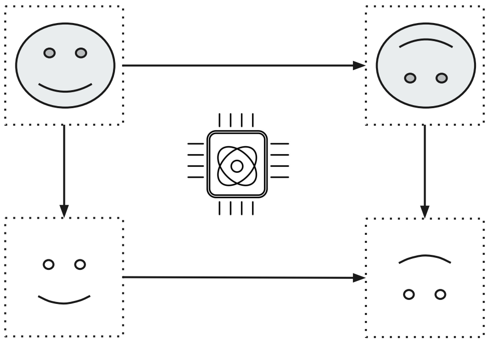
UNIVERSITAT DE BARCELONA | ICCUB | Institut de Ciències del Cosmos UNIVERSITAT DE BARCELONA | EXCELENCIA MARÍA DE MAEZTU 04/2025-03/2031 | HadNucAtUB

Adding equivariance to Variational Quantum Eigensolvers



Joan Ainaud Fondevila

Advisors: Arnau Rios, Javier Rozalén, Alejandro Romero-Ros

HADNUCMAT workshop **2026**

27 January 2026

Quick Recap: VQE

Finding the ground state is **hard**

Exploit Variational Principle →

$$E_0 = \min_{\psi} \frac{\langle\psi| \hat{H} |\psi\rangle}{\langle\psi|\psi\rangle} \leq E_{\psi}$$

Quick Recap: VQE

Finding the ground state is **hard**
Exploit Variational Principle →

$$E_0 = \min_{\psi} \frac{\langle\psi| \hat{H} |\psi\rangle}{\langle\psi|\psi\rangle} \leq E_\psi$$

Propose *ansatz*:
Minimize over θ

$$\left|\psi(\vec{\theta})\right\rangle$$

- *Classical*
  Analytical expression, Neural Quantum State, etc

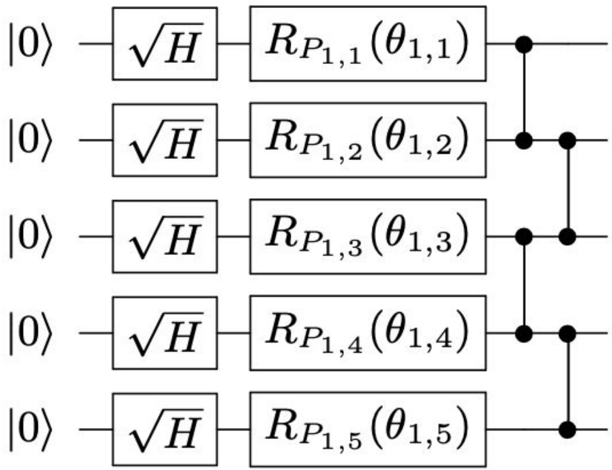
Quick Recap: VQE

Finding the ground state is **hard**
Exploit Variational Principle →

$$E_0 = \min_{\psi} \frac{\langle\psi| \hat{H} |\psi\rangle}{\langle\psi|\psi\rangle} \leq E_\psi$$

Propose *ansatz*: $\left|\psi(\vec{\theta})\right\rangle$
Minimize over θ

- *Classical*
- Use *Quantum Computer*



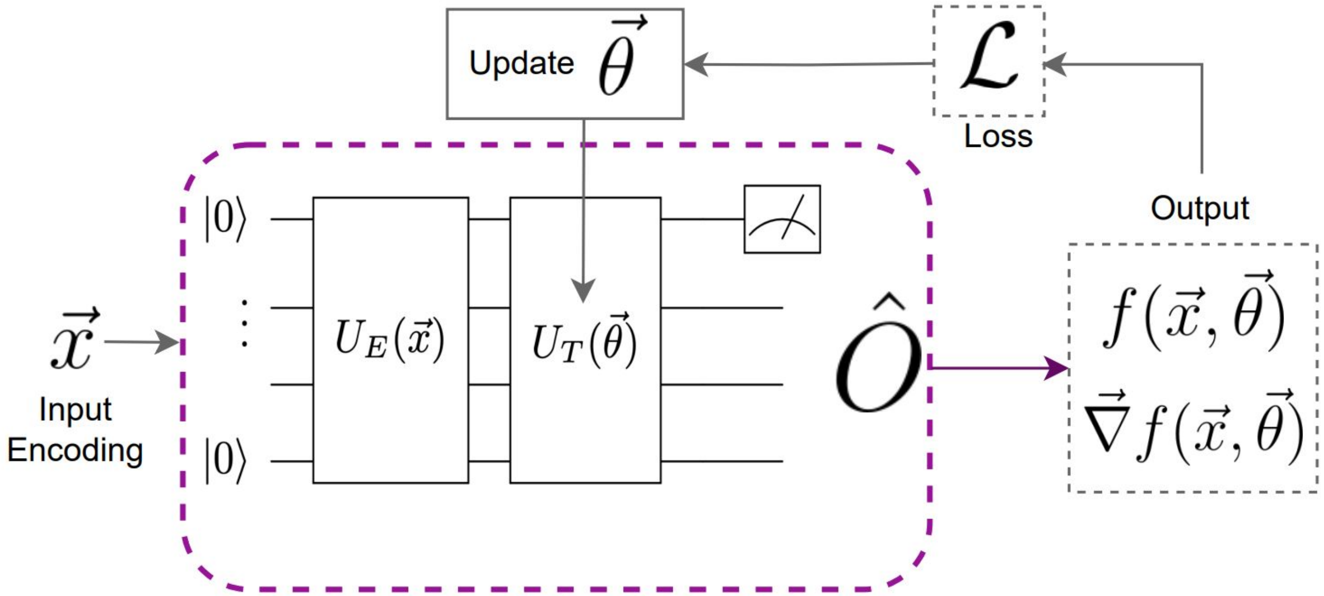
Simple (polynomial) quantum circuits can be **hard to simulate classically**
Ansatz expressibility: Quantum advantage?

McClean, J.R., Boixo, S., Smelyanskiy, V.N. *et al.* Nat Commun 9, 4812 (2018)

Peruzzo, A. *et al. Nat. Commun.* 5, 4213 (2014)

The Variational Model Loop

Training [Classical]

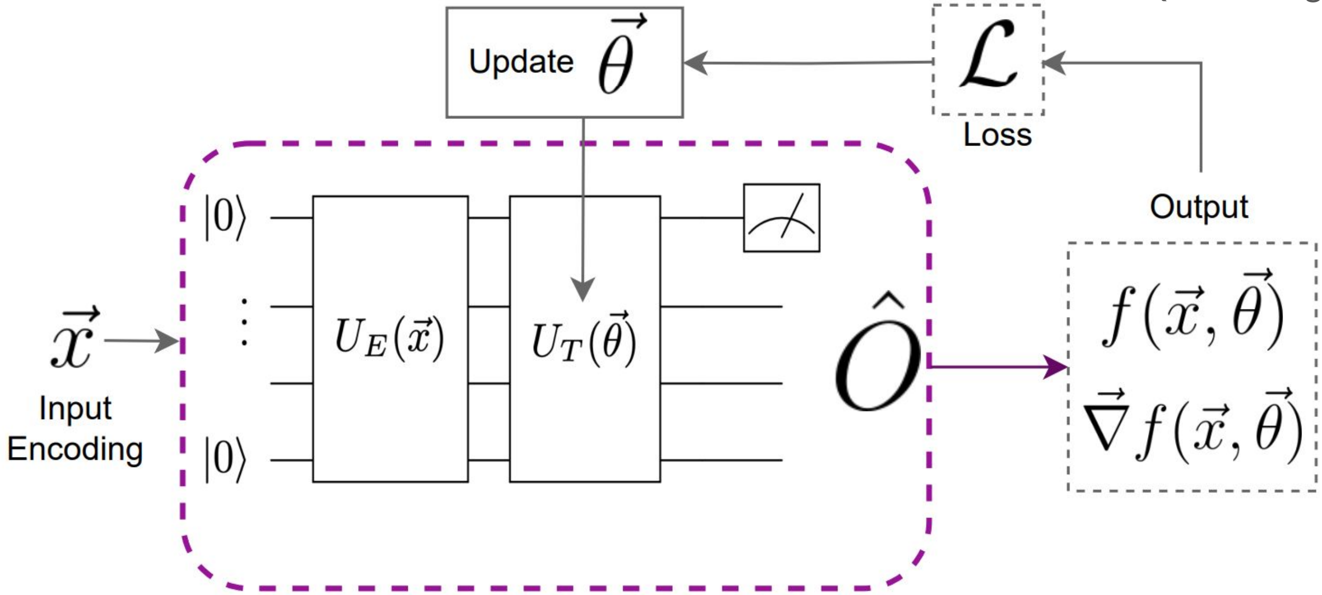


Parametrized Quantum Circuit (PQC) [Quantum]

The Variational Quantum Algorithm Loop

Training [Classical]

Loss (classical post-processing):
Quantifies goodness of model



f: Quantum Circuit output.
Train to solve problem

Gradients can be computed
reusing circuit

Parametrized Quantum Circuit (PQC) [Quantum]

More general (Quantum Machine Learning)

Quantify the problems

In practice, convergence is hard. How to characterize?

Metrics: Expressivity, trainability, accuracy, depth

Quantify the problems

In practice, convergence is hard. How to characterize?

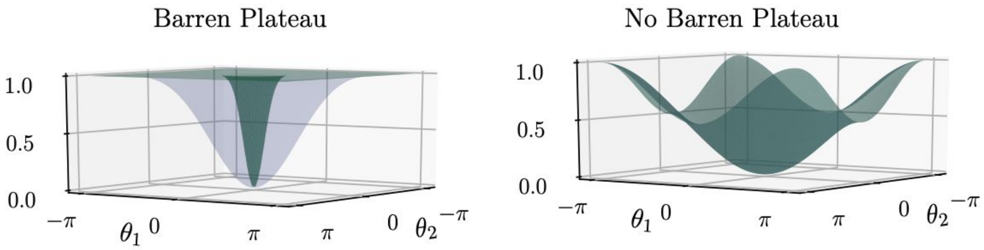
Metrics: Expressivity, trainability, accuracy, depth

Barren Plateaus

Exponential concentration of loss landscape (with increasing qubits)

Equivalent to concentration of

$$\mathrm{Var}_{\vec{\theta}}[\partial_\mu \mathcal{L}]$$



A. Arrasmith, Z. Holmes, M. Cerezo and P. Coles, Quantum Sci. Technol. 7, 045015 (2022)

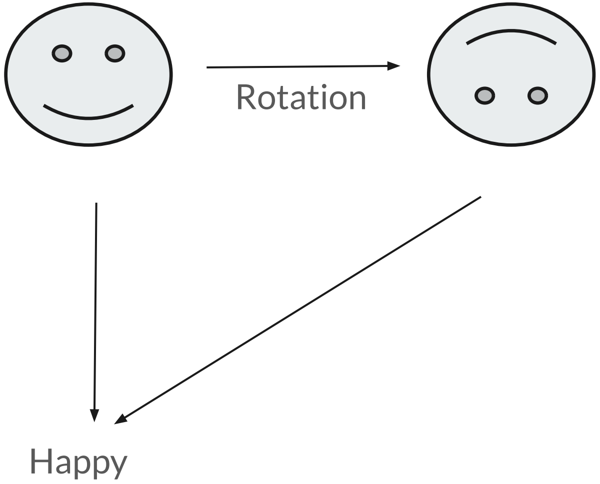
Larocca, M., Thanasilp, S., Wang, S. et al.. Nat Rev Phys 7, 174–189 (2025)

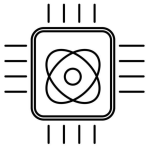
Common problem: Very expressive → Barren Plateau!

Invariance and Equivariance

Invariance: Property of our system and solution

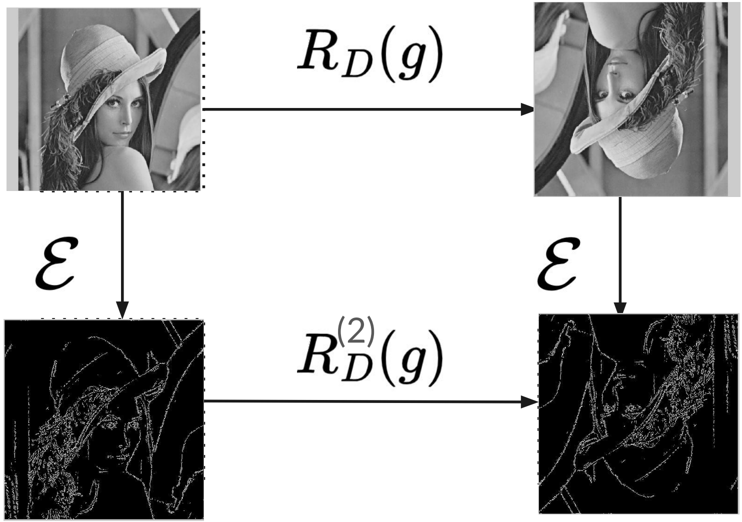
$F($☺$)$ Happy/Sad?





Equivariance:

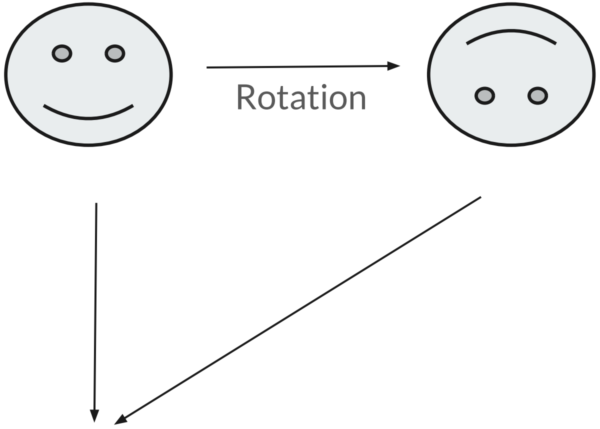
Ex. Layer of Convolutional Neural Network



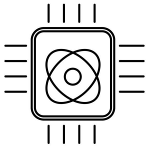
Invariance and Equivariance

Invariance: Property of our system and solution

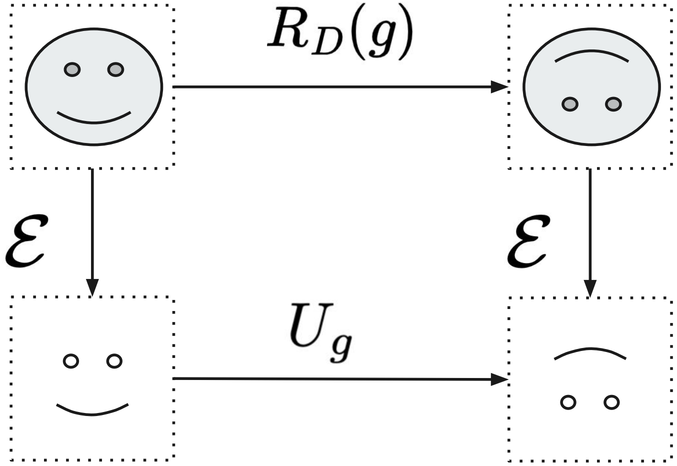
Happy/Sad?



Happy



Equivariance:



For simple quantum circuit:
(same unitary representation)

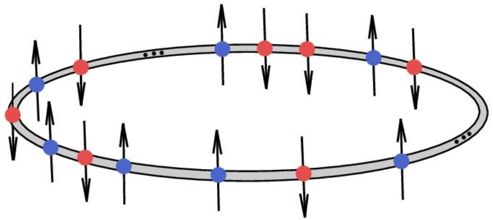
$$[R(g), U_{\boldsymbol{\theta}}] = [R(g), \hat{O}] = 0$$

Implementation: Problem

Ground state of Transverse Field Ising Model (TFIM)

$$H_{\text{TFIM}} = -J \sum_{i=1}^{N} Z_i Z_{i+1} - g \sum_{i=1}^{N} X_i$$

$$|\uparrow\rangle \rightarrow |0\rangle \,, |\downarrow\rangle \rightarrow |1\rangle$$

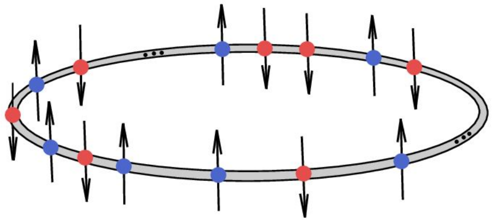


Implementation: Problem

Ground state of Transverse Field Ising Model (TFIM)

$$H_{\text{TFIM}} = -J\sum_{i=1}^{N} Z_i Z_{i+1} - g\sum_{i=1}^{N} X_i$$

$$|\uparrow\rangle \rightarrow |0\rangle, |\downarrow\rangle \rightarrow |1\rangle$$



<u>Symmetries</u>

Translation Symmetry (Periodic conditions)

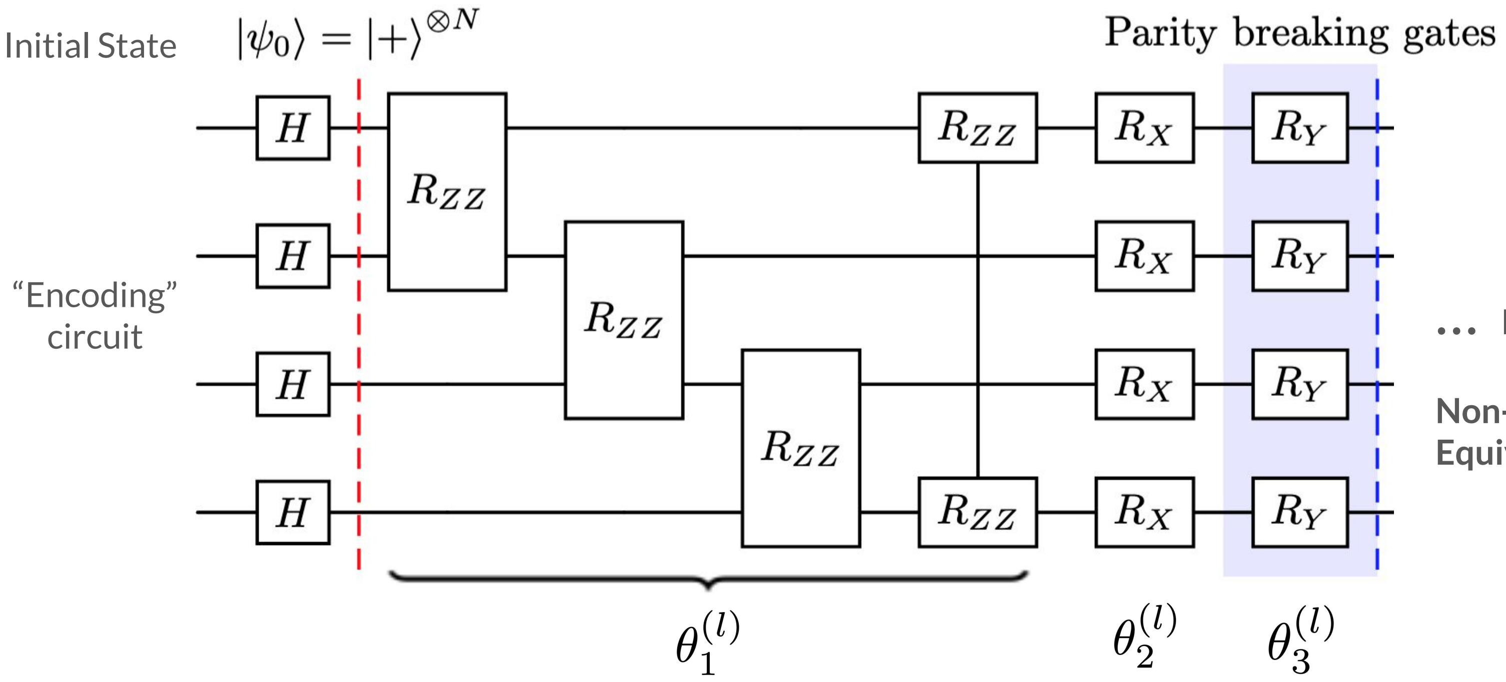
Parity Symmetry: $E(|\uparrow\uparrow\downarrow\uparrow\ldots\rangle) = E(|\downarrow\downarrow\uparrow\downarrow\ldots\rangle)$

Representations

$$\hat{T} = \text{SWAP}_{n-1,n}\text{SWAP}_{n-2,n-1}\ldots\text{SWAP}_{2,3}\text{SWAP}_{1,2}$$

$$\hat{P} = \prod_{i=1}^{N} X_i$$

Implementation: Ansatz

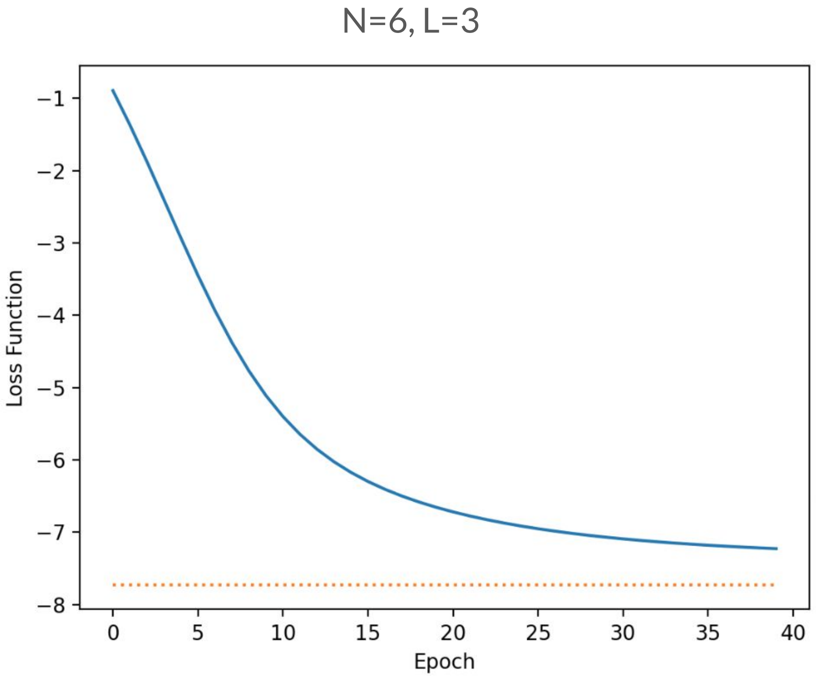


Non-Equivariant: As shown
Equivariant: Without Ry

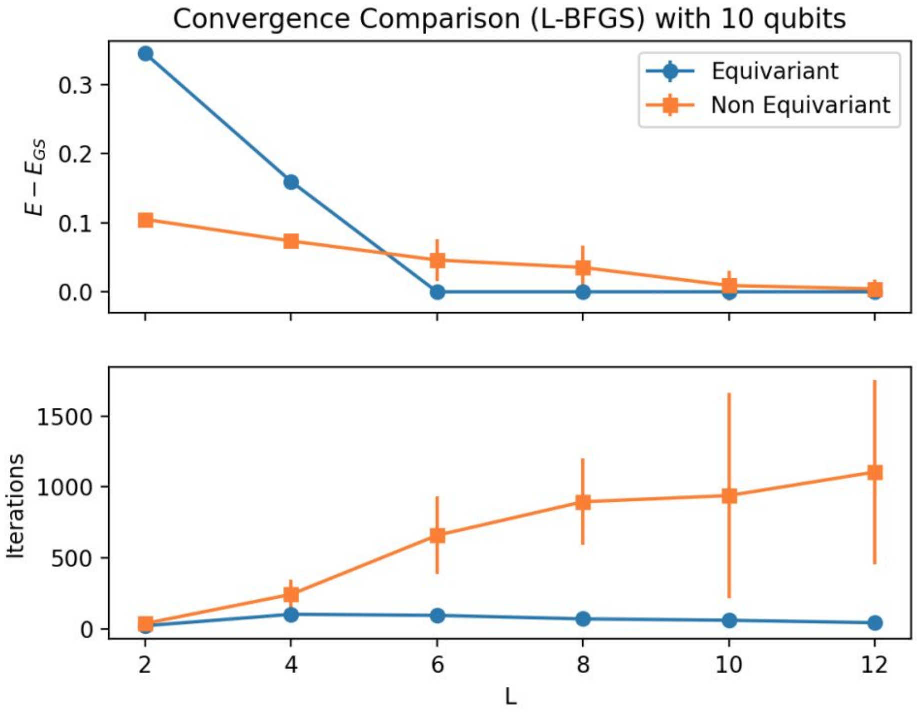
Same parameter for all qubits (translational equivariance)

$$R_Y(\theta) = e^{-i\frac{\theta}{2}Y}, R_X(\theta) = e^{-i\frac{\theta}{2}X}, R_{ZZ}(\theta) = e^{-i\frac{\theta}{2}ZZ}$$

Better convergence



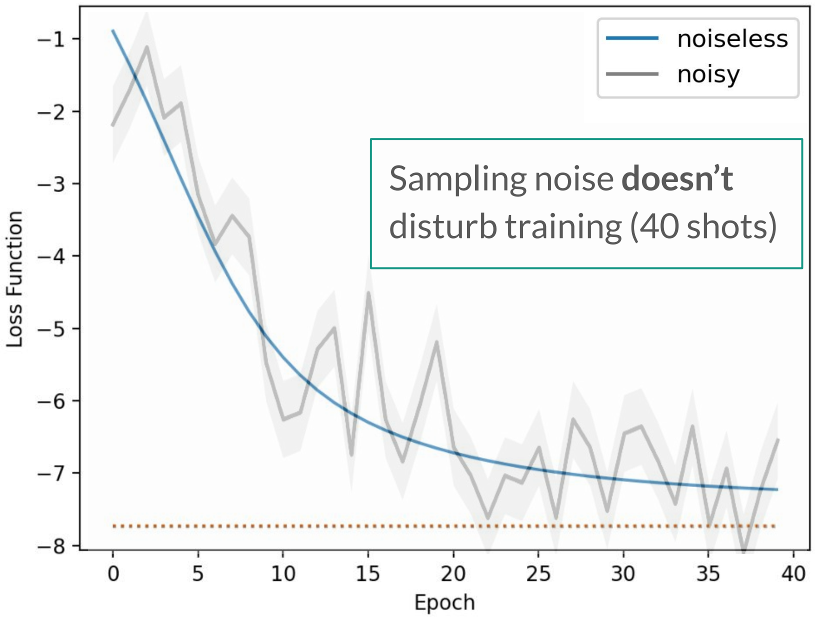
Example of training Equivariant model using **Gradient Descent**



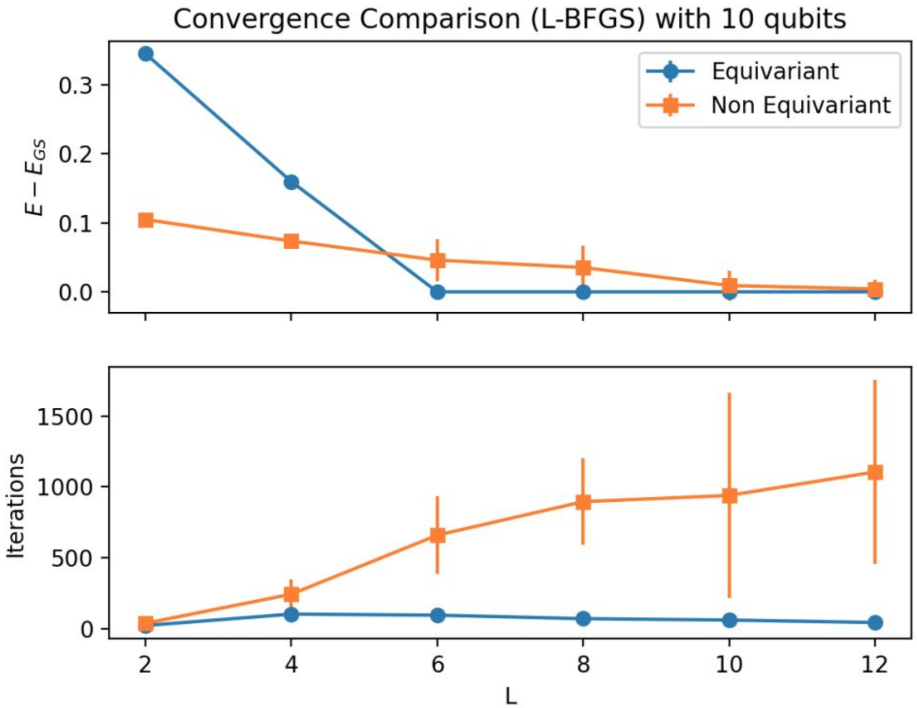
Comparison of equivariant and non-equivariant models, trained with **L-BFGS**

Better convergence

N=6, L=3

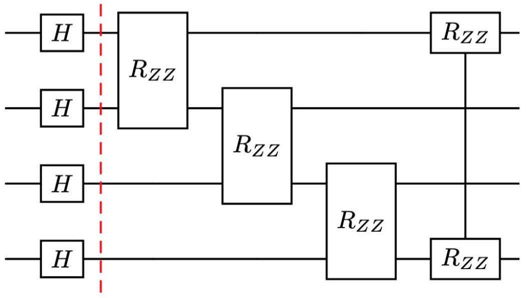


Example of training Equivariant model using **Gradient Descent**



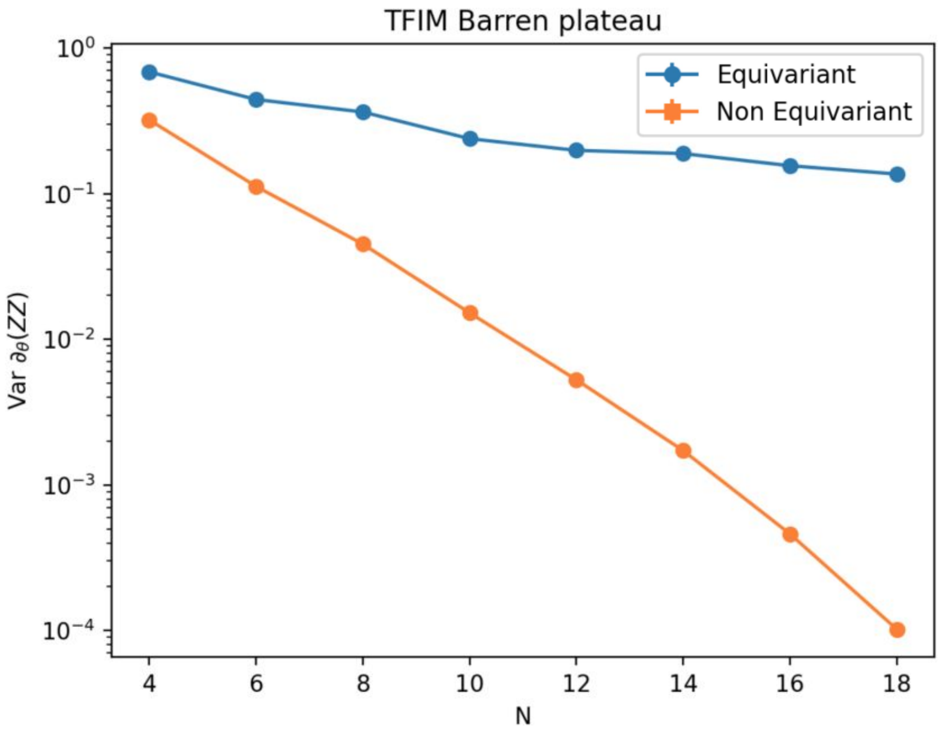
Comparison of equivariant and non-equivariant models, trained with **L-BFGS**

Results: Barren Plateaus



Shown: variance of ZZ measurement

$$H_{\text{TFIM}} = -J \sum_{i=1}^{N} Z_i Z_{i+1} - g \sum_{i=1}^{N} X_i$$

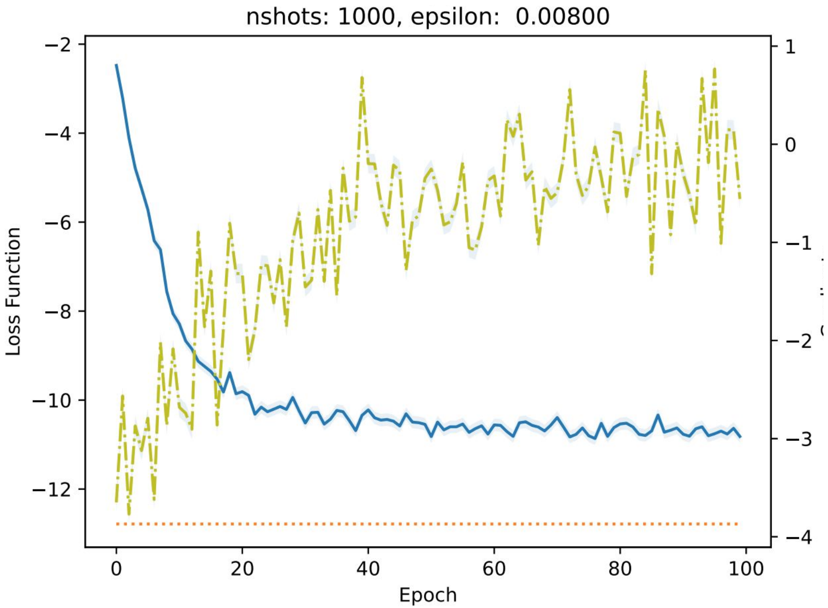


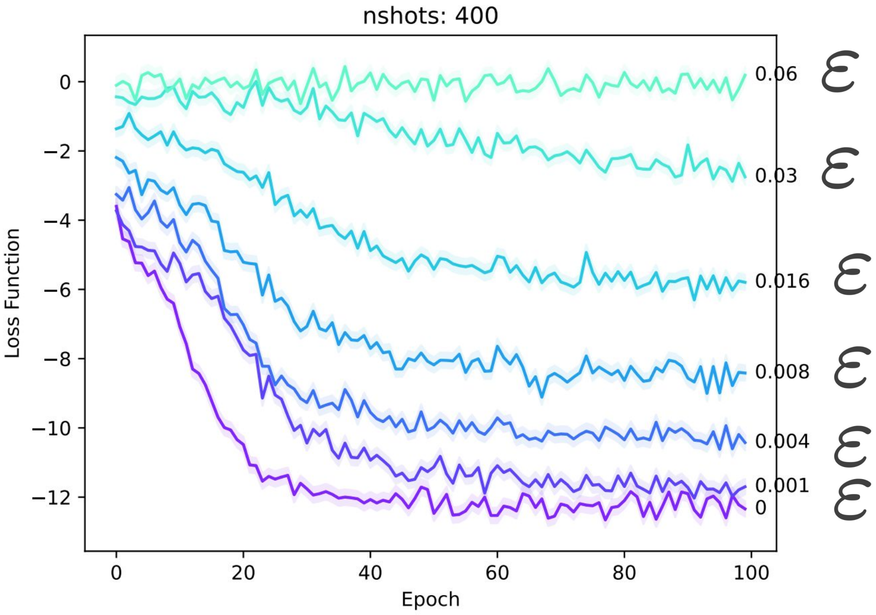
Barren Plateaus:
- non-equivariant model **yes!**
- equivariant model **no!** → (In fact "Dynamical Lie Algebra" N^3 scaling)

Results: Noise Model

Random X and Z noise, probability ε

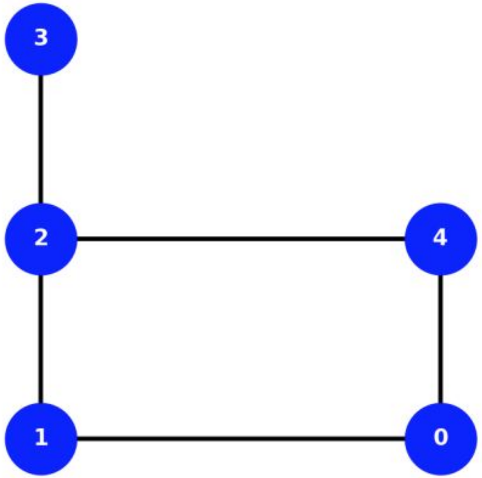
N=10, L=2





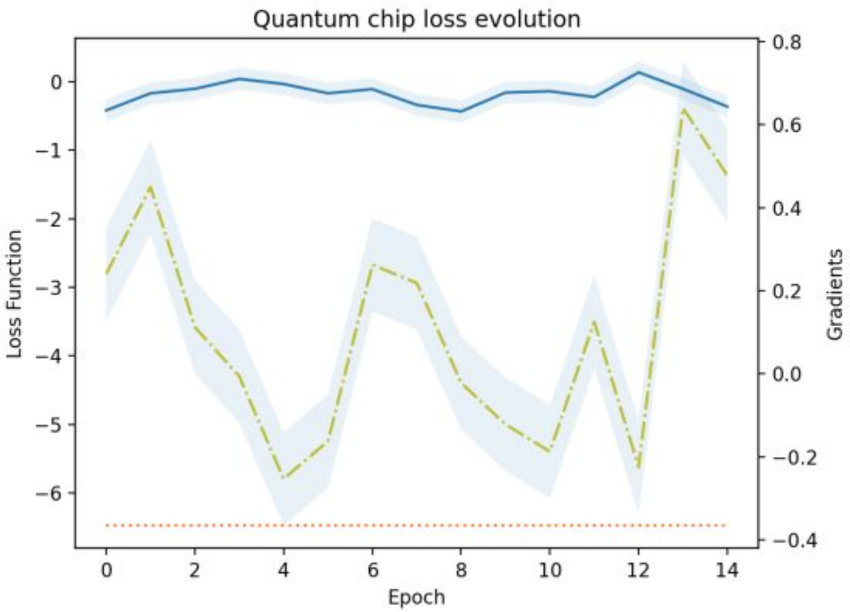
Noisy barren plateaus

Results: BSC chip



Old Chip topology
(New one is worse!)

N=5, L=2

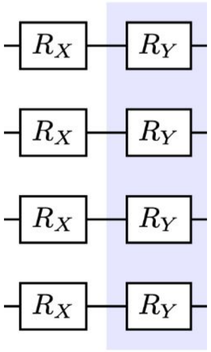


Noise dominated

Conclusions and Outlook

Exploit symmetries through equivariance, for example TFI model

- Better behaviour than non equivariant model, absence of Barren Plateaus
- Trainability limited by noise



Outlook:

- Interplay between more representations
- Links absence of barren plateaus with classical simulabilty! (Bad)
- Smart initialisation strategies
- Symmetry breaking

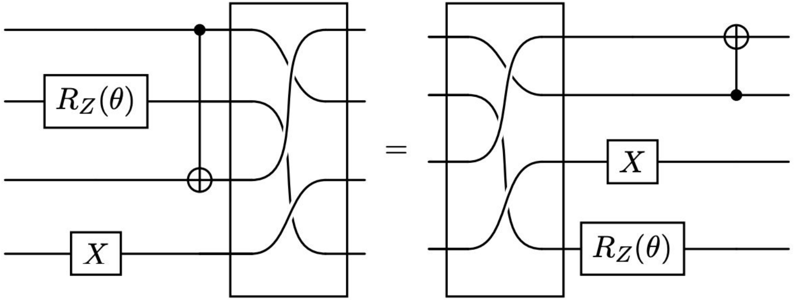
Implementation: Twirling

$$\mathcal{T}_G(A) = \frac{1}{|G|}\sum_{g\in G} R(g) A R(g)^\dagger$$

Implementation: Twirling

$$\mathcal{T}_G(A) = \frac{1}{|G|} \sum_{g \in G} R(g) A R(g)^\dagger$$

Example for translation

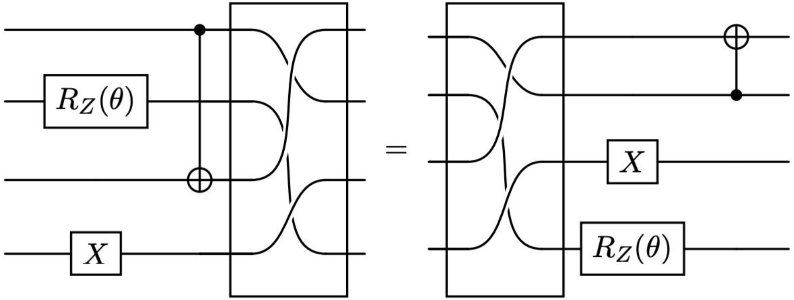


$$X \to \sum_{i=1}^{N} X_i, \quad X \to \sum_{i=1}^{N} Z_i Z_{i+1}$$

Implementation: Twirling

$$\mathcal{T}_G(A) = \frac{1}{|G|} \sum_{g \in G} R(g) A R(g)^\dagger$$

Example for translation



$$X \to \sum_{i=1}^{N} X_i, \quad X \to \sum_{i=1}^{N} Z_i Z_{i+1}$$

$$[R(g), A] = 0 \longrightarrow [R(g), e^{iA}] = 0$$

Introduction: Invariance and Equivariance

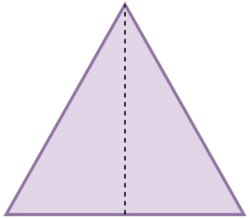
Goal: Exploit structure, symmetries

Symmetries → Group Theory

Symmetry groups identified **abstractly**

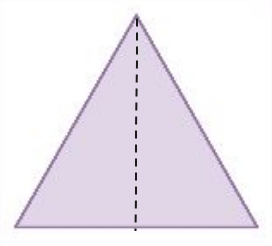
$g \circ g' \in G$, here D_3

Example



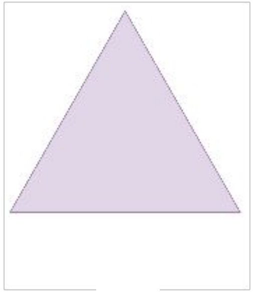
1

Reflection



a

Rotation



r

Introduction: Invariance and Equivariance

Goal: Exploit structure, symmetries

Symmetries → Group Theory

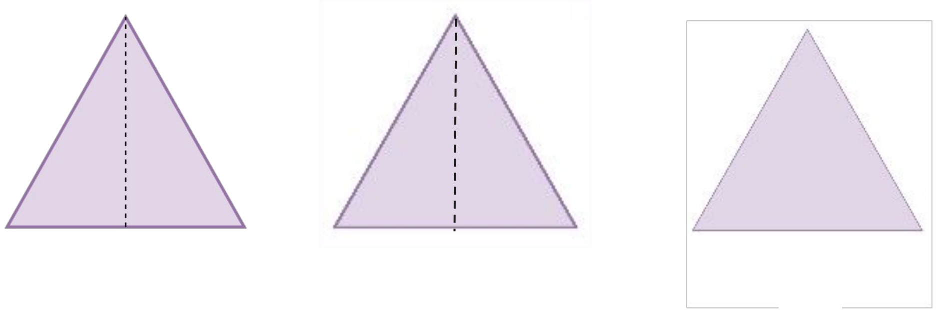
Symmetry groups identified **abstractly**

$g \circ g' \in G$, here D_3

Example

Reflection

Rotation



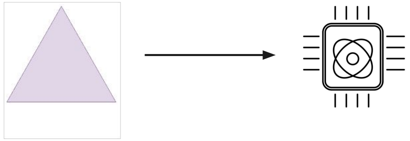
1 $\quad$ a $\quad$ r

Representation $R\colon G \to \mathrm{Aut}(V)$ $\qquad R(g_1)R(g_2) = R(g_1 \circ g_2)$

In 2d data space

$$R_D(r) = \begin{pmatrix} \cos 60° & \sin 60° \\ -\sin 60° & \cos 60° \end{pmatrix}$$

Unitary representation
- Dependent on Encoding!
- For quantum states: Wigner Theorem



$g \to U_g$

Extra: Parameter Shift Rule

PSR:

$$\partial_{\theta_k} C(\boldsymbol{\theta}) = r\left(C(\boldsymbol{\theta} + \frac{\pi}{4r}\boldsymbol{e}_k) - C(\boldsymbol{\theta} - \frac{\pi}{4r}\boldsymbol{e}_k)\right)$$

Derivation:

$$C(\theta) = \langle\psi|U(\theta)^\dagger \hat{O} U(\theta)\,|\psi\rangle$$

with

$$G^2 = r^2 I$$

$$e^{i\theta G} = \cos(r\theta) + i\sin(r\theta)\frac{G}{r}$$

Plugging this in

$$C(\theta) = \langle\psi|U(\theta)^\dagger \hat{O} U(\theta)\,|\psi\rangle =$$

$$= \cos^2(r\theta)\langle\psi|\hat{O}|\psi\rangle + \sin^2(r\theta)\langle\psi|\frac{G}{r}\hat{O}\frac{G}{r}|\psi\rangle + 2\sin(r\theta)\cos(r\theta)\mathfrak{Re}(-i\langle\psi|\frac{G}{r}\hat{O}|\psi\rangle) =$$

$$= a + b\sin 2r\theta + c\cos 2r\theta$$

Extra: Dynamical Lie Algebra

If we build our circuits from some generator set

$$U(\boldsymbol{\theta}) = \prod_{l=1}^{L} e^{-i\theta_l \hat{G}_l}$$

$$e^A e^B = e^{A+B+\frac{1}{2}[A,B],+\frac{1}{12}[X,[X,Y]]+\dots}$$

We can bound accessible gates from generators by computing all nested commutators: **Dynamical Lie Algebra (DLA)**

$$\mathfrak{g} = \langle \{ i\hat{G}_l \}_l \rangle_{\text{Lie}} \subseteq \mathfrak{su}(2^n)$$

Extra: Dynamical Lie Algebra

$$\mathfrak{g} = \langle \{iG_l\}_l \rangle_{\text{Lie}} \subseteq \mathfrak{su}(2^n)$$

Expressivity can be understood Information theoretically:

- A (noiseless) channel has maximum capacity → all **outputs are equally likely**

But DLA exponential scaling → Average Inner product exponentially decaying

$$\text{Tr}[\mathcal{U}_{\boldsymbol{\theta}}(\rho) O \ ]$$

Double-edged sword: Can exploit DLA to simulate classically