Generalization with quantum geometry for learning unitaries

Tobias Haug

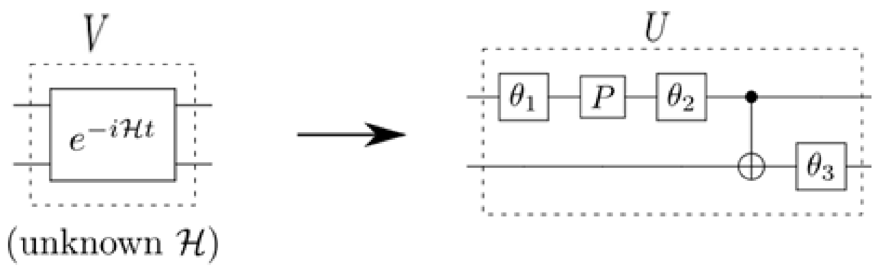
20.11.2023

T. Haug, M.S. Kim, Generalization with quantum geometry for learning unitaries arXiv:2303.13462 (2023)

T. Haug, K. Bharti, M.S. Kim, Capacity and quantum geometry of parametrized quantum circuits PRX Quantum 2 (4), 040309 (2021)

QML routine: Learn unitary

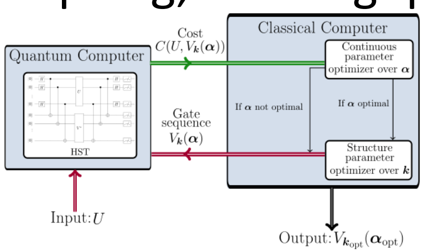
Task: Given target unitary *V*, represent it with **M** parameter unitary *U(***θ***)*



- Target unitary *V* not known directly, only its action on ***L*** training states $|\psi_\ell\rangle \in W$ drawn from distribution *W:* $S_L = \{|\psi_\ell\rangle, V|\psi_\ell\rangle\}_{\ell=1}^{L}$
- Goal: Find $U(\boldsymbol{\theta})\,|\psi_\ell\rangle = V\,|\psi_\ell\rangle$
- Training: Optimize fidelity on training data

$$C_{\text{train}}(\boldsymbol{\theta}, S_L) = 1 - \frac{1}{L}\sum_{\ell=1}^{L} |\langle\psi_\ell|V^\dagger U(\boldsymbol{\theta})|\psi_\ell\rangle|^2$$

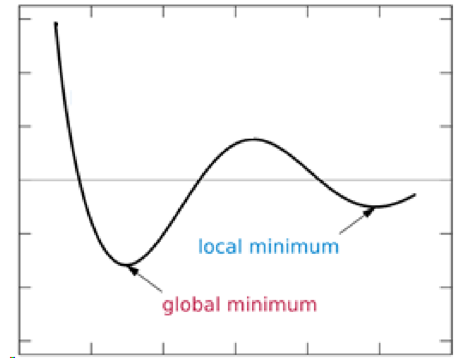
- Routine used for quantum compiling, learning quantum dynamics, quantum autoencoder, …

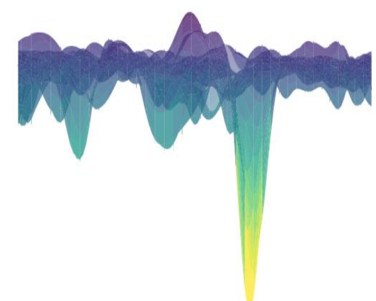


Khatri, Sumeet, et al. "Quantum-assisted quantum compiling." *Quantum* 3 (2019): 140.

Challenge 1: Converge to global minima

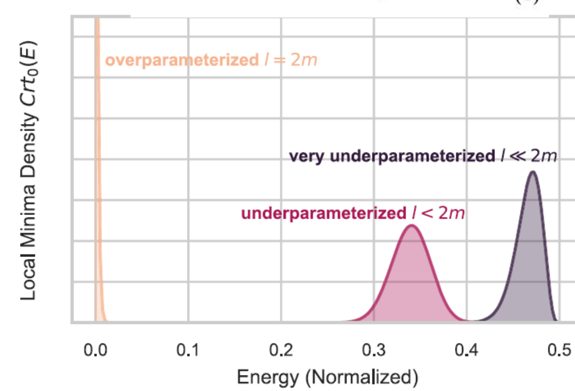
- Train circuit parameters **θ** (using **L** training datapoints) by minimizing training cost $C_{\text{train}}(\boldsymbol{\theta})$
- Success depends on circuit parameters **M**
  - Underparameterized models swamped with bad local minima (far away from global minima)
    →Optimization gets stuck in bad solutions
  - Overparameterized: Local minimas become global minimas
    →Easy to find good solutions





- Critical number of circuit parameters $\mathbf{M_c(L)}$ to overparameterize?
  - $\mathbf{M_c(L=1)}$ from rank of "quantum Fisher information metric"

- What is $\mathbf{M_c}$ **(L)** for L>1?

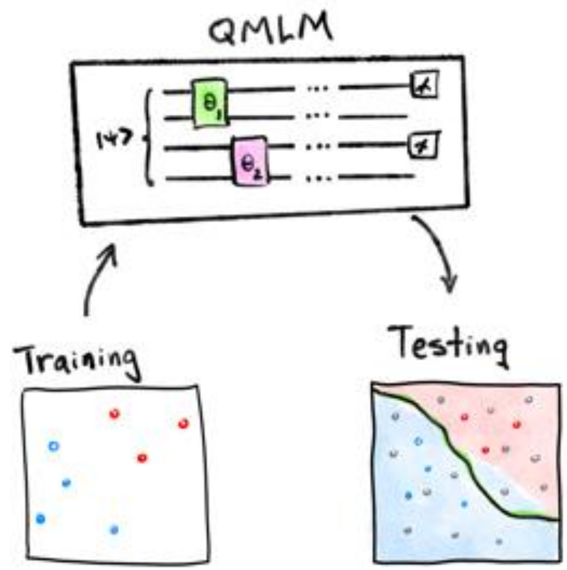


X. You, S. Chakrabarti, and X Wu. "A convergence theory for over-parameterized variational quantum eigensolvers." *arXiv:2205.12481* (2022).
E. Anschuetz, and B. Kiani. "Quantum variational algorithms are swamped with traps." *Nature Communications* 13.1 (2022)
M. Larocca, N. Ju, D. Garcıa-Martın, P. J. Coles, and M. Cerezo, Nature Computational Science 3, 542 (2023)
T. Haug, K. Bharti, M.S. Kim, Capacity and quantum geometry of parametrized quantum circuits PRX Quantum 2 (4), 040309 (2021)

Challenge 2: Generalization

- Generalization: Performance of trained model on unseen test data (test error)

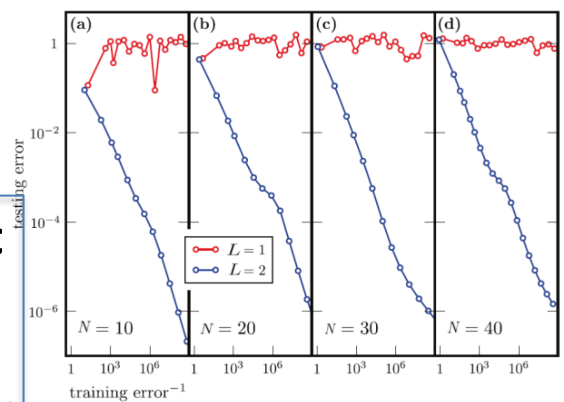
$$C_{\text{test}}(\boldsymbol{\theta}, W) = 1 - \mathop{\mathbb{E}}_{|\psi\rangle \in W}[|\langle\psi|V^{\dagger}U(\boldsymbol{\theta})|\psi\rangle|^2]$$



- More training data **L** → better test error

- Loose uniform generalization bounds: $\Delta C_{\text{test}} \sim \frac{1}{\sqrt{L}}$

- In practice: Numerics suggest we can achieve $\mathbf{L_c}$ **=const** data to generalize, but not clear when or why
- Can we get accurate values for $\mathbf{L_c}$



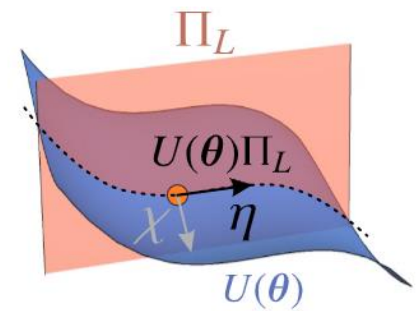
E. Gil-Fuster, J. Eisert, and C. Bravo-Prieto, Understanding quantum machine learning also requires rethinking generalization, arXiv:2306.13461

L. Banchi, J. Pereira, and S. Pirandola. "Generalization in quantum machine learning: A quantum information standpoint." *PRX Quantum* 2.4 (2021)

C. Matthias et al. "Generalization in quantum machine learning from few training data." *Nature communications* 13.1 (2022): 4919.

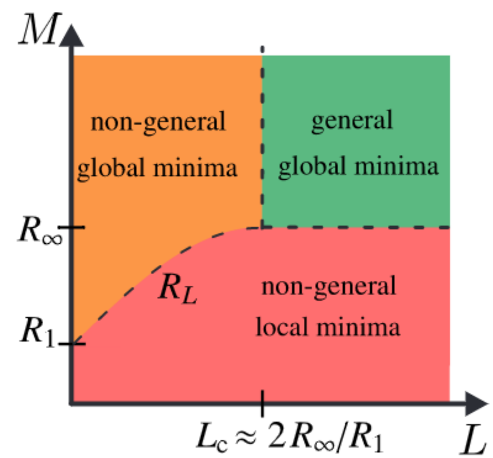
Our contributions

- We give theory of overparameterization and generalization for learning unitaries via
  “Data quantum Fisher information metric” (DQFIM)



$\dim(\eta) \sim R_L$

- Explain generalization from few data via rank of DQFIM and dynamical Lie algebra



- Surprising implications:
  - Symmetries can hurt generalization
  - Out-of-distribution learning (learning from “wrong distribution”) can improve generalization

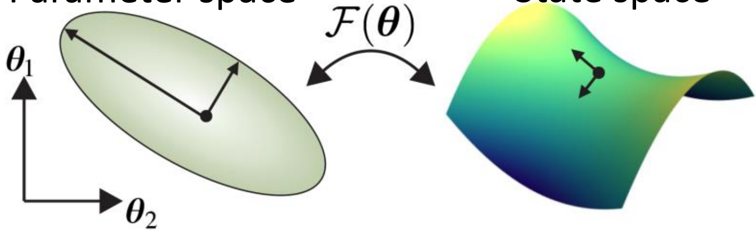
T. Haug, M.S. Kim, Generalization with quantum geometry for learning unitaries arXiv:2303.13462 (2023)
T. Haug, K. Bharti, M.S. Kim, Capacity and quantum geometry of parametrized quantum circuits PRX Quantum 2 (4), 040309 (2021)

Data quantum Fisher information metric

- Quantum Fisher information metric: Describes how change in parameter **θ** affects quantum state ("Hessian of fidelity")

$$\mathcal{F}_{nm}(|\psi(\boldsymbol{\theta})\rangle) = 4\mathrm{Re}(\langle\partial_n\psi|\partial_m\psi\rangle - \langle\partial_n\psi|\psi\rangle\langle\psi|\partial_m\psi\rangle)$$

$$|\langle\psi(\boldsymbol{\theta})|\psi(\boldsymbol{\theta}+\mathrm{d}\boldsymbol{\theta})\rangle|^2 = 1 - \frac{1}{4}\mathcal{F}_{nm}\mathrm{d}\boldsymbol{\theta}^n\mathrm{d}\boldsymbol{\theta}^m$$



- Maximal rank of $\mathcal{F}_{nm}$ determines overparameterization
- Problem: Incomplete, does not characterize learning from L>1 datapoints

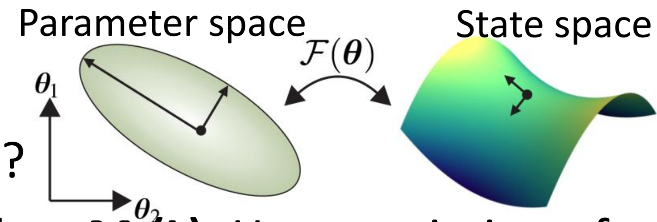
- Solution: We introduce "Data quantum Fisher information metric" (DQFIM)

$$Q_{nm}(S_L, U(\boldsymbol{\theta})) = 4\mathrm{Re}(\mathrm{tr}(\partial_n U^\dagger \tilde{\Pi}_L \partial_m U) - \mathrm{tr}(\partial_n U^\dagger \tilde{\Pi}_L U)\mathrm{tr}(U^\dagger \tilde{\Pi}_L \partial_m U))$$

- "Quantum Fisher information metric for parameterized unitary *U(***θ***)* projected onto dataset $\tilde{\Pi}_L$"
- *L=1* reduces to quantum Fisher information metric

T. Haug, and M. S. Kim. "Generalization with quantum geometry for learning unitaries." *arXiv:2303.13462* (2023).
T. Haug, K. Bharti, M.S. Kim *Capacity and quantum geometry of parametrized quantum circuits*, PRX Quantum 2 (2021)

Overparameterization and rank of DQFIM



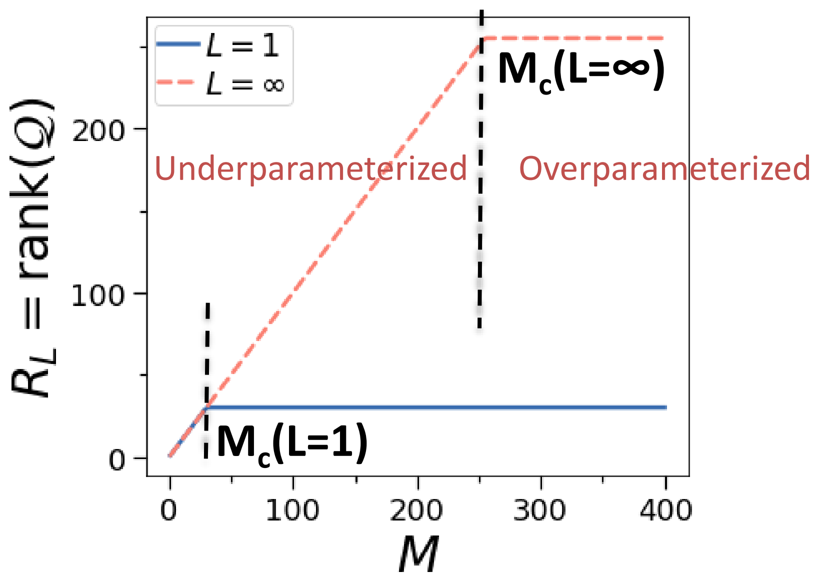
How many circuit parameters **M** to overparameterize?

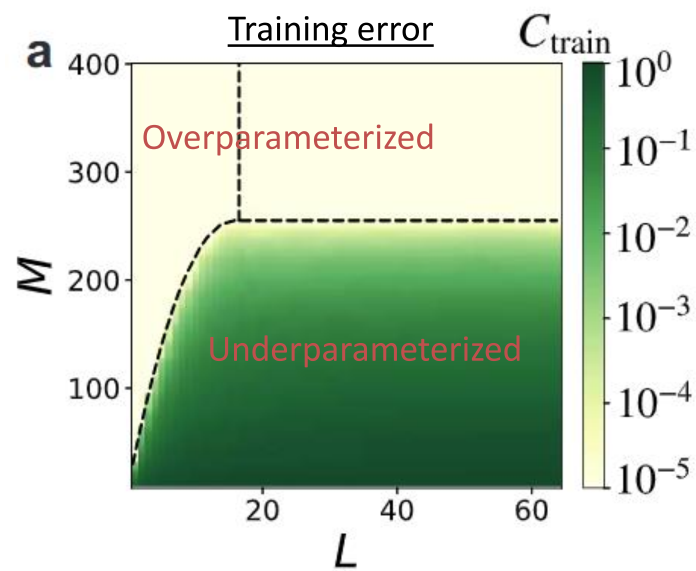
- Increase **M** until rank of DQFIM becomes maximal, at **M_c(L)**: Here, variation of **θ** explores every direction of unitary space projected onto data

→ Critical number of circuit parameters **M_c(L)** to overparameterize

$$R_L = \max_{M \geq M_c, \boldsymbol{\theta}} \operatorname{rank}[\mathcal{Q}(S_L, U(\boldsymbol{\theta}, M))]$$

- $M_c(L) \sim R_L$





Generalization and rank of DQFIM

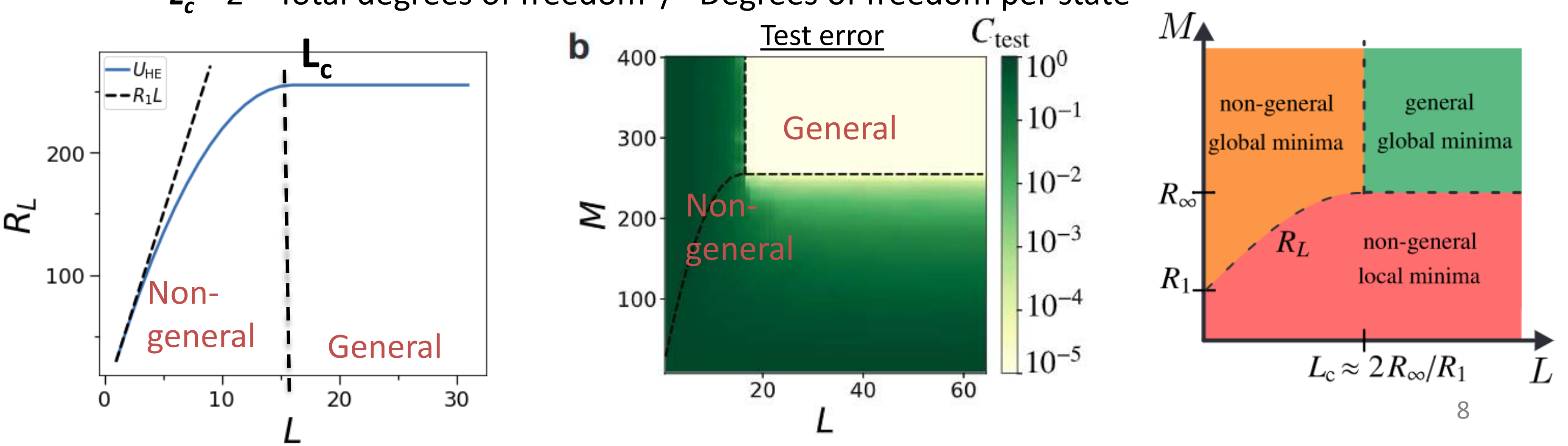
- Dataset size $\boldsymbol{L_c}$ to generalize?
- Rank DQFIM R_L describes degrees of freedom we can learn with dataset size $\boldsymbol{L}$
- Increase $\boldsymbol{L}$ until rank of DQFIM R_L becomes maximal, at $\boldsymbol{L_c}$

$$R_\infty = \max_{L \geq L_c} R_L(S_L, U)$$

- Maximal rank of DQFIM→ Dataset has complete information → Generalisation

  (We assume data without noise)

- Approximation: $L_c \approx 2R_\infty / R_1$
  - $\boldsymbol{L_c}$~ 2*"Total degrees of freedom"/ "Degrees of freedom per state"



Learning unitaries and dynamical Lie algebra

- Dynamical Lie algebra characterizes dimensionality of unitary
- We prove upper bounds on rank R_L in terms of dynamical Lie algebra

Theorem 1. *The maximal rank R_L is bounded by the dimension of the dynamical Lie algebra (DLA)*

$$R_L \leq \dim(\mathfrak{g}) \,, \tag{10}$$

where $\mathfrak{g} = \mathrm{span}\,\langle iH_1, \dots, iH_K \rangle_{Lie}$ is generated by the repeated nested commutators of the generators H_k of $U(\boldsymbol{\theta})$.

→ Polynomial sized dynamical Lie algebra implies generalisation with polynomial depth and dataset size

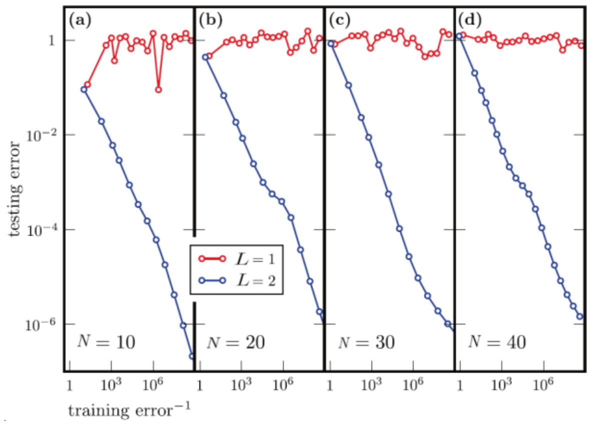
$$\dim(\mathfrak{g}) \sim \mathrm{poly}(N) \rightarrow \quad \begin{aligned} M_{\mathrm{c}} &\sim \mathrm{poly}(N) \\ L_{\mathrm{c}} &\sim \mathrm{poly}(N) \end{aligned}$$

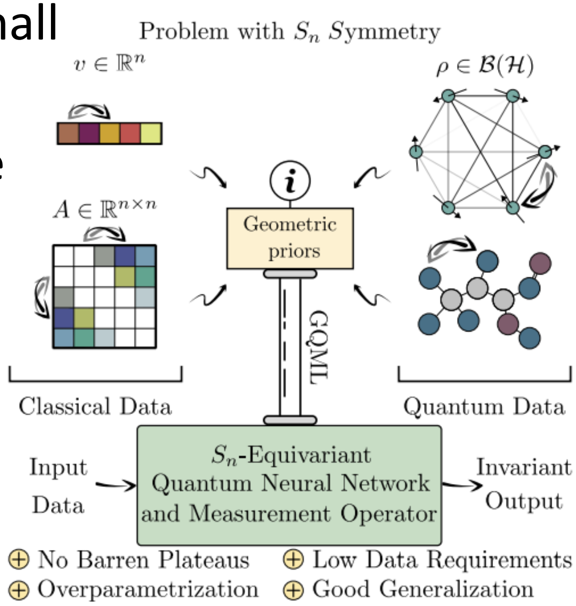
Symmetries and QML

- QML problems with symmetries

→ Encode symmetries into ansatz circuit and data→Small dynamical Lie algebra

- Previous numerics found L_c=const data to generalize





- Explained with DQFIM: Small dynamical Lie algebra implies low rank of DQFIM

→ $L_\text{c} \approx 2R_\infty / R_1 = \text{const}$

C. Matthias et al. "Generalization in quantum machine learning from few training data." *Nature communications* 13.1 (2022): 4919.

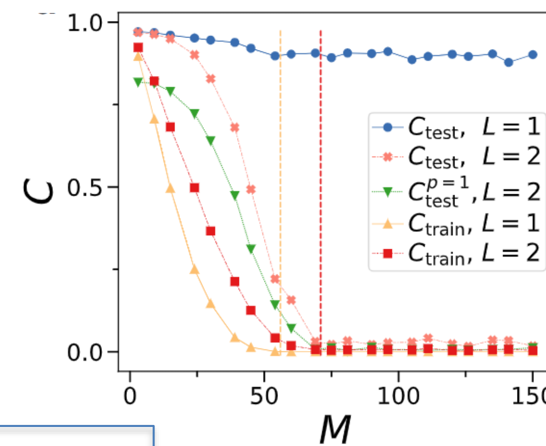
J. Gibbs, et al. "Dynamical simulation via quantum machine learning with provable generalization"*arXiv:2204.10269*

M. Larocca, et al. "Group-invariant quantum machine learning." *PRX Quantum* 3.3 (2022): 030341.

J.J. Meyer, et al. "Exploiting symmetry in variational quantum machine learning." *PRX Quantum* 4.1 (2023)

Are symmetries always good?

- Symmetry operator *P*
- Parameterized unitary respects symmetry $[U(\boldsymbol{\theta}), P] = 0$
- Data
  - Symmetric $P\left|\psi_i\right\rangle = \left|\psi_i\right\rangle$
  - Non-symmetric $P\left|\psi_i\right\rangle \neq \left|\psi_i\right\rangle$
- Generalization for excitation number symmetry:
  - Symmetric data: $L_c \sim N$
  - Non-symmetric: $L_c \sim 2$ datapoints → no symmetry better!



→ Encoding symmetry of problem onto data may hurt generalization!

→ Why? Non-symmetric data "explores" more degrees of freedom of unitary

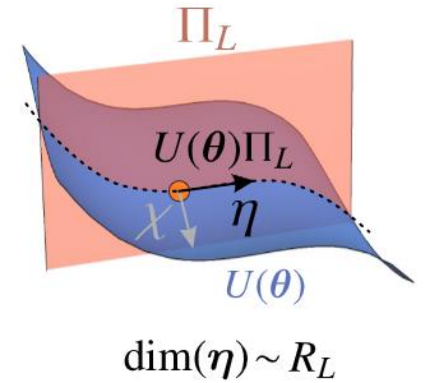
- Catch: Non-symmetric data needs more circuit parameters $\mathbf{M_c}$

We also show: Out-of-distribution learning (train and test data from different distributions) can generalize better than in-distribution (train and test data from same distribution)

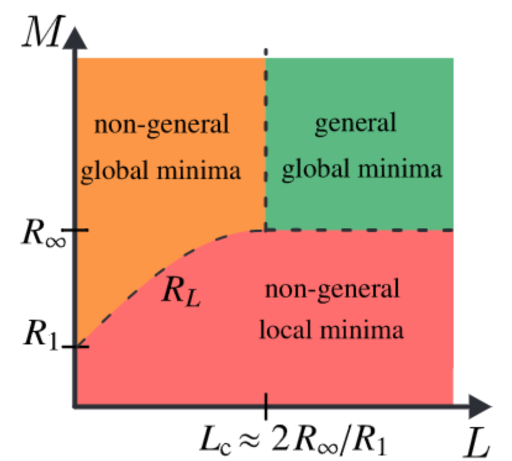
C. Matthias et al. "Generalization in quantum machine learning from few training data." *Nature communications* 13.1 (2022): 4919.
J. Gibbs, et al. "Dynamical simulation via quantum machine learning with provable generalization"*arXiv:2204.10269*

Conclusion

- "Data quantum Fisher information metric" (DQFIM) provides complete theory for generalisation and overparameterization for learning unitaries
- More symmetries may need more data to generalise
- Out-of-distribution learning can be superior



- Next
  - Extension beyond learning unitary problem
  - Effect of noise
  - Classify symmetries with constant dataset size for generalisation?
  - Phase transition in generalization?
  - Combine with information-theoretic perspectives?



T. Haug, M.S. Kim, Generalization with quantum geometry for learning unitaries arXiv:2303.13462 (2023)
T. Haug, K. Bharti, M.S. Kim, Capacity and quantum geometry of parametrized quantum circuits PRX Quantum 2 (4), 040309 (2021)