \tag{5}$$

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

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



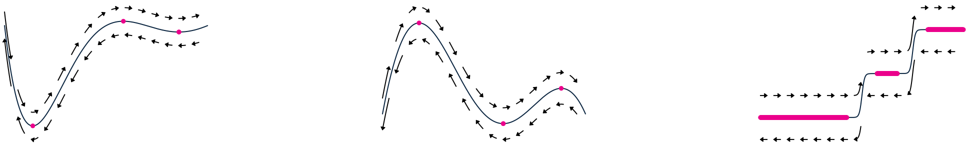
- Noise generated by constant *Kraus* operators $\{\mathcal{K}_{\gamma_\alpha}\}$

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

$$\rho \;\rightarrow\; \rho_{\Lambda_{\theta\gamma}} = \prod_t^M \mathcal{N}_\gamma \circ \mathcal{U}_\theta^{(t)}(\rho) = \prod_t^M \left[\sum_\alpha \mathcal{K}_{\gamma_\alpha}\, U_\theta^{(t)}\, \rho\, U_\theta^{(t)\dagger}\, \mathcal{K}_{\gamma_\alpha}^{\dagger} \right] \tag{7}$$

Learning Phenomena

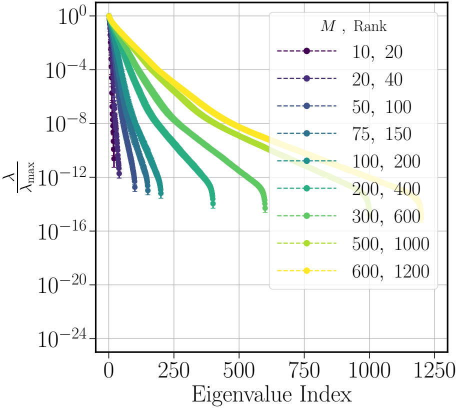
- How do optimization algorithms *learn*, and traverse the *objective landscape*?



- 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))

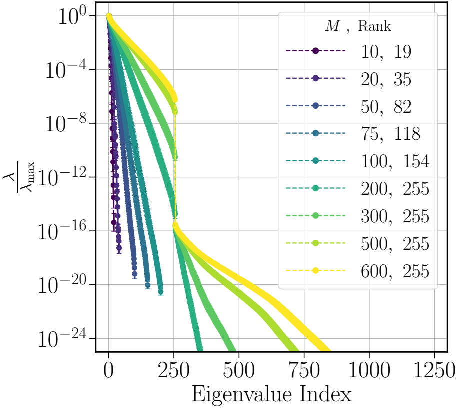
Overparameterization Phenomena

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



(f) Hessian Rank Saturation

$$\mathcal{H}_{\mu\nu} = \partial_{\mu\nu}\mathcal{L}_\theta$$

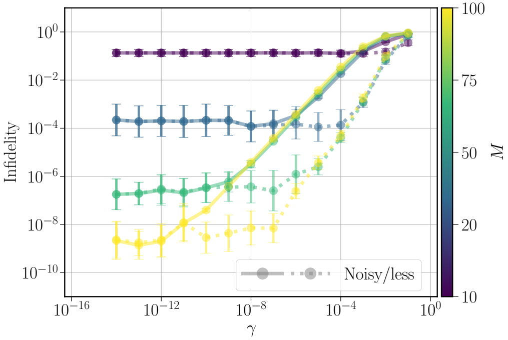


(g) Fisher Information Rank Saturation

$$\mathcal{F}_{\mu\nu} = \frac{1}{n}\mathrm{tr}\left(\partial_\mu U_\theta^\dagger \partial_\nu U_\theta\right) - \frac{1}{n^2}\mathrm{tr}\left(\partial_\mu U_\theta^\dagger U_\theta\right)\mathrm{tr}\left(U_\theta^\dagger \partial_\nu U_\theta\right)$$

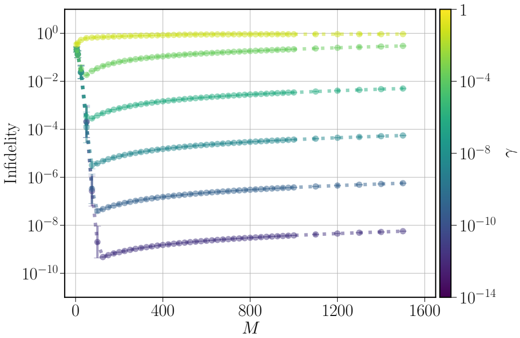
Noisy Optimization

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



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

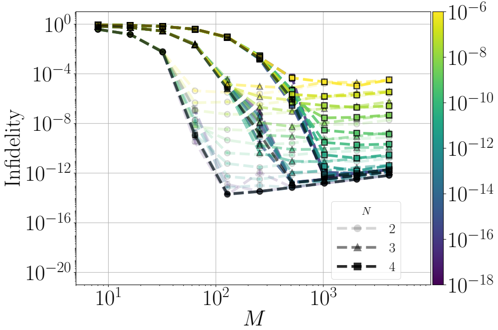
$$\partial \mathcal{L}_{\theta\gamma} \sim \sum_{\eta} \alpha_{\eta}\, \mathcal{L}_{\theta+\eta\ \gamma}$$



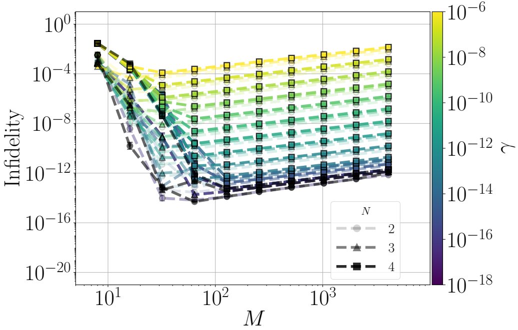
(i) Critical Depth for Noisy Infidelity

Universal Effects of Noise

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

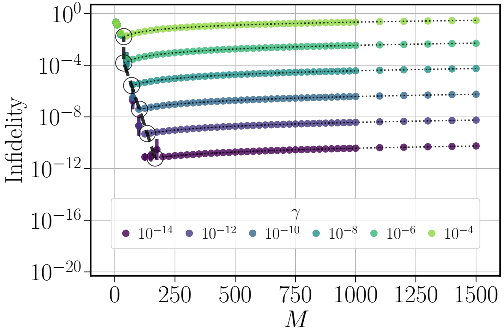


(j) Classical floating point noise for unitary compilation
$$|\mathcal{L}_{\theta_\gamma} - \mathcal{L}_\theta| \leq n^{O(NM)}\left|(1+\gamma/n)^{O(NM)} - 1\right|$$

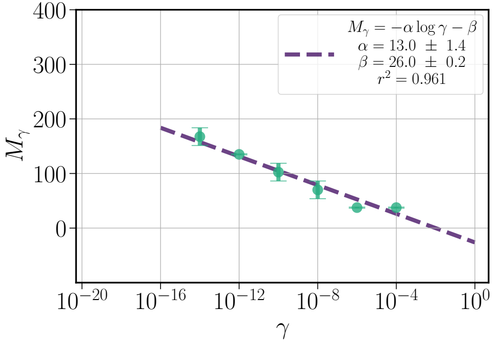


(k) Quantum dephasing noise for state preparation
$$|\mathcal{L}_{\theta_\gamma} - \mathcal{L}_\theta| \leq 2\left|(1-\gamma)^{NM} - 1\right|$$

Noise Induced Critical Depth



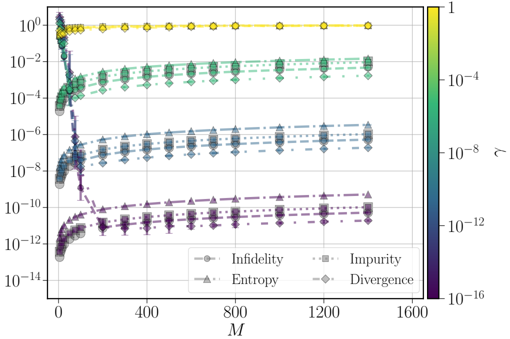
(l) Piecewise Fit of Noisy Infidelity



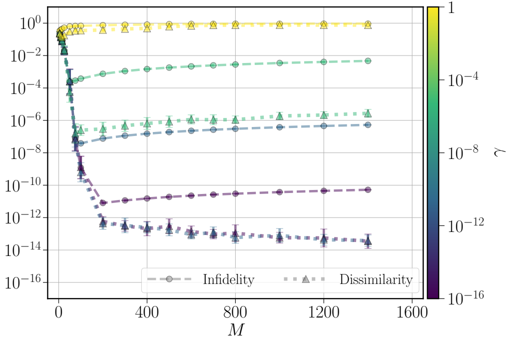
(m) Linear-Log Fit of Critical Depth

Correlated Quantities

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



(n) Impurity, Entropy, Divergence



(o) Cosine Dissimilarity