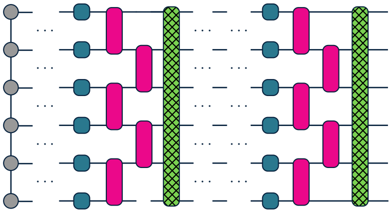
Measures of Expressivity for Quantum Channels and their Operational Meaning

Matt Duschenes*, Diego García-Martín, Martín Larocca, Zoë Holmes, Marco Cerezo

Los Alamos National Laboratory

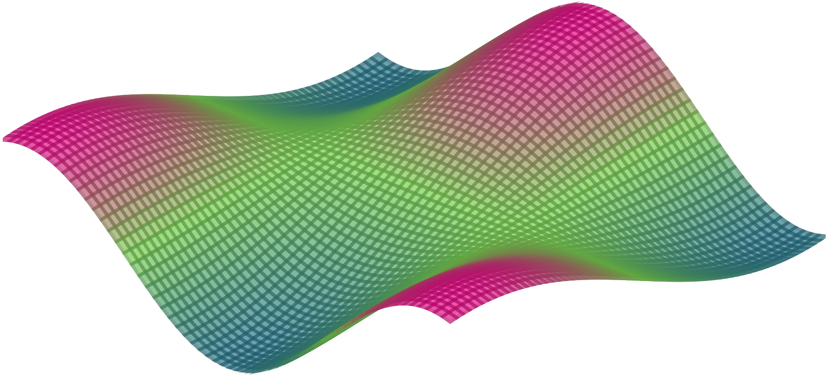
March 7, 2024

arXiv:2403.XXXXX

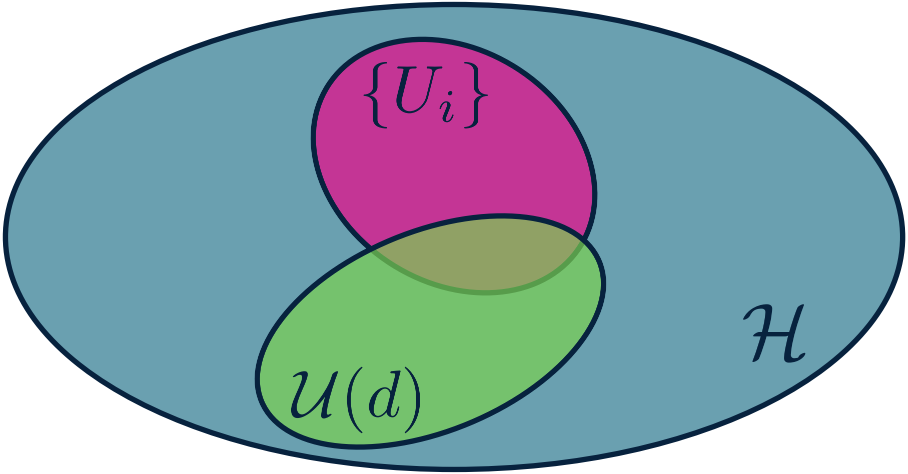


APS March Meeting 2024

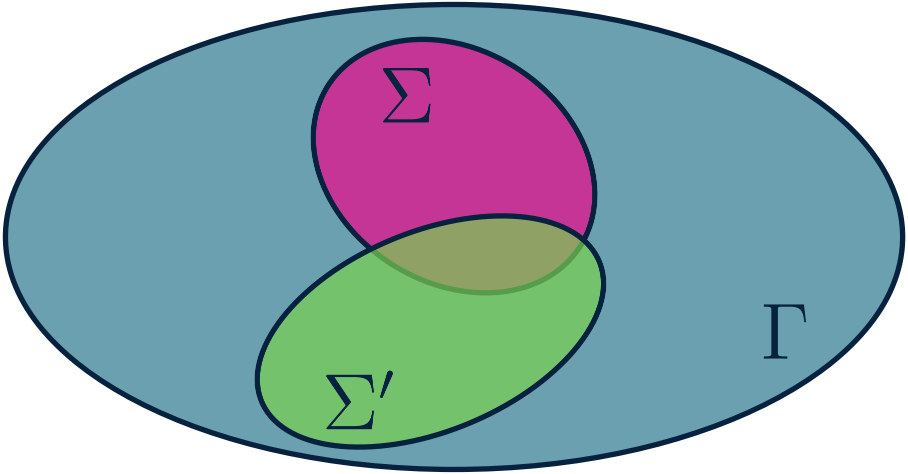
How Expressive are our Ansatze?



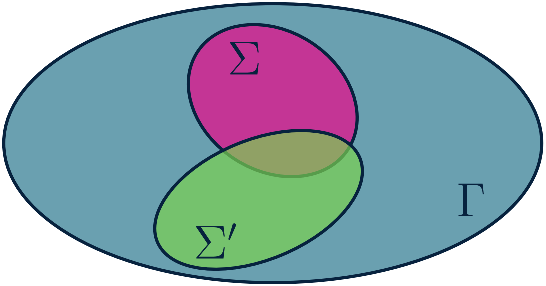
How Expressive are our Ansatze?

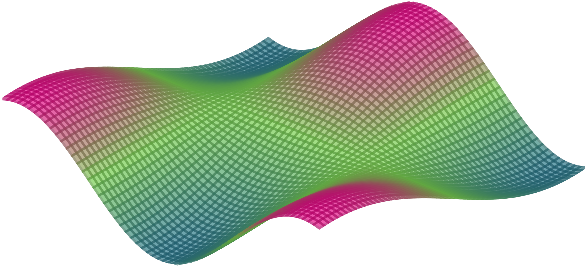


How Expressive are our Ansatze?



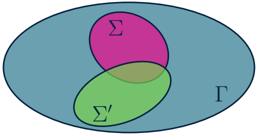
How Expressive are our Ansatze?





How Expressive are our Ansatze?

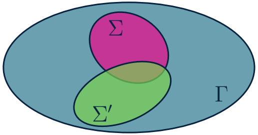
- Expressivity and trainability of *unitary ansatze* are well understood (Holmes, *et al.* , PRX Quantum, 2021)



How Expressive are our Ansatze?

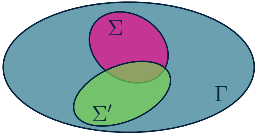
- Expressivity and trainability of *unitary ansatze* are well understood (Holmes, *et al.* , PRX Quantum, 2021)

- How does an ansatz compare to a *maximally expressive* reference ansatz?



How Expressive are our Ansatze?

- Expressivity and trainability of *unitary ansatze* are well understood (Holmes, *et al.* , PRX Quantum, 2021)
- How does an ansatz compare to a *maximally expressive* reference ansatz?
- How do generalized expressivity measures for channels depend on:
  - Noise induced phenomena
  - Underlying (parameterized) unitary evolution
  - Coupling with the environment



Expressivity Measures

- Let an *ensemble* of channels $\Lambda \sim \Sigma$ over a space $\mathcal{H}$ of dimension d have an average behaviour defined by the t-order *twirl*

$$\mathcal{T}_{\Sigma}^{(t)}(\cdot) = \int_{\Sigma} d\Lambda\ \Lambda^{\otimes t}(\cdot) \tag{1}$$

Expressivity Measures

- Let an *ensemble* of channels $\Lambda \sim \Sigma$ over a space $\mathcal{H}$ of dimension d have an average behaviour defined by the t-order *twirl*

$$\mathcal{T}_{\Sigma}^{(t)}(\cdot) = \int_{\Sigma} d\Lambda \; \Lambda^{\otimes t}(\cdot) \tag{1}$$

- This allows us to define an *expressivity* measure between ensembles

$$\mathcal{E}_{\Sigma\Sigma'}^{(t)}(\cdot) = \left\| \mathcal{T}_{\Sigma}^{(t)}(\cdot) - \mathcal{T}_{\Sigma'}^{(t)}(\cdot) \right\| \tag{2}$$

Expressivity Measures

- Let an *ensemble* of channels $\Lambda \sim \Sigma$ over a space $\mathcal{H}$ of dimension d have an average behaviour defined by the t-order *twirl*

$$\mathcal{T}_{\Sigma}^{(t)}(\cdot) = \int_{\Sigma} d\Lambda \; \Lambda^{\otimes t}(\cdot) \tag{1}$$

- This allows us to define an *expressivity* measure between ensembles

$$\mathcal{E}_{\Sigma\Sigma'}^{(t)}(\cdot) = \left\| \mathcal{T}_{\Sigma}^{(t)}(\cdot) - \mathcal{T}_{\Sigma'}^{(t)}(\cdot) \right\| \tag{2}$$

- Twirls over (quasi) *invariant* measures Σ are (quasi) *projections* onto its (quasi) *commutant* $\mathcal{S}_{\Sigma}^{(t)}$, allowing adapted *Weingarten* calculus approaches

$$\mathcal{T}_{\Sigma}^{(t)}(\cdot) \;=\; \boxed{\frac{\mathrm{tr}\,(\cdot)}{d^t} I} + \boxed{\Delta_{\Sigma}^{(t)}(\cdot)} \tag{3}$$

Expressivity Measures

- Let an *ensemble* of channels $\Lambda \sim \Sigma$ over a space $\mathcal{H}$ of dimension d have an average behaviour defined by the t-order *twirl*

$$\mathcal{T}_{\Sigma}^{(t)}(\cdot) = \int_{\Sigma} d\Lambda \; \Lambda^{\otimes t}(\cdot) \tag{1}$$

- This allows us to define an *expressivity* measure between ensembles

$$\mathcal{E}_{\Sigma\Sigma'}^{(t)}(\cdot) = \left\| \mathcal{T}_{\Sigma}^{(t)}(\cdot) - \mathcal{T}_{\Sigma'}^{(t)}(\cdot) \right\| \tag{2}$$

- Twirls over (quasi) *invariant* measures Σ are (quasi) *projections* onto its (quasi) *commutant* $\mathcal{S}_{\Sigma}^{(t)}$, allowing adapted *Weingarten* calculus approaches

$$\boxed{\mathcal{E}_{\Sigma}^{(t)}(\cdot) = \left\| \Delta_{\Sigma}^{(t)}(\cdot) \right\|} \tag{4}$$

Reference Ensembles

Channels present several choices for reference ensembles

Reference Ensembles

Channels present several choices for reference ensembles

- *Haar* $\sim$ Unitary Haar measure (uniformly random unitaries)

$$\mathcal{T}^{(t)}_{\Sigma_{\mathrm{Haar}}}(\cdot) = \int_{\mathcal{U}(d_{\mathcal{H}})} dU \; U^{\otimes t} \; \cdot \; U^{\otimes t\,\dagger} \tag{5}$$

Reference Ensembles

Channels present several choices for reference ensembles

- *Haar* ∼ Unitary Haar measure (uniformly random unitaries)

$$\mathcal{T}^{(t)}_{\Sigma_{\text{Haar}}}(\cdot) = \int_{\mathcal{U}(d_{\mathcal{H}})} dU\; U^{\otimes t} \;\cdot\; U^{\otimes t\,\dagger} \tag{5}$$

- *cHaar* ∼ Stinespring Unitary Haar measure (Kukulski, J. Math. Phys, 2020)

$$\mathcal{T}^{(t)}_{\Sigma_{\text{cHaar}}}(\cdot) = \mathrm{tr}_{\mathcal{E}}\left(\int_{\mathcal{U}(d_{\mathcal{H}\otimes\mathcal{E}})} dU\; U^{\otimes t} \;\cdot \otimes\, \nu\; U^{\otimes t\,\dagger}\right) \tag{6}$$

Reference Ensembles

Channels present several choices for reference ensembles

- *Haar* $\sim$ Unitary Haar measure (uniformly random unitaries)

$$\mathcal{T}^{(t)}_{\Sigma_{\mathrm{Haar}}}(\cdot) = \int_{\mathcal{U}(d_{\mathcal{H}})} dU \; U^{\otimes t} \;\cdot\; U^{\otimes t\,\dagger} \tag{5}$$

- *cHaar* $\sim$ Stinespring Unitary Haar measure (Kukulski, J. Math. Phys, 2020)

$$\mathcal{T}^{(t)}_{\Sigma_{\mathrm{cHaar}}}(\cdot) = \mathrm{tr}_{\mathcal{E}}\left(\int_{\mathcal{U}(d_{\mathcal{H}\otimes\mathcal{E}})} dU \; U^{\otimes t} \;\cdot \otimes \nu \; U^{\otimes t\,\dagger}\right) \tag{6}$$

- *Depolarizing* $\sim I$ (all trace preserving operations contain the identity)

$$\mathcal{T}^{(t)}_{\Sigma_{\mathrm{Depolarize}}}(\cdot) = \frac{\mathrm{tr}\,(\cdot)}{d^t_{\mathcal{H}}} I \tag{7}$$

Reference Ensembles

Channels present several choices for reference ensembles

- *Haar* $\sim$ Unitary Haar measure (uniformly random unitaries)

$$\mathcal{T}^{(t)}_{\Sigma_{\mathrm{Haar}}}(\cdot) = \int_{\mathcal{U}(d_{\mathcal{H}})} dU \; U^{\otimes t} \; \cdot \; U^{\otimes t\,\dagger} \tag{5}$$

- *cHaar* $\sim$ Stinespring Unitary Haar measure (Kukulski, J. Math. Phys, 2020)

$$\mathcal{T}^{(t)}_{\Sigma_{\mathrm{cHaar}}}(\cdot) = \mathrm{tr}_{\mathcal{E}} \left(\int_{\mathcal{U}(d_{\mathcal{H}\otimes\mathcal{E}})} dU \; U^{\otimes t} \; \cdot \otimes \nu \; U^{\otimes t\,\dagger} \right) \tag{6}$$

- *Depolarizing* $\sim I$ (all trace preserving operations contain the identity)

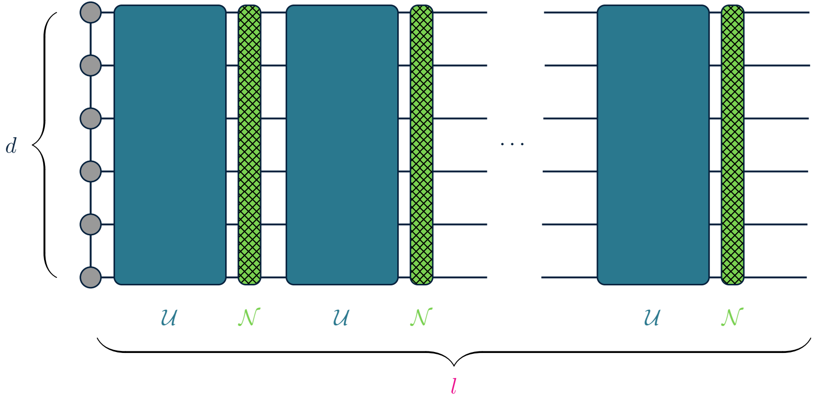
$$\mathcal{T}^{(t)}_{\Sigma_{\mathrm{Depolarize}}}(\cdot) = \frac{\mathrm{tr}\,(\cdot)}{d^t_{\mathcal{H}}} I \tag{7}$$

$$1 = \boxed{\left\| \mathcal{T}^{(t)}_{\Sigma_{\mathrm{Depolarize}}} \right\|^2} \lesssim \boxed{\left\| \mathcal{T}^{(t)}_{\Sigma_{\mathrm{cHaar}}} \right\|^2} \lesssim \boxed{\left\| \mathcal{T}^{(t)}_{\Sigma_{\mathrm{Haar}}} \right\|^2} = \left| \mathcal{S}^{(t)}_{\Sigma} \right| \tag{8}$$

Relationships between Noise and Expressivity

Analytical *expressivities* for l layers of channels

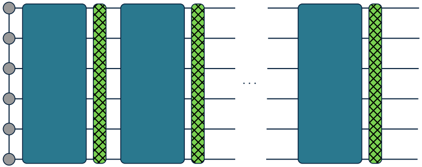
$$\Lambda^{(l)}_{\mathcal{U}_\gamma} = (\mathcal{N}_\gamma \circ \mathcal{U})^l = \frac{\mathrm{tr}\,(\cdot)}{d} I + \Delta^{(l)}_\gamma(\cdot) \tag{9}$$



Relationships between Noise and Expressivity

Haar Random Unitaries + *Unital* Pauli Noise:

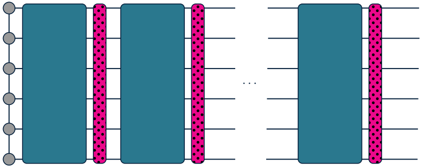
$$\mathcal{E}^{(t)}_{(\mathcal{N}_\gamma \circ \mathcal{U})^l} \sim O\left(\binom{t}{2}\gamma^l\right) \qquad \text{(Unital noise } \textit{increases} \text{ expressivity)} \tag{10}$$



Relationships between Noise and Expressivity

Haar Random Unitaries + *Non-Unital* Pauli Noise:

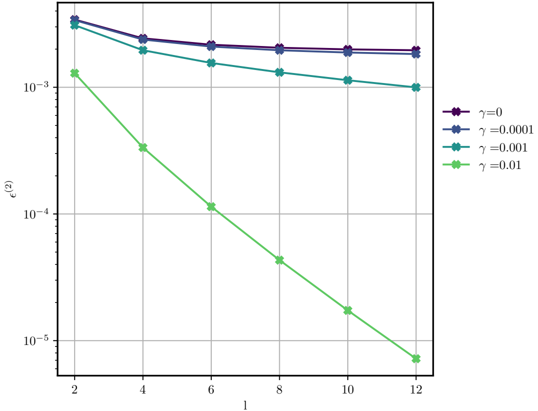
$$\mathcal{E}^{(t)}_{(\mathcal{N}_\gamma \circ \mathcal{U})^l} \sim O\left(t\gamma\right) \qquad \text{(Non-unital noise \textit{decreases} expressivity)} \tag{11}$$



Relationships between Noise and Expressivity

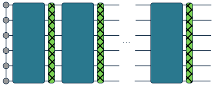
Objective $\mathcal{L}$ and Gradient $\partial\mathcal{L}$ *Concentration:* $p\left(|\mathcal{L}-\mu_{\mathcal{L}}|\geq\epsilon\right)\leq\sigma_{\mathcal{L}}^2/\epsilon^2$

$$\sigma_{\mathcal{L}}^2 \sim O\left(\frac{1}{d_{\mathcal{H}}d_{\mathcal{E}}}\right) \;+\; \|O\|_2^2\;\mathcal{E}^{(2)}_{\Sigma\Sigma'}[\rho] \quad \text{(with \textit{caveats} on } \Sigma',\rho,O \text{ locality)} \qquad (12)$$



TFIM vs. Depolarize $t=2$ expressivity as a function of depth l,
for $n=6$ qubits in the initial $|+\rangle$ state, and local depolarizing noise γ

Relationships between Noise and Expressivity



- *Haar* Random Unitaries + *Unital* Pauli Noise:

$$\mathcal{E}^{(t)}_{(\mathcal{N}_\gamma \circ \mathcal{U})^l} \sim O\left(\binom{t}{2}\gamma^l\right) \qquad \text{(Unital noise \textit{increases} expressivity)} \tag{10}$$

- *Haar* Random Unitaries + *Non-Unital* Pauli Noise:

$$\mathcal{E}^{(t)}_{(\mathcal{N}_\gamma \circ \mathcal{U})^l} \sim O\left(t\gamma\right) \qquad \text{(Non-unital noise \textit{decreases} expressivity)} \tag{11}$$

- Objective $\mathcal{L}$ and Gradient $\partial\mathcal{L}$ *Concentration*: $p\left(|\mathcal{L}-\mu_{\mathcal{L}}| \geq \epsilon\right) \leq \sigma^2_{\mathcal{L}}/\epsilon^2$

$$\sigma^2_{\mathcal{L}} \sim O\left(\frac{1}{d_{\mathcal{H}} d_{\mathcal{E}}}\right) + \|O\|_2^2\ \mathcal{E}^{(2)}_{\Sigma\Sigma'}[\rho] \qquad \text{(with \textit{caveats} on } \Sigma', \rho, O \text{ locality)} \tag{12}$$

Operational Meaning of Channel Expressivity Measures

- *Noise induced* phenomena are actually channel expressivity phenomena!

Operational Meaning of Channel Expressivity Measures

- *Noise induced* phenomena are actually channel expressivity phenomena!
- Many *subtle* differences between ensembles of unitary and non-unitary channels

Operational Meaning of Channel Expressivity Measures

- *Noise induced* phenomena are actually channel expressivity phenomena!
- Many *subtle* differences between ensembles of unitary and non-unitary channels
- Choice of reference ensemble reflects retained *coherent information*

Operational Meaning of Channel Expressivity Measures

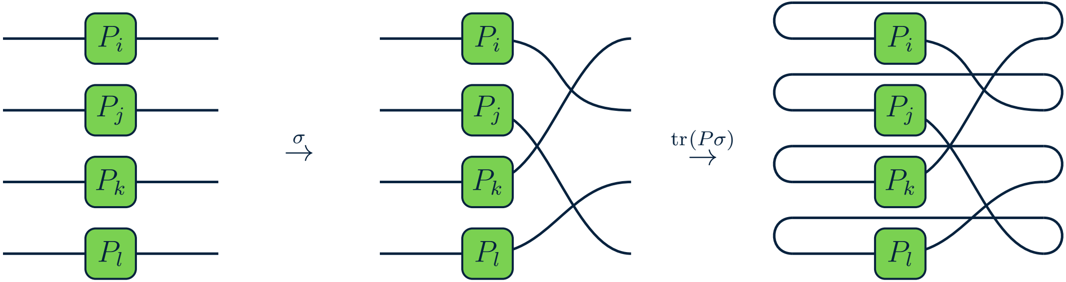
- *Noise induced* phenomena are actually channel expressivity phenomena!
- Many *subtle* differences between ensembles of unitary and non-unitary channels
- Choice of reference ensemble reflects retained *coherent information*
- Are there *practical* ensembles of channels that approach t-designs?

Operational Meaning of Channel Expressivity Measures

- *Noise induced* phenomena are actually channel expressivity phenomena!
- Many *subtle* differences between ensembles of unitary and non-unitary channels
- Choice of reference ensemble reflects retained *coherent information*
- Are there *practical* ensembles of channels that approach t-designs?
- Are there relationships between channel expressivity and their *simulability*?

Appendix

Diagrammatic Expansions of Permutations



$P = P_i \otimes P_j \otimes P_k \otimes P_l$

$\sigma = (ijlk)$

$\mathrm{tr}\,(P\sigma) = \mathrm{tr}\left(\prod_{\mu \in \sigma} P_\mu\right)$

$$\mathcal{T}_{\Sigma}^{(t)}(\cdot) = \frac{\mathrm{tr}\,(\cdot)}{d^t} I \; + \; \frac{1}{d^t} \sum_{P \in \mathcal{P}_d^{(\mathcal{S}_{\Sigma}^{(t)})} \setminus \{I\}} \tau_d^{(t)}(P, \cdot)\; P \tag{13}$$

Haar, cHaar, and Depolarizing Ensembles

Σ \ t	1	2
Haar	$\frac{\mathrm{tr}(\cdot)}{d_{\mathcal{H}}} I$	$\frac{\mathrm{tr}(\cdot)}{d_{\mathcal{H}}^2} I + \frac{1}{d_{\mathcal{H}}^2} \sum_{P \in \mathcal{P}_{d_{\mathcal{H}}}^{(\mathcal{S}_{\Sigma}^{(2)})} \setminus \{I\}} \tau_{d_{\mathcal{H}}}^{(2)}(P, \cdot) \; P \otimes P^{-1}$
cHaar	$\frac{\mathrm{tr}(\cdot)}{d_{\mathcal{H}}} I$	$\frac{\mathrm{tr}(\cdot)}{d_{\mathcal{H}}^2} I + \frac{1}{d_{\mathcal{H}}^2} \sum_{P \in \mathcal{P}_{d_{\mathcal{H}}}^{(\mathcal{S}_{\Sigma}^{(2)})} \setminus \{I\}} \tau_{d_{\mathcal{H}}, d_{\mathcal{E}}}^{(2)}(P, \cdot) \; P \otimes P^{-1}$
Depolarizing	$\frac{\mathrm{tr}(\cdot)}{d_{\mathcal{H}}^t} I$	

Table 1: Twirls $\mathcal{T}_{\Sigma}^{(t)}$ for various ensembles and moments

Subtleties with Concentration of Channel Ansatz

Given $\Lambda \sim \mathcal{U}_\theta$ with involutory generators $U_\theta = e^{-i\theta G}$, and pure inputs ρ:

Objective $\mathcal{L}_\Lambda^O$ variance concentrates as

$$\sigma^2_{\mathcal{L}_\Lambda^O|\Sigma}[\rho] \leq \begin{cases} O\left(\frac{1}{d_{\mathcal{H}} d_{\mathcal{E}}}\right) \; + \; \min\left\{\|O\|_2^2 \mathcal{E}^{(2)}_{\Sigma\Sigma'}[\rho] \; , \; \|O\|_\infty^2 \mathcal{E}^{(2|\diamond)}_{\Sigma\Sigma'}\right\} & \left\{O_{\text{Pauli}} \; , \; \Sigma'_{\text{cHaar}}\right\} \\ O\left(\frac{1}{d_{\mathcal{H}}^2}\right) \; + \; \min\left\{\|O\|_2^2 \mathcal{E}^{(2)}_{\Sigma\Sigma'}[\rho] \; , \; \|O\|_\infty^2 \mathcal{E}^{(2|\diamond)}_{\Sigma\Sigma'}\right\} & \left\{O_{\text{Projector}} \; , \; \Sigma'_{\text{cHaar}}\right\} \\ 0 \; + \; \min\left\{\|O\|_2^2 \mathcal{E}^{(2)}_{\Sigma\Sigma'}[\rho] \; , \; \|O\|_\infty^2 \mathcal{E}^{(2|\diamond)}_{\Sigma\Sigma'}\right\} & \left\{O_{\text{Pauli}} \; , \; \Sigma'_{\text{Depolarize}}\right\} \\ O\left(\frac{1}{d_{\mathcal{H}}^2}\right) \; + \; \min\left\{\|O\|_2^2 \mathcal{E}^{(2)}_{\Sigma\Sigma'}[\rho] \; , \; \|O\|_\infty^2 \mathcal{E}^{(2|\diamond)}_{\Sigma\Sigma'}\right\} & \left\{O_{\text{Projector}} \; , \; \Sigma'_{\text{Depolarize}}\right\} \end{cases} \tag{14}$$

Subtleties with Concentration of Channel Ansatz

Given $\Lambda \sim \mathcal{U}_\theta$ with involutory generators $U_\theta = e^{-i\theta G}$, and pure inputs ρ:

Objective gradient $\partial_\mu \mathcal{L}^O_\Lambda$ variance concentrates as

$$\sigma^2_{\partial_\mu \mathcal{L}^O_\Lambda | \Sigma\Sigma'_{RL}}[\rho] \leq \sigma^{2RL}_{\partial_\mu \mathcal{L}^O_\Lambda | \Sigma'_{\mu RL}}[\rho] + \begin{cases} \begin{pmatrix} O\left(\frac{1}{d_\mathcal{E}}\right) \mathcal{E}^{(2|\diamond)}_{\Sigma_{\mu R}\Sigma'_{\mu R}} + O\left(\frac{1}{d_\mathcal{E} d^2_\mathcal{H}}\right) \mathcal{E}^{(2|\diamond)}_{\Sigma_{\mu l}\Sigma'_{\mu l}} + \\ O(1)\, \mathcal{E}^{(2|\diamond)}_{\Sigma_{\mu R}\Sigma'_{\mu R}} \mathcal{E}^{(2|\diamond)}_{\Sigma_{\mu l}\Sigma'_{\mu l}} \end{pmatrix} & \left\{ \begin{matrix} O_{\text{Pauli}} , \\ \Sigma'_{\text{cHaar}} \end{matrix} \right\} \\ \begin{pmatrix} O\left(\frac{1}{d_\mathcal{E} d_\mathcal{H}}\right) \mathcal{E}^{(2|\diamond)}_{\Sigma_{\mu R}\Sigma'_{\mu R}} + O\left(\frac{1}{d_\mathcal{E} d^2_\mathcal{H}}\right) \mathcal{E}^{(2|\diamond)}_{\Sigma_{\mu l}\Sigma'_{\mu l}} + \\ O(1)\, \mathcal{E}^{(2|\diamond)}_{\Sigma_{\mu R}\Sigma'_{\mu R}} \mathcal{E}^{(2|\diamond)}_{\Sigma_{\mu l}\Sigma'_{\mu l}} \end{pmatrix} & \left\{ \begin{matrix} O_{\text{Projector}} , \\ \Sigma'_{\text{cHaar}} \end{matrix} \right\} \\ \left(O(1)\, \mathcal{E}^{(2|\diamond)}_{\Sigma_{\mu R}\Sigma'_{\mu R}} \mathcal{E}^{(2|\diamond)}_{\Sigma_{\mu R}\Sigma'_{\mu l}} \right) & \left\{ \begin{matrix} O_{\text{Pauli}} , \\ \Sigma'_{\text{Depolarize}} \end{matrix} \right\} \\ \left(O(1)\, \mathcal{E}^{(2|\diamond)}_{\Sigma_{\mu R}\Sigma'_{\mu R}} \mathcal{E}^{(2|\diamond)}_{\Sigma_{\mu R}\Sigma'_{\mu l}} \right) & \left\{ \begin{matrix} O_{\text{Projector}} , \\ \Sigma'_{\text{Depolarize}} \end{matrix} \right\} \end{cases} \tag{15}$$

where the *left* (l) and *right* (R) 2-design gradient variance is

$$\sigma^{2RL}_{\partial_\mu \mathcal{L}^O_\Lambda | \Sigma'_{\mu RL}}[\rho] = \begin{cases} O\left(\frac{1}{d^2_\mathcal{E} d_\mathcal{H}}\right) & \{O_{\text{Pauli}} , \Sigma'_{\text{cHaar}}\} \\ O\left(\frac{1}{d^2_\mathcal{E} d^2_\mathcal{H}}\right) & \{O_{\text{Projector}} , \Sigma'_{\text{cHaar}}\} \\ 0 & \{O_{\text{Pauli}} , \Sigma'_{\text{Depolarize}}\} \\ 0 & \{O_{\text{Projector}} , \Sigma'_{\text{Depolarize}}\} \end{cases} \tag{16}$$