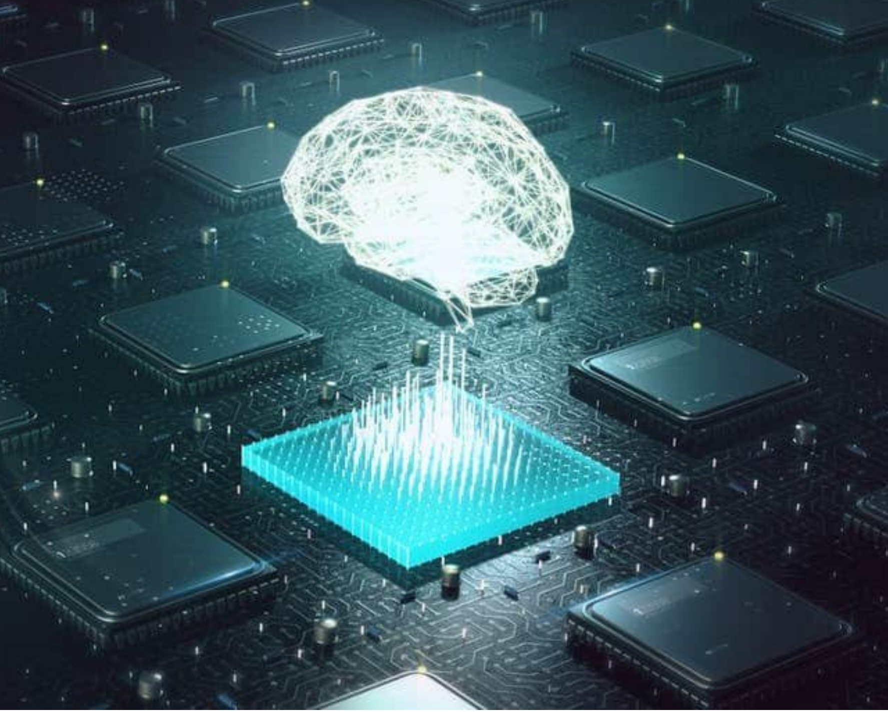
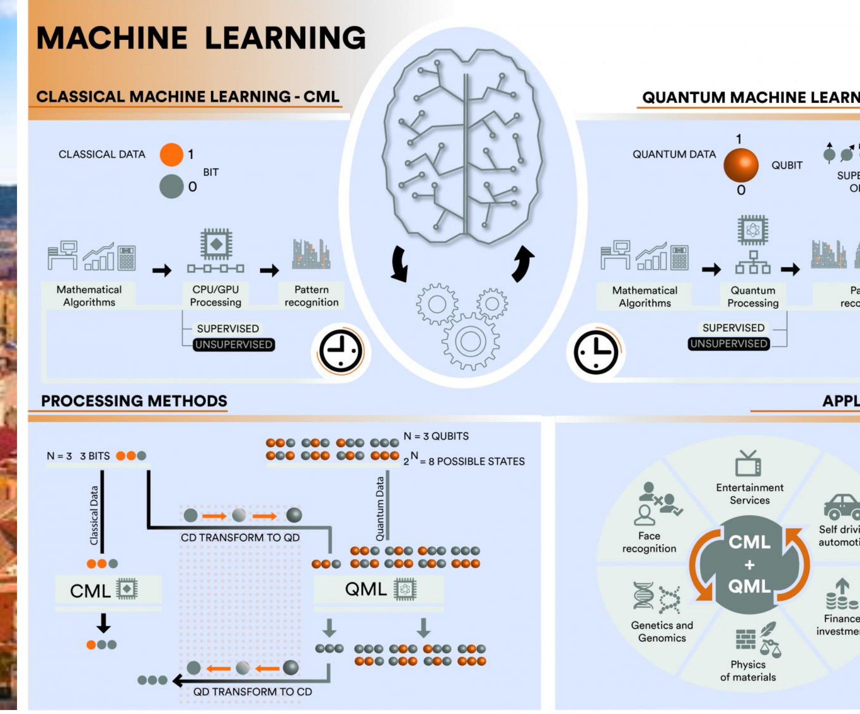
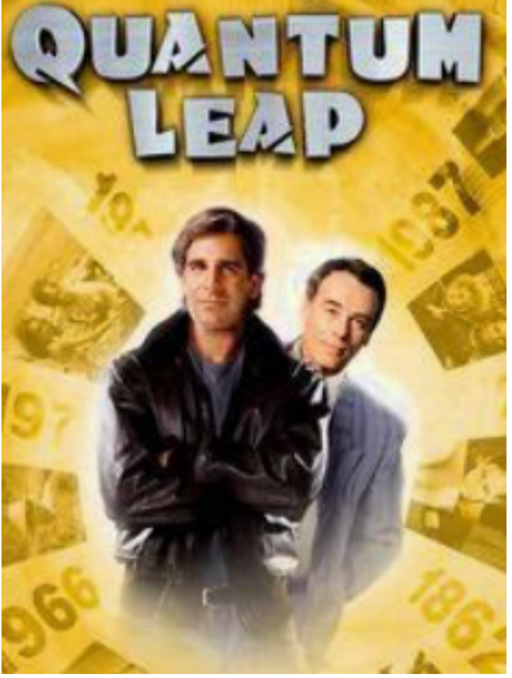




Quantum Machine Learning for Fundamental Physics

Michael Spannowsky

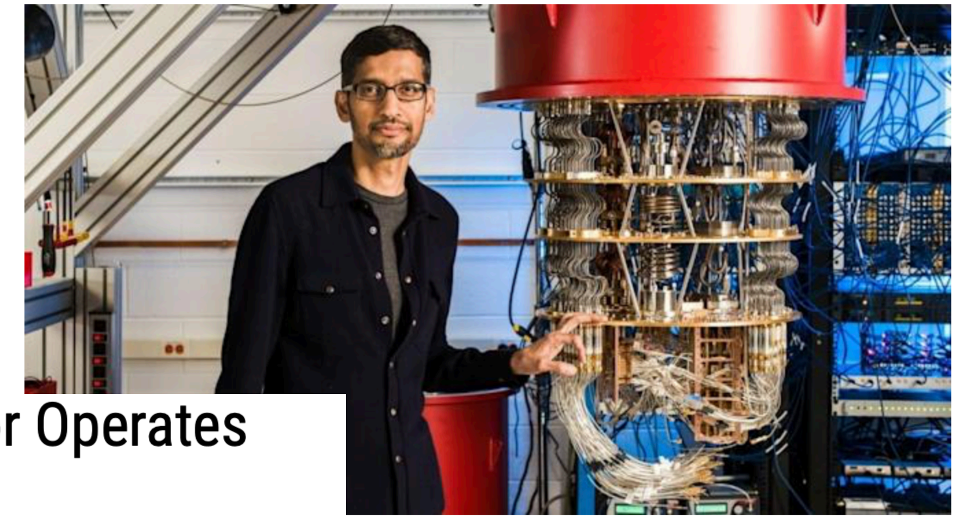
IPPP, Durham University



The Morning After: Google claims 'quantum supremacy'

And a controversial 'Ghost in the Shell' trailer.

R. Lawler
@Rjcc
October 24th, 2019



Master in Elektrotechnik, Informatik, Robotik, Maschinenwesen o. ä. (w/m/d)
German Aerospace Center (DLR) · Oberpfaffenhofen, Bavaria, Germany (On-site)
4 company alumni

Professor Cyber Security im Online Fernstudium (m/w/d)
IU International University of Applied Sciences · Germany (Remote)
Actively recruiting

Expertin für Post-Quanten-Kryptographie (w/m/d)
Deutsche Bahn · Frankfurt, Hesse, Germany (On-site)
Actively recruiting

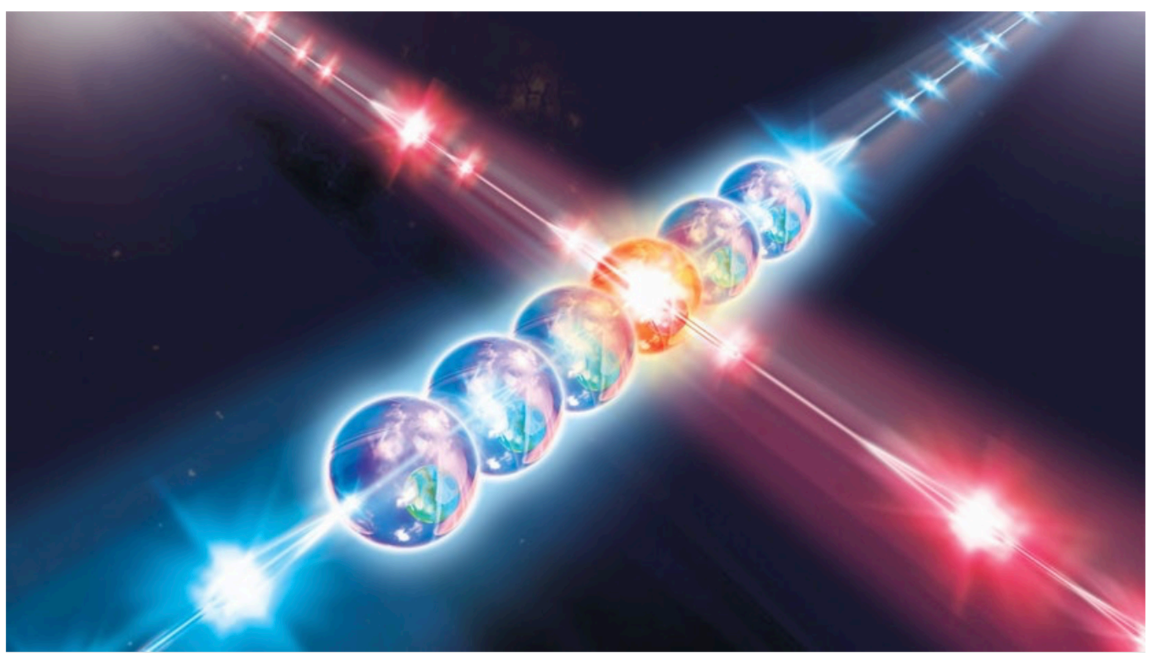
Master Thesis: Design of digitally enhanced power management circuits for Future Quantum Computers
Forschungszentrum Jülich · Jülich, North Rhine-Westphalia, Germany (On-site)
1 company alum

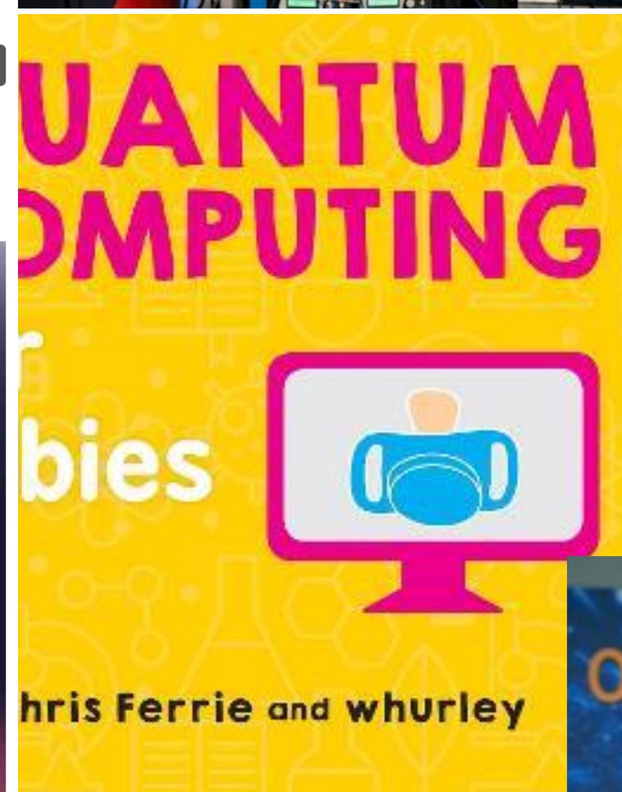
Expertin für Quantenkommunikation (w/m/d)
Deutsche Bahn · Frankfurt, Hesse, Germany (On-site)
Actively recruiting

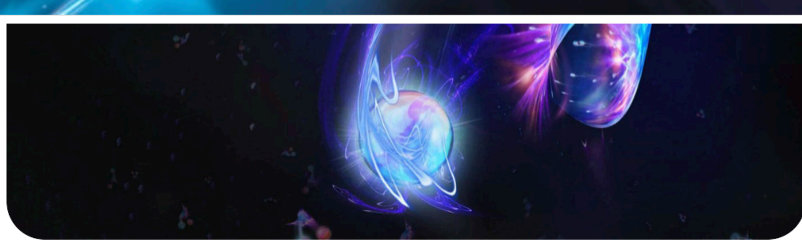
First Quantum Computer Simulator Operates The Speed Of Light

Share

Kristen Philipkoski
Published 10 years ago: September 2, 2011 at 7:02 am - Filed to: COMPUTING







uantum Computers Will Be Incredibly Useful For

Computers don't exist in a vacuum. They serve to solve problems, and the type of problems they can solve are influenced by their hardware. Graphics processors are specialized for rendering images; artificial intelligence processors for AI; and quantum computers designed for... what? While the power of quantum computing is impressive, it does not mean that existing ...

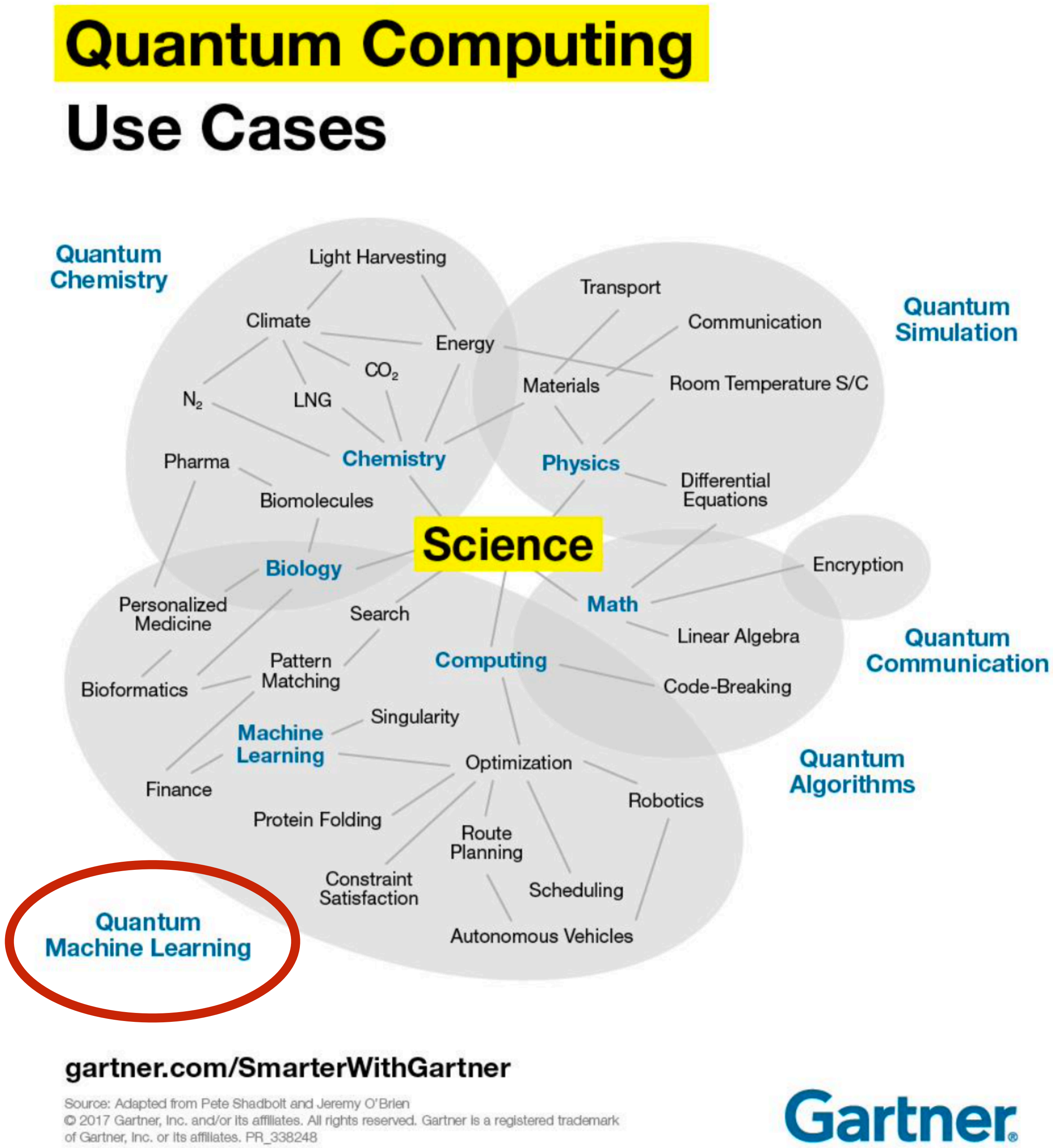
Futuristic News
208 followers

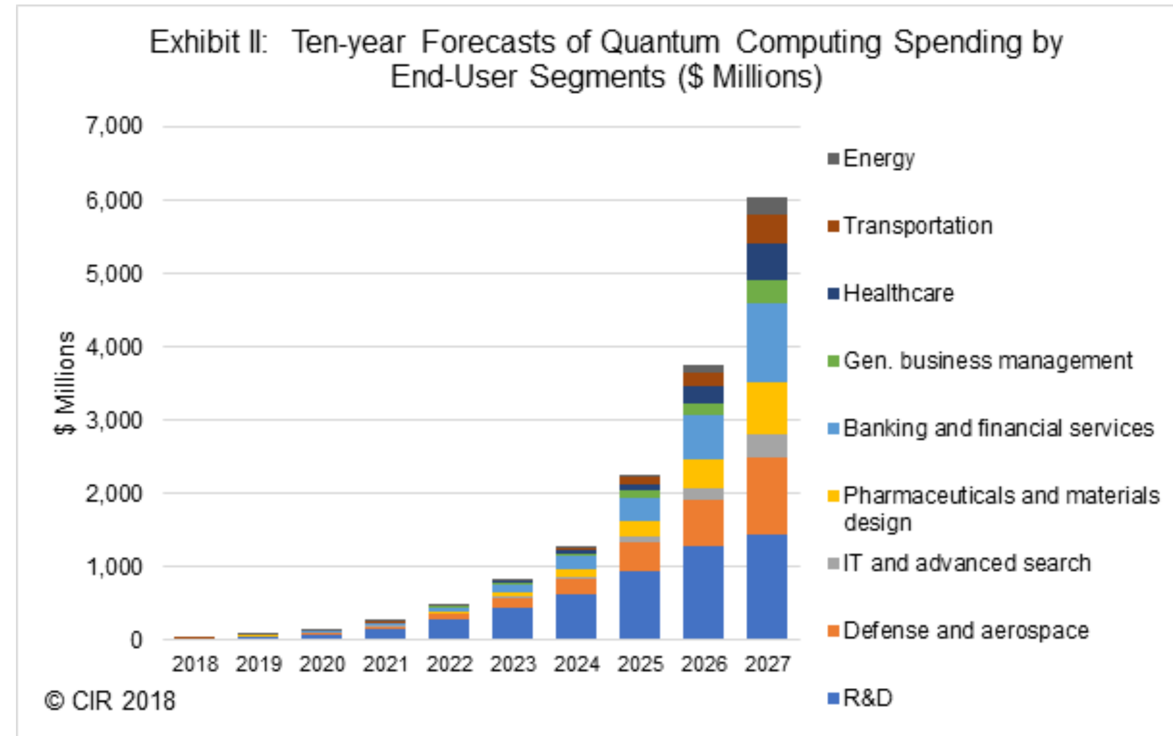
"Nature is quantum [...]
so if you want to
simulate it, you need a
quantum computer"
– Richard Feynman
(1982)

Easily said … so how do we do that?

Beginning of a scientific journey that accelerated in recent years tremendously….

Private and Public Sector is placing big bets on Quantum Computing

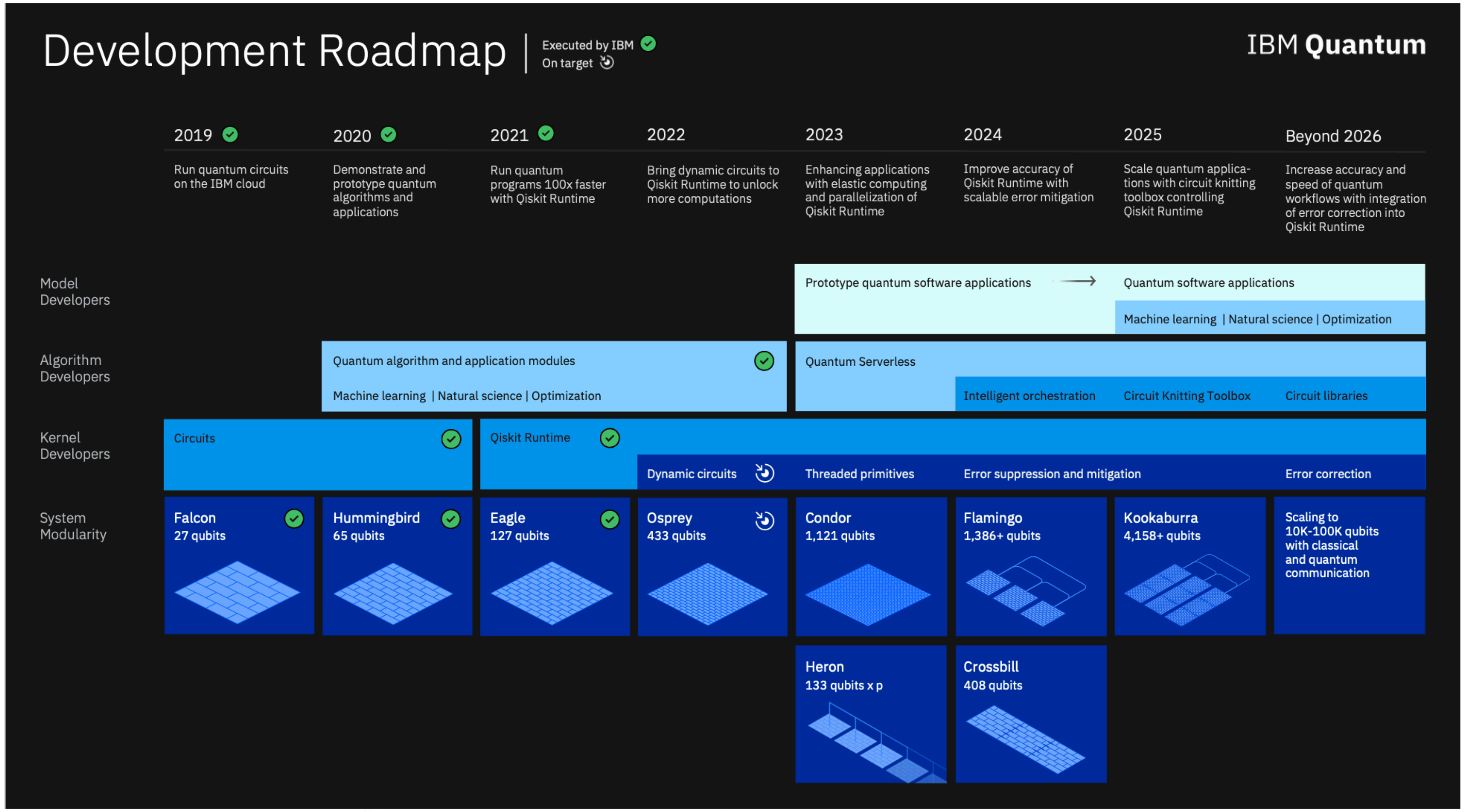




Significant financial investment expected across many sectors

In US, already now higher financial investment from private than public sector

All national and international labs have QC programmes (Fermilab, BNL, LBNL, DESY, CERN, Singapur, Abu Dhabi, …)

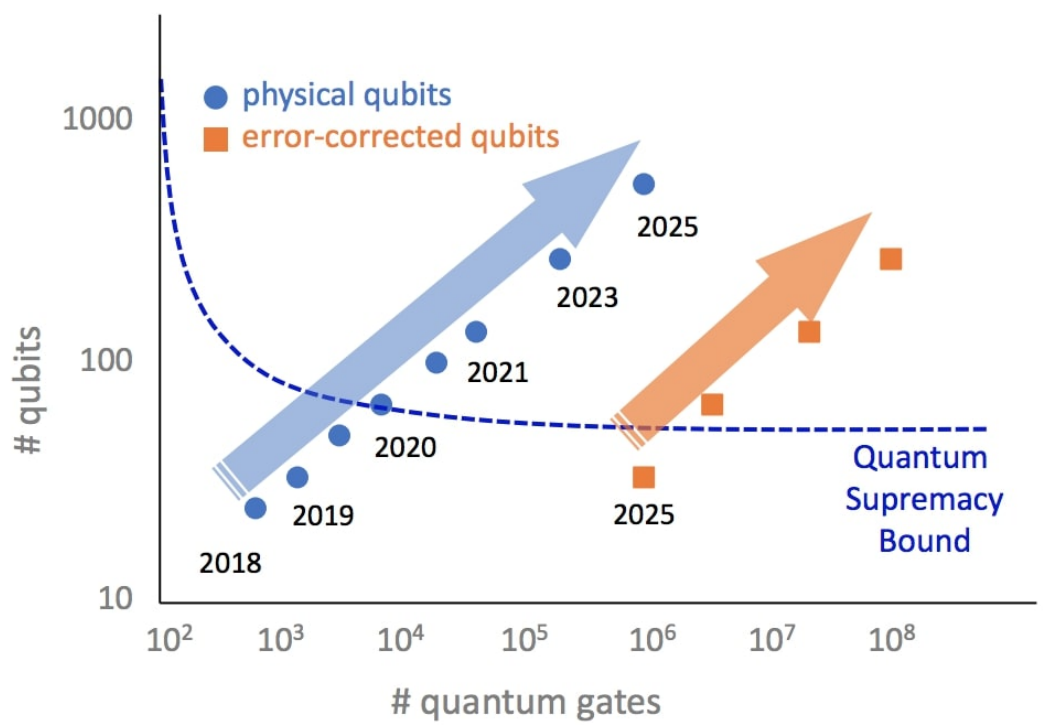


Future Circular Collider

1 mio physical qubits
->
1k logical qubits

by 2029

IonQ Roadmap into the Future



Design Principle

$$Qubits = \sqrt{Gates}$$

timescales much smaller than large-scale science projects

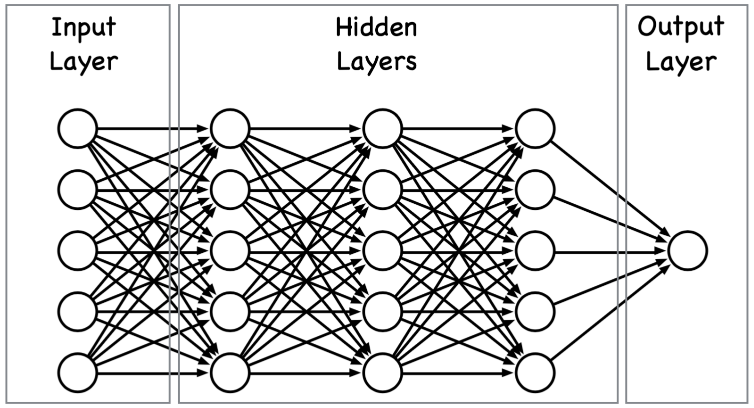
Algorithm Type

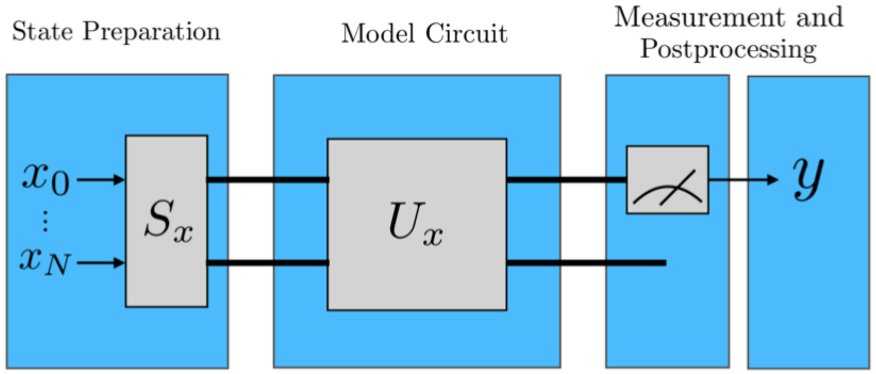
Data Type	Classical	Quantum
Classical	CC	CQ
Quantum	QC	QQ

Classical
ML Algorithms

Quantum
ML Algorithms

1. an adaptable complex system that allows approximating a complicated function





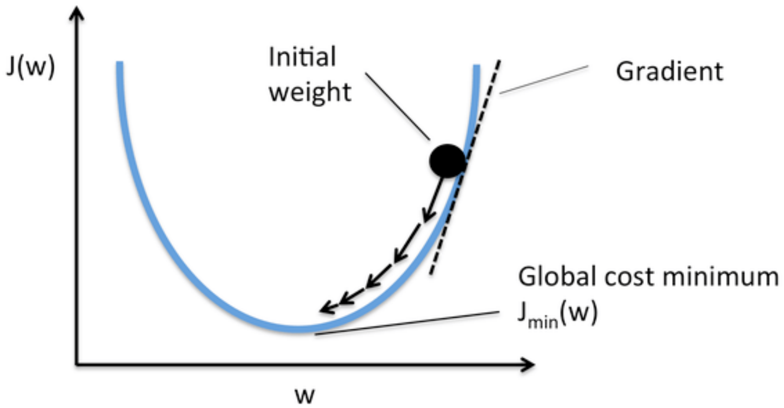
2. the calculation of a loss function used to define the task the method

$$E(y, y') = \frac{1}{2}|y - y'|^2$$

ground state

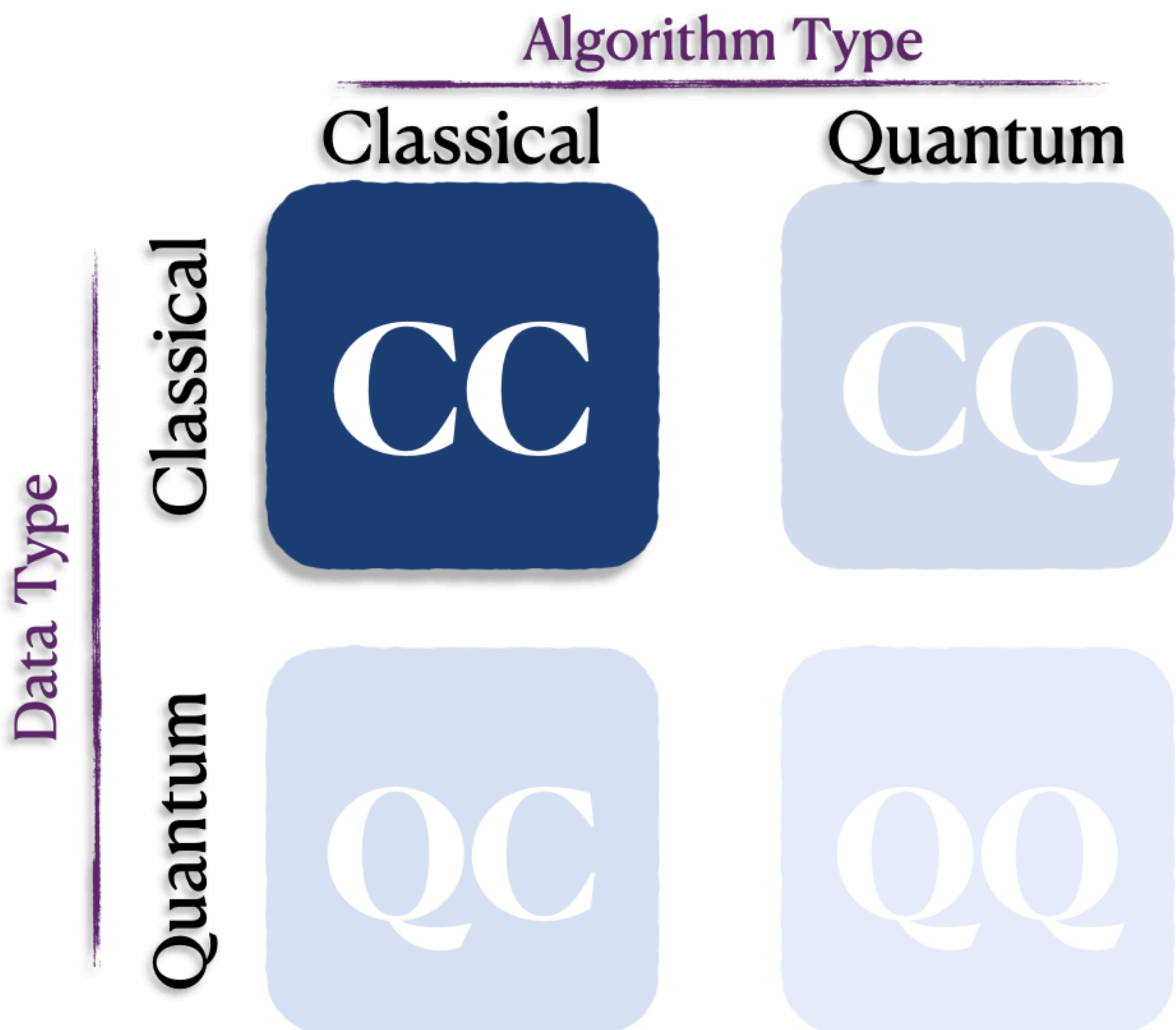
$$|\Gamma\rangle := \arg\min_{|\psi\rangle \in \mathcal{D}} \frac{\langle\psi|H|\psi\rangle}{\langle\psi|\psi\rangle}$$

3. a way to update 1. while minimising the loss function

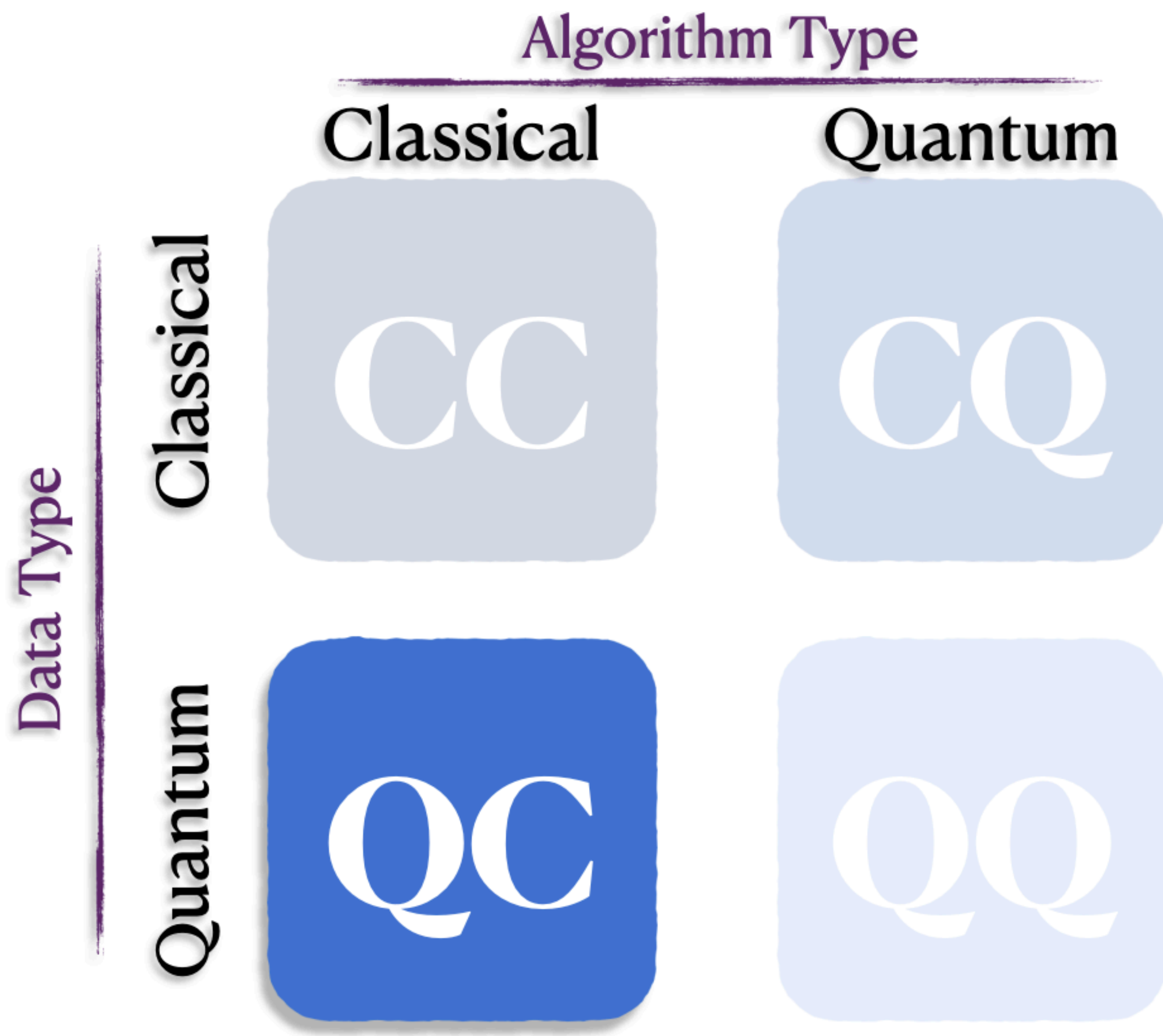


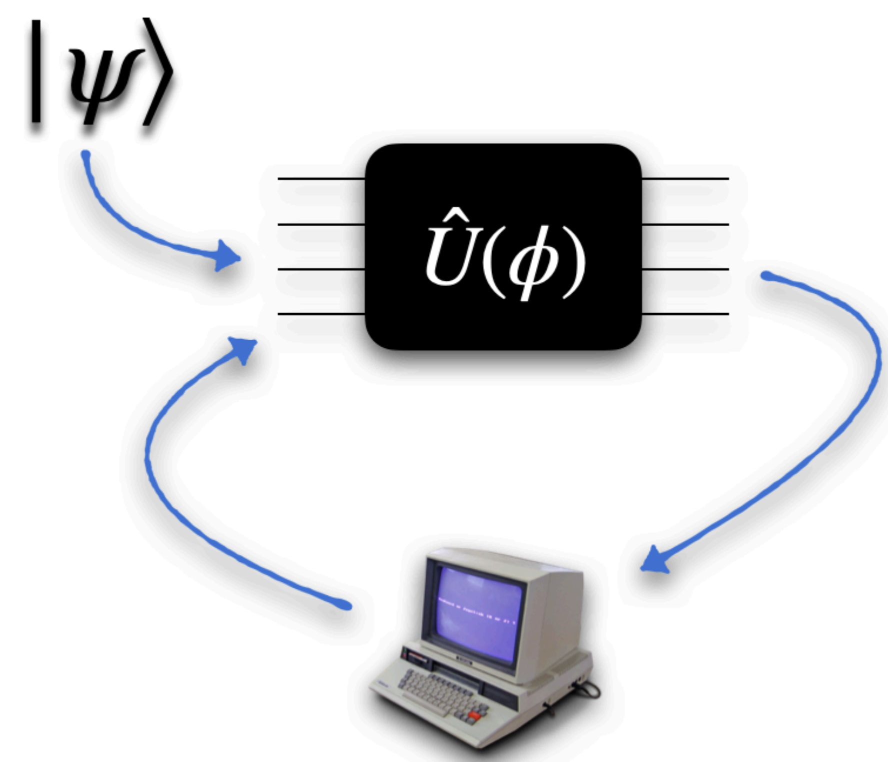
quantum: annealing

hybrid: classical opti.



Almost all of the methods used and shown at EUCAIFCON

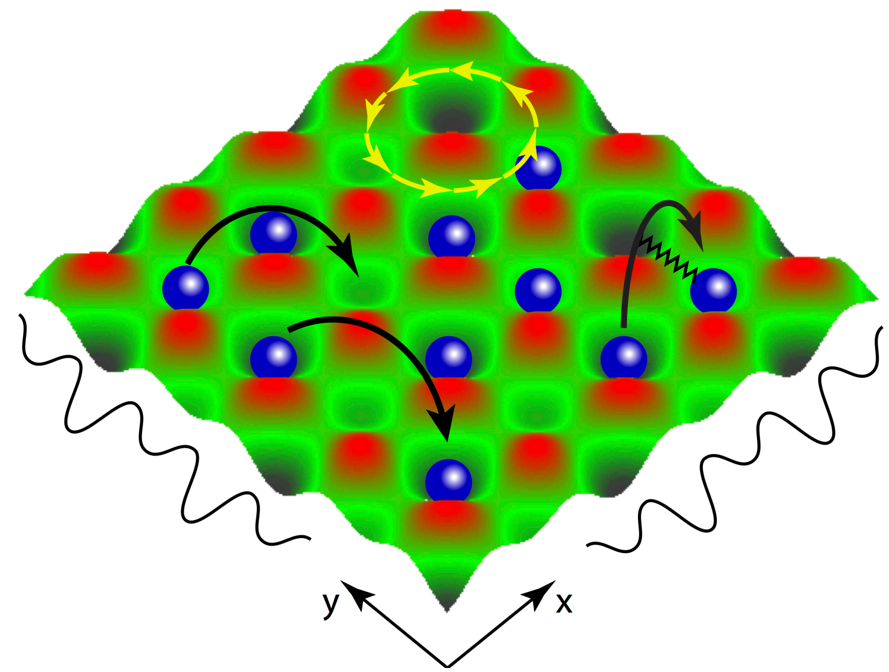




- Also Tensor Networks

Calculating observables in many-body systems with Neural Quantum States

- Challenges in solving quantum many-body problems
  - ➡ Exponential growth of Hilber space with system size
  - ➡ Computational intractability of classical methods
- The promise of machine learning
  - ➡ Efficient representation of complex systems
  - ➡ Advances in optimization and scalability
- Neural Quantum States as a bridge [Carleo, Troyer '17]
  - ➡ Combining neural networks with variational quantum methods
  - ➡ Alternative approach to Tensor Networks



Universal Approximation Theorem(s)

width $\rightarrow \infty \quad \Rightarrow |f^* - f_W| \sim \mathrm{width}^{-1}$

[Cybenk '89] [Leshno et al '93] [Hornik '91]

Univ. Approx. Conjecture(s)

$D \rightarrow \infty \quad \Rightarrow \quad |f^* - f_W| \sim \exp[-D]$

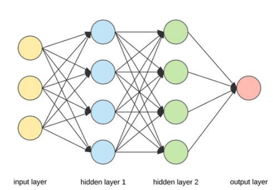
[Z. Lu et al, '17]

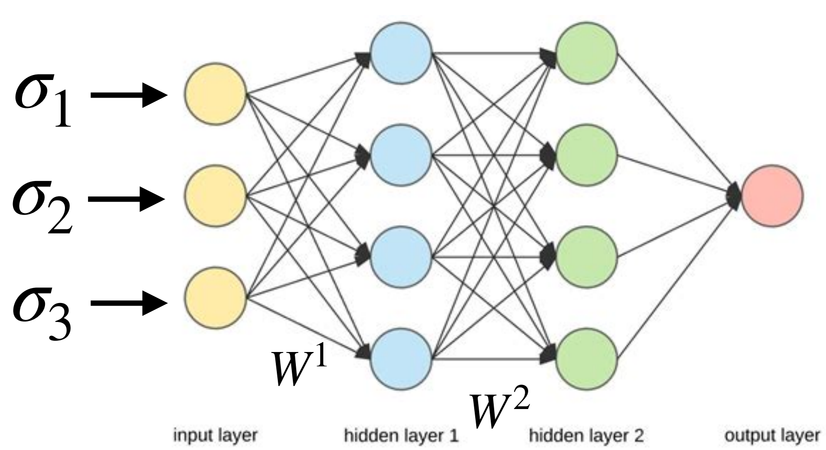
- To calculate the groundstate of a quantum system, choose a variational ansatz that approximates the quantum state of the system $|\psi\rangle \approx |\psi_\theta\rangle$ and minimize the energy computed with respect to the parameters θ.

- Rayleigh-Ritz variational principle guarantees convergence to approximation of groundstate

$$E_\theta = \frac{\langle \psi_\theta | H | \psi_\theta \rangle}{\langle \psi_\theta | \psi_\theta \rangle} \geq E_{\mathrm{GS}}$$

- In NQS, we approximate the wave function with a neural network

$$|\psi_W\rangle = \sum_{\sigma_1,\ldots,\sigma_N}$$

$$(\sigma_1, \ldots, \sigma_N)\, |\sigma_1, \ldots, \sigma_N\rangle$$



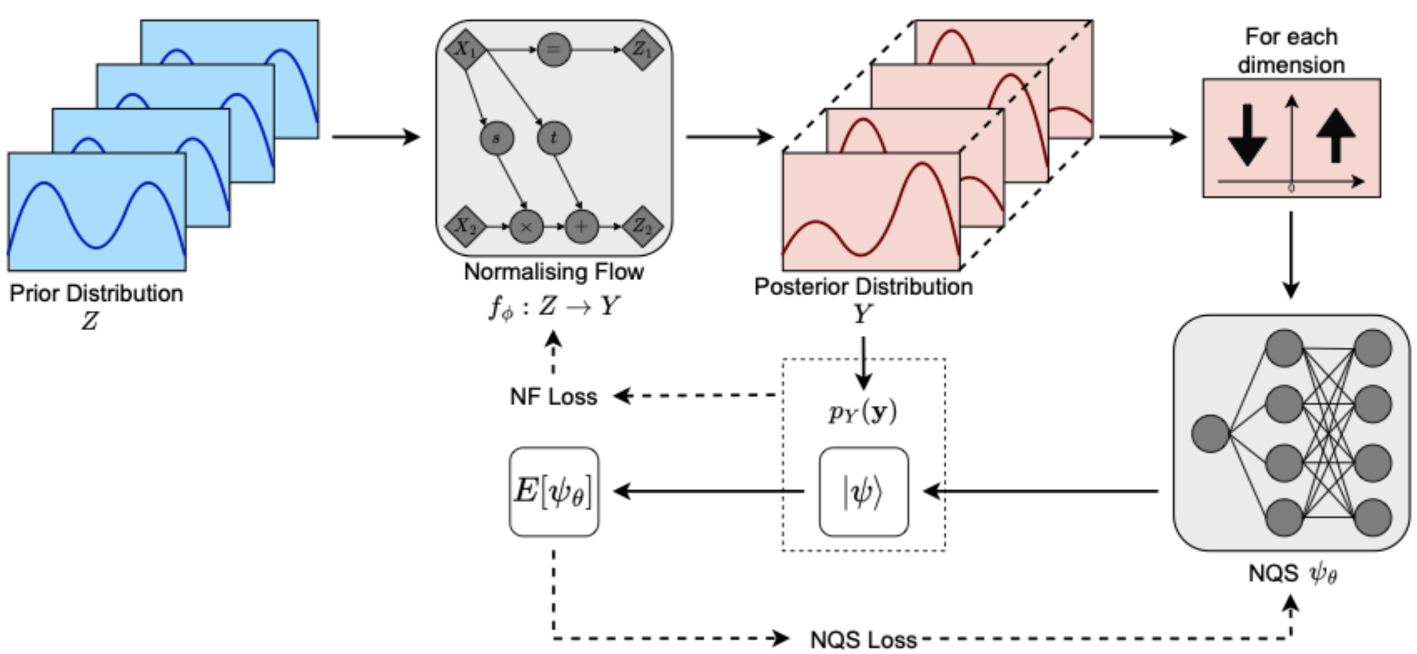
$$\psi(\sigma_1, \sigma_2, \ldots) \in \mathbb{C}$$

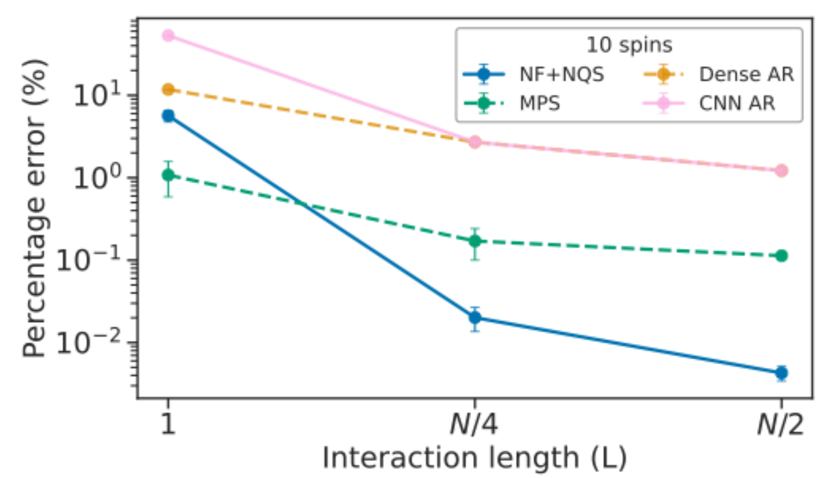
- Problem is that Hilbert space scales exponentially and selecting the entire basis for ψ becomes prohibitive.

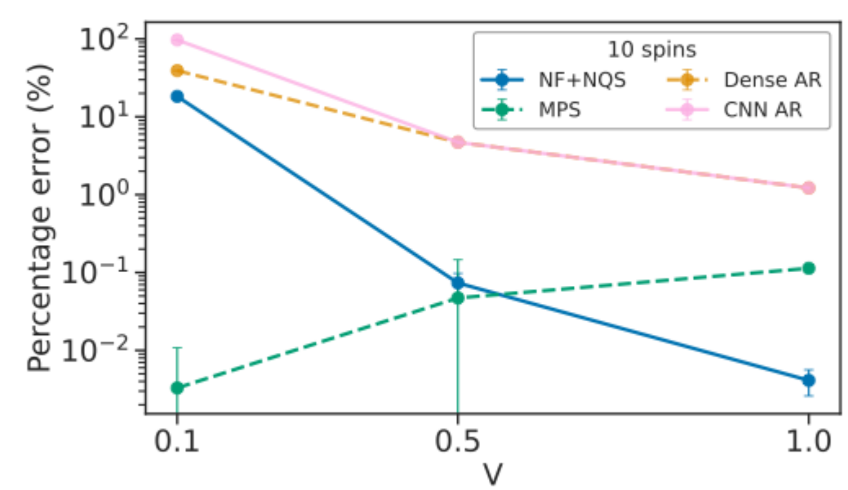
  ➡ Monte Carlo sampling of basis states $\left|\sigma_1, \ldots, \sigma_N\right\rangle$ to approximate energy ground state

  ➡ NN variation and energy optimisation are necessary for each choice of set of $\left|\sigma_1, \ldots, \sigma_N\right\rangle$

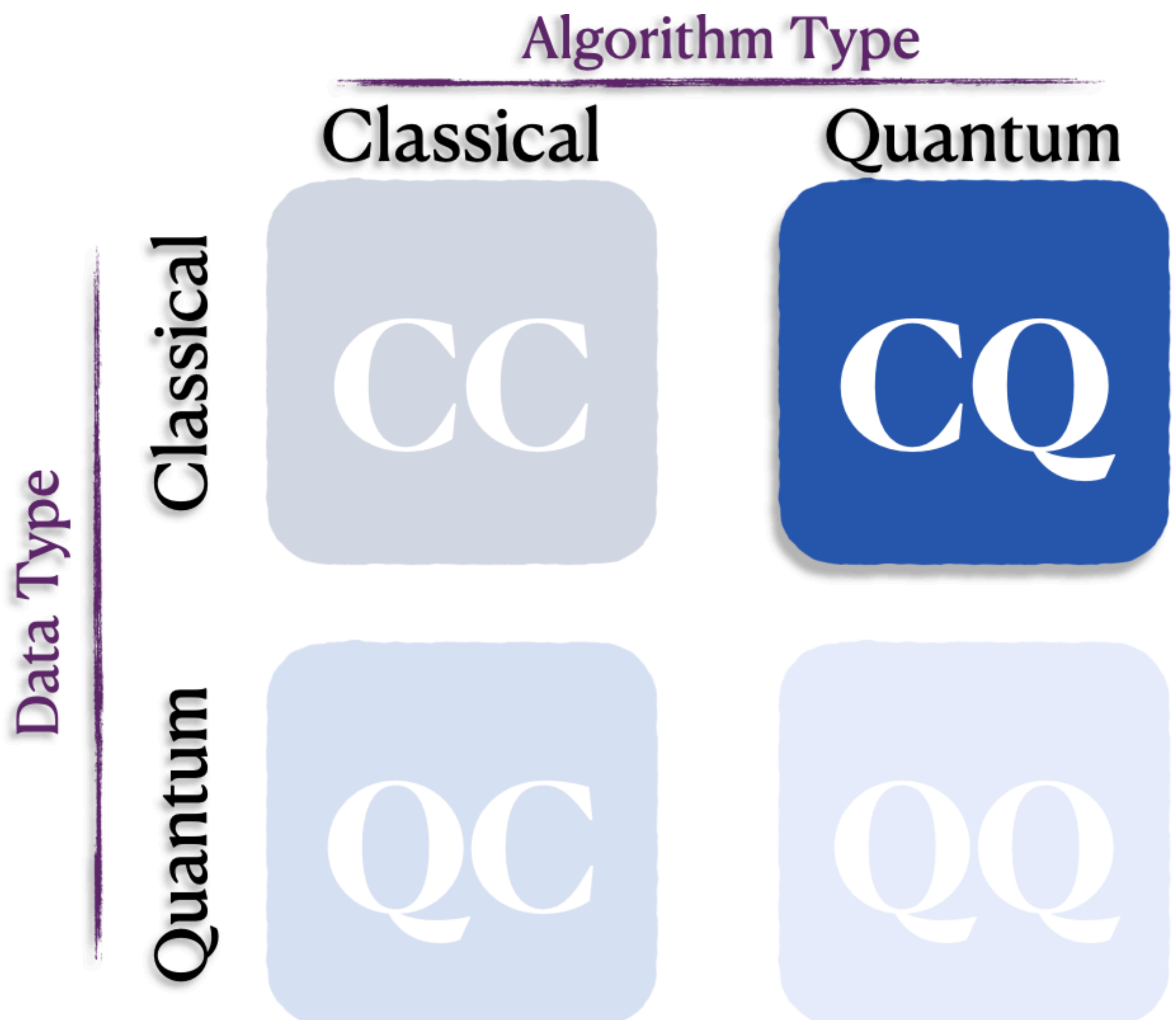
Example: Non-local transverse Ising model $H = -V \sum_{\{i,j\}\in\mathscr{E}} \sigma_i^z \sigma_j^z - \sum_{i\in\mathscr{N}} \sigma_i^x$



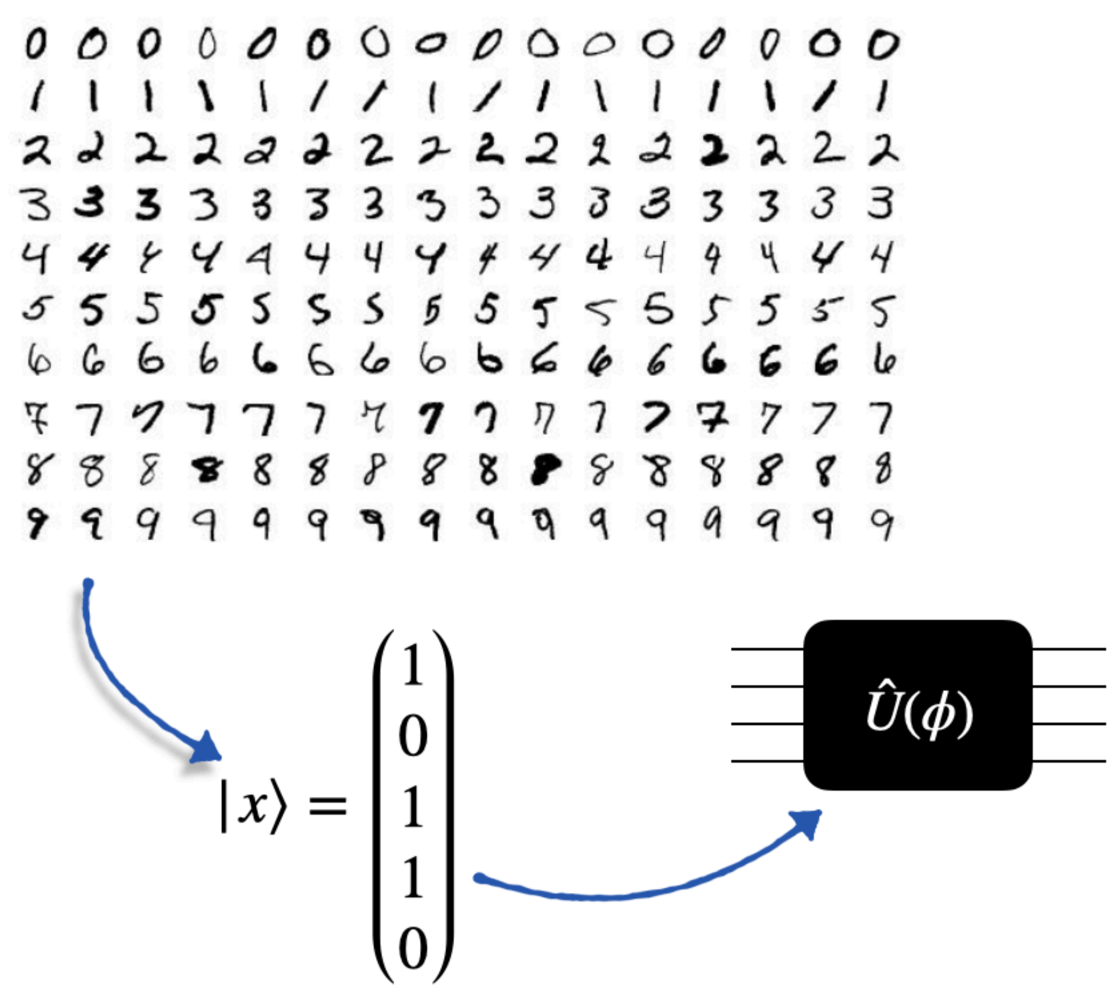




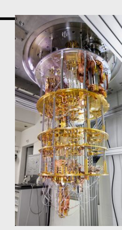
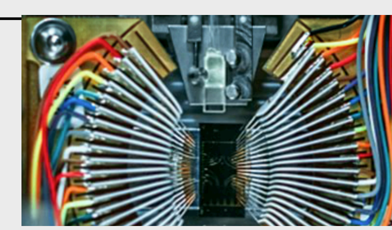
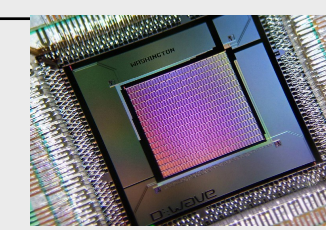
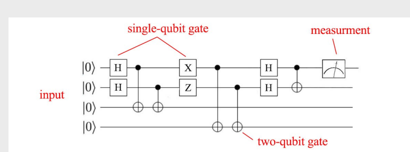
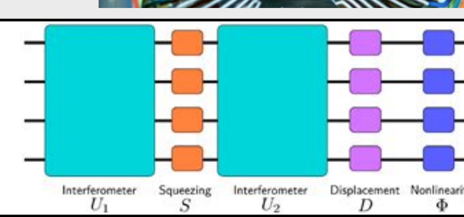
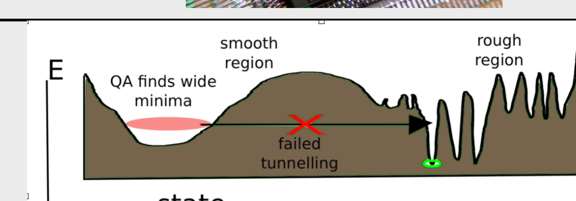
[Ngairangbam, MS, Sypchenko '25]



Classical data processed via quantum algorithms on quantum devices



Popular Quantum Computing paradigms

Type	Discrete Gate (DG)	Continuous Variable (CV)	Quantum Annealer (QA)
Computing	Digital	Digital/Analog	Analog
Property	Universal (any quantum algorithm can be expressed)	Universal - GBS non-Universal	Not universal — certain quantum systems
Advantage	most algorithms and tech support	uncountable Hilbert (configuration) space	continuous time quantum process
How?	IBM - Qiskit ~500 Qubits	Xanadu	DWave - LEAP ~7000 Qubits
What?			
			

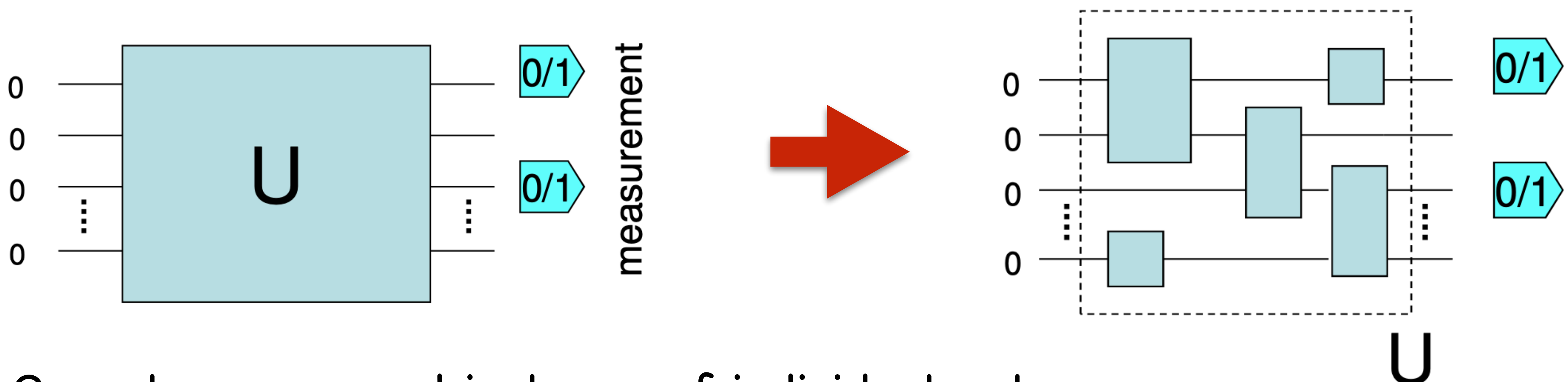
How most quantum algorithms work

operator acts on Hilbert space states $U\,|x\rangle = |\Psi_1\rangle$

measurement of observable $\hat{U}$ corresponds to exp. value of operator U $\left\langle \hat{U} \right\rangle_\Psi = \frac{\langle\Psi|U|\Psi\rangle}{\langle\Psi|\Psi\rangle}$

statistical statement need to evaluate often

Need to encode Hilbert space and operator suitable for quantum system

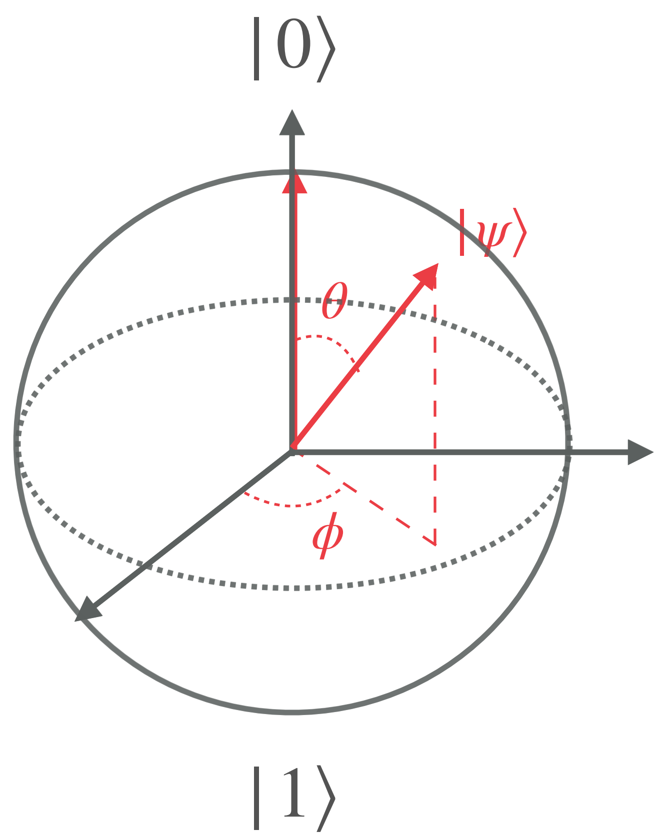


- Operator expressed in terms of individual gates
- Often 'Trotterization' (Suzuki-Trotter decomposition) needed:

For $H = \sum_{j=1}^{m} H_j$ $\longrightarrow$ $e^{iHt} = \left(\prod_{j=1}^{m} e^{-iH_j t/r} \right)^r + \mathcal{O}(m^2 t^2 / r)$

Rotation about the Bloch Sphere and state parametrisation

$$|\psi\rangle = \cos\frac{\theta}{2}|0\rangle + e^{i\phi}\sin\frac{\theta}{2}|1\rangle = \begin{pmatrix} \cos\frac{\theta}{2} \\ \sin\frac{\theta}{2}e^{i\phi} \end{pmatrix}$$



Measure

$|0\rangle$

$|1\rangle$ Prob$\left(|1\rangle\right) = \left(e^{i\phi}\sin\frac{\theta}{2}\right)^2$

$|0\rangle$ Prob$\left(|0\rangle\right) = \left(\cos\frac{\theta}{2}\right)^2$

$|1\rangle$

Apply Unitary rotation $U_3|0\rangle$: $U_3(\theta,\phi,\lambda) = \begin{pmatrix} \cos\left(\frac{\theta}{2}\right) & -e^{i\lambda}\sin\left(\frac{\theta}{2}\right) \\ e^{i\phi}\sin\left(\frac{\theta}{2}\right) & e^{i(\phi+\lambda)}\cos\left(\frac{\theta}{2}\right) \end{pmatrix}$

Extending this to a system of N qubits forms a 2^N-dimensional Hilbert Space

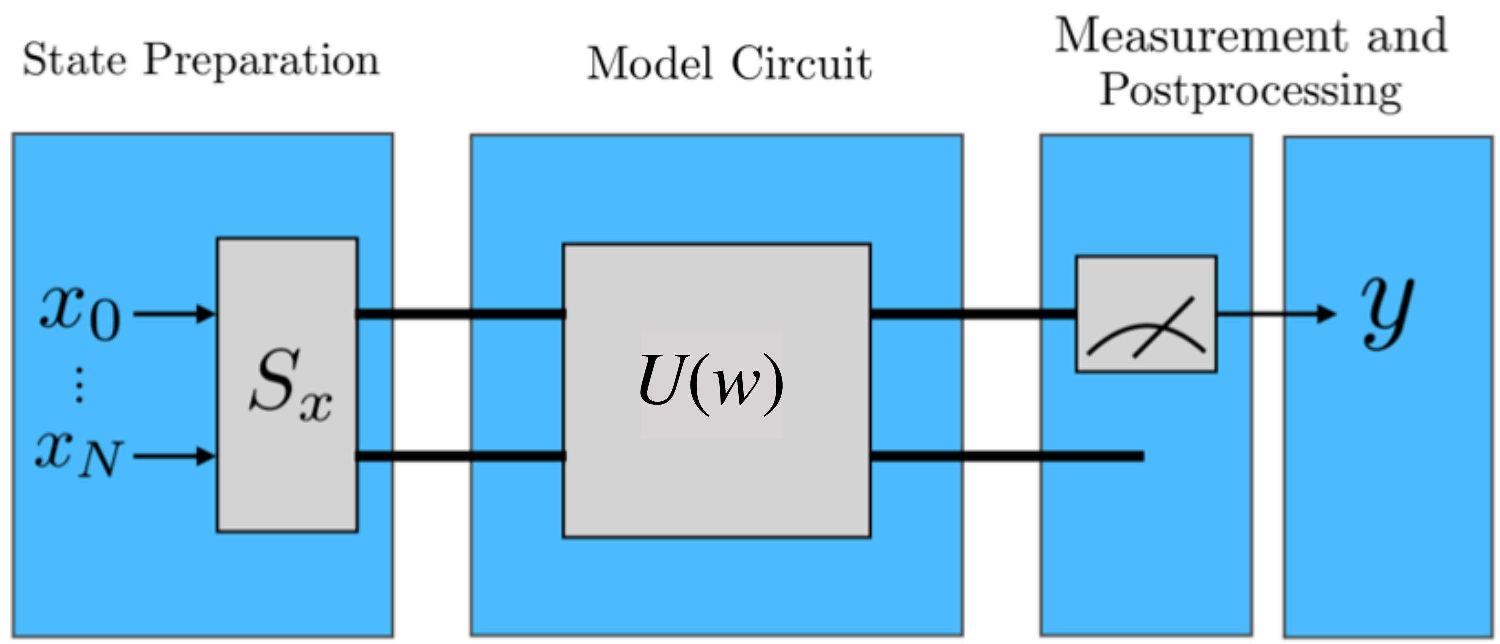
Quantum Machine Learning with a Variational Quantum Circuit

[McClean et al '16]

[Farhi, Neven '18]

[Schuld et al '20]

[Blance, MS '20]



Encoding step

Ports data into quantum state

Linear Unitary Operation

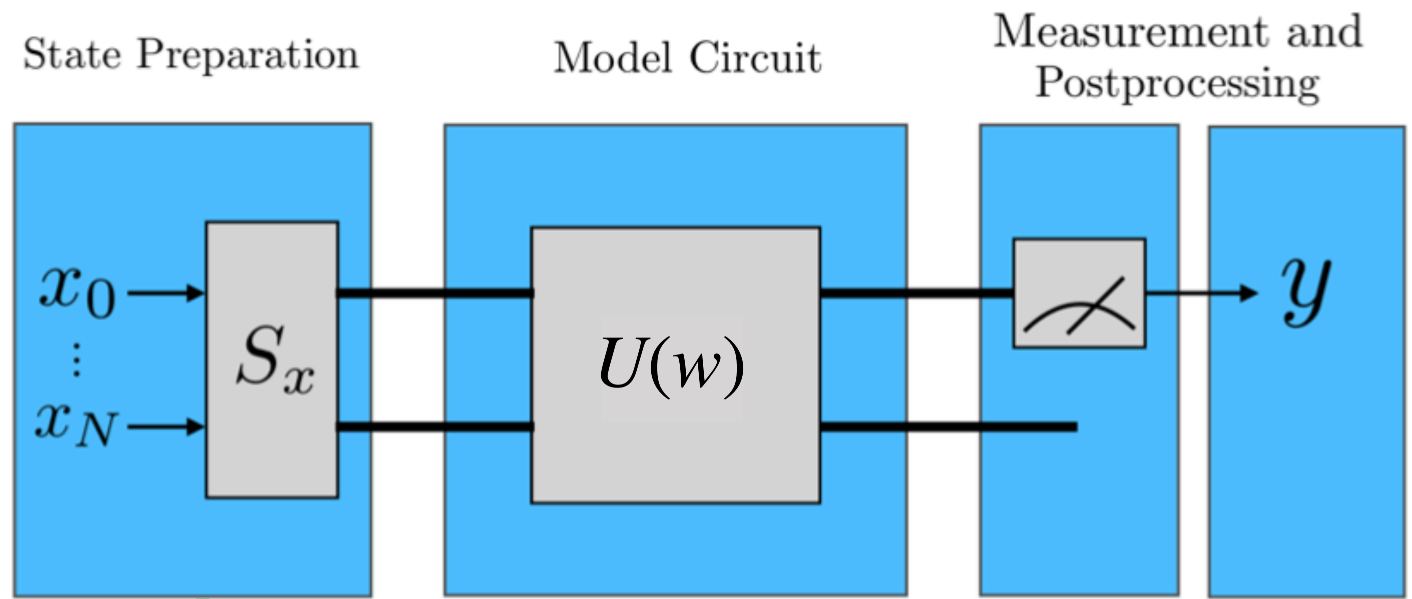
realises entanglement etc

Measurement with respect to operator, e.g. σ_z

expectation value of operator <-> many shots

difference to classical NN - expressibility relies on encoding step

Quantum Machine Learning
with a Variational Quantum Circuit



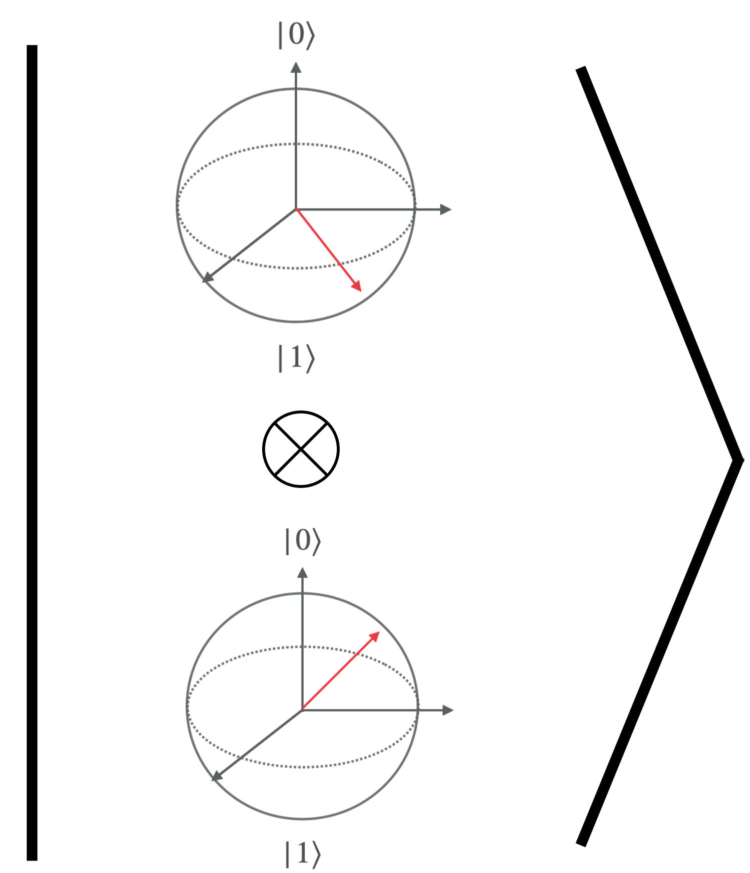
state preparation

n corresponds
to # features

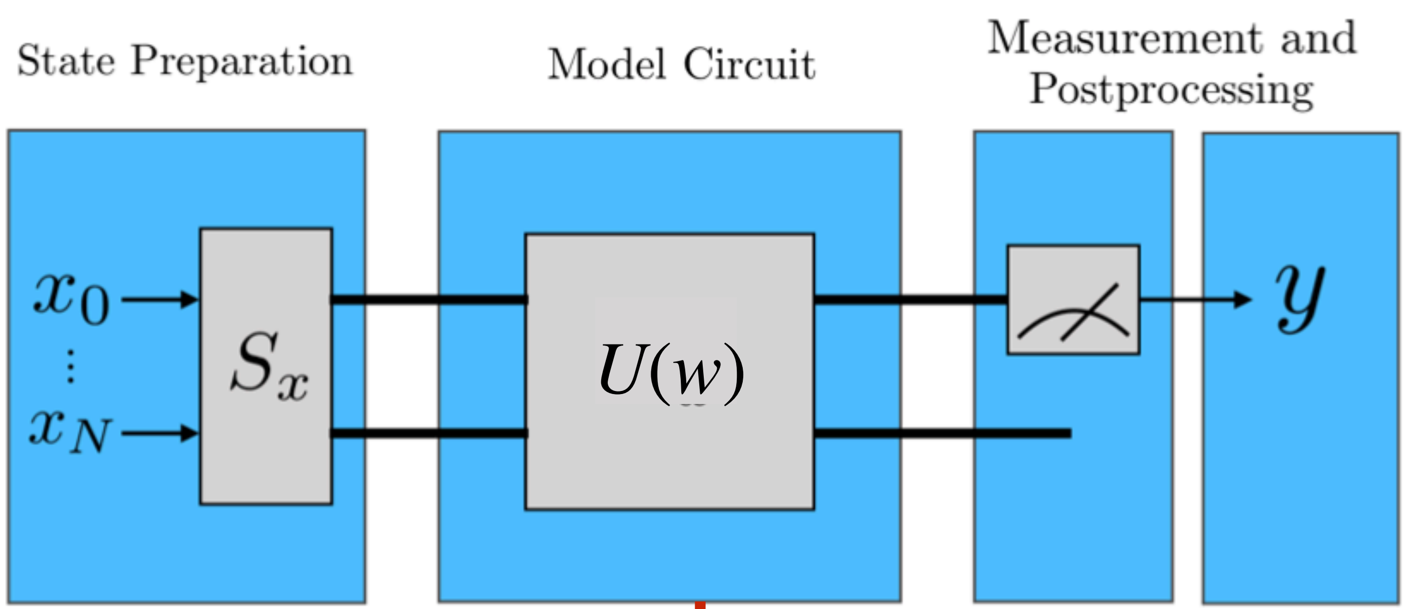
$$x \mapsto S_x|\phi\rangle = S_x|0\rangle^{\otimes n} = |x\rangle$$

e.g. angle encoding

$$|x\rangle = \bigotimes_{i=1}^{n} \cos(x_i)|0\rangle + \sin(x_i)|1\rangle$$



Quantum Machine Learning with a Variational Quantum Circuit



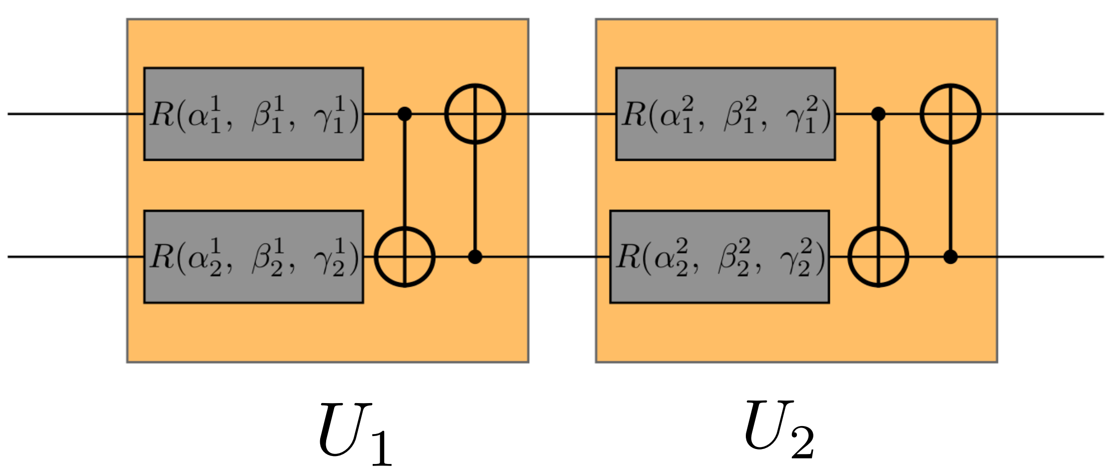
$$|\psi\rangle = U(w)|x\rangle$$ with $$U(w) = U_{l_{\max}}(w_{l_{\max}})...U_l(w_l)...U_1(w_1)$$

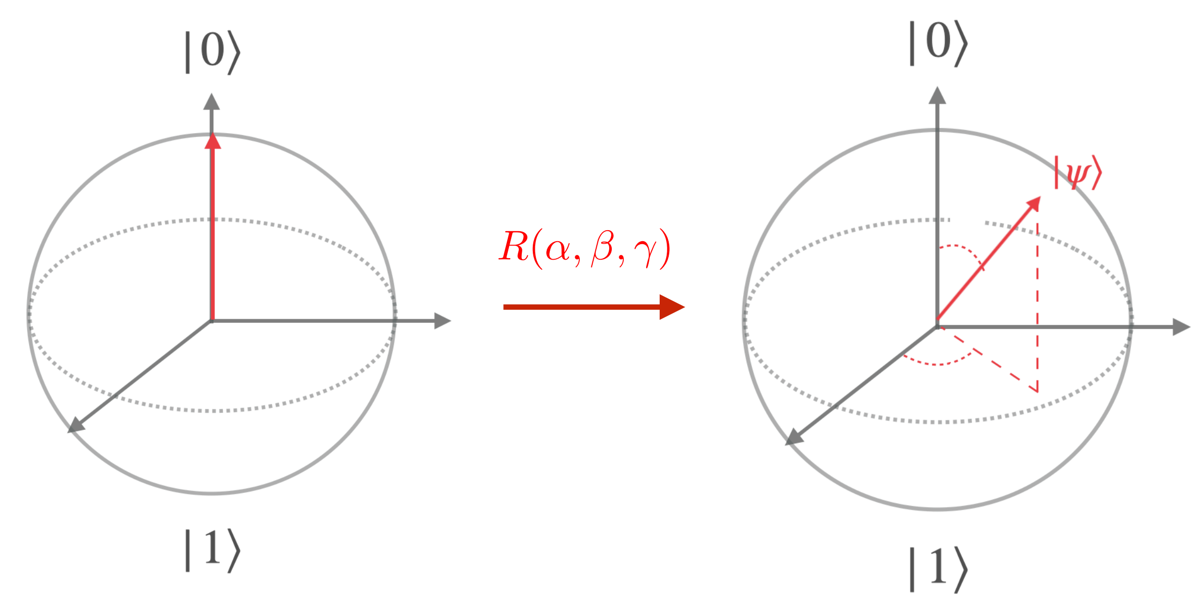
model circuit

trainable parameters

prepared state

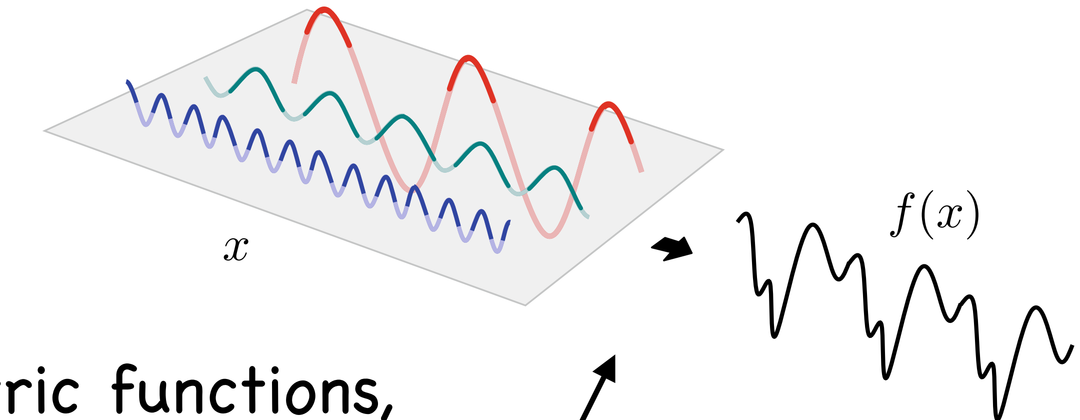
2-layer Variational Quantum Circuit





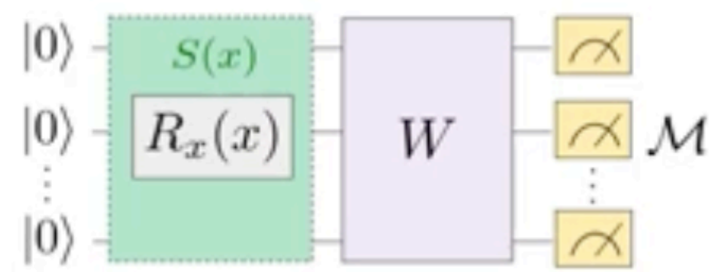
→ Rotation + CNOT -> Entanglement

Expressibility of model and encoding



- Most encodings result in sum of trigonometric functions, e.g. angle encoding, time evolution encoding

- Fourier series is universal approximator, but for many encoding strategies quantum models are linear combinations of functions composed of few frequencies



- Pendant to activation functions in the encoding step.

- Encoding + W operator give functional form

$$\boxed{f_\theta(x) = \langle \mathscr{M} \rangle_{x,\theta} = A + B\cos(x) - C\sin(x)}$$

A, B, C coefficients from parametrised circuit W

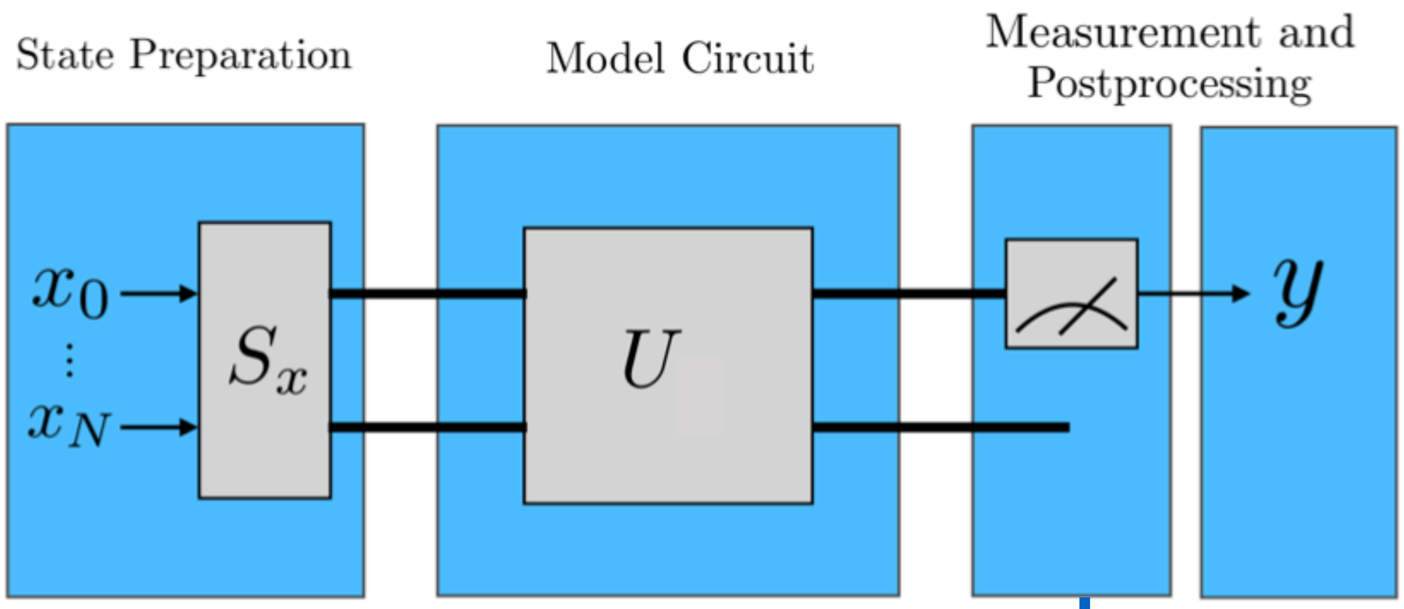
trigonometric structure from data encoding

- Data reuploading can increase expressivity

➡ Universal Approximation Theorem

[Perez-Salina et al 2020] [Schuld et al '20]

Quantum Machine Learning with a Variational Quantum Circuit



- Entangled state shares information across qubits
- Evaluate expectation value of qubits to construct loss

  for supervised S vs B classification one qubit sufficient

  $$\mathbb{E}(\sigma_z) = \langle 0|S_x(x)^\dagger U(w)^\dagger \hat{O} U(w) S_x(x)|0\rangle = \pi(w,x) \quad \text{for} \quad \hat{O} = \sigma_z \otimes \mathbb{I}^{\otimes(n-1)}$$

- Quantum network output: $f(w,b,x) = \pi(w,x) + b$
- Changing operator and loss => VQE, VQT, ... (simulate QFT)

Simple example:

$$|0\rangle \;-\boxed{R_x(x)}-\boxed{\mathrm{Rot}(\theta_1, \theta_2, \theta_3)}-\boxed{\nearrow}\quad \sigma_z$$

gives the model output $\quad f_\theta(x) = \langle 0|R_x(x)^\dagger \mathrm{Rot}(\theta_1, \theta_2, \theta_3)^\dagger \sigma_z \mathrm{Rot}(\theta_1, \theta_2, \theta_3) R_x(x)|0\rangle$

data encoding

$$|\phi(x)\rangle = R_x(x)|0\rangle = \begin{pmatrix} \cos(\frac{x}{2}) & -i\,\sin(\frac{x}{2}) \\ -i\,\sin(\frac{x}{2}) & \cos(\frac{x}{2}) \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} \cos(\frac{x}{2}) \\ -i\,\sin(\frac{x}{2}) \end{pmatrix}$$

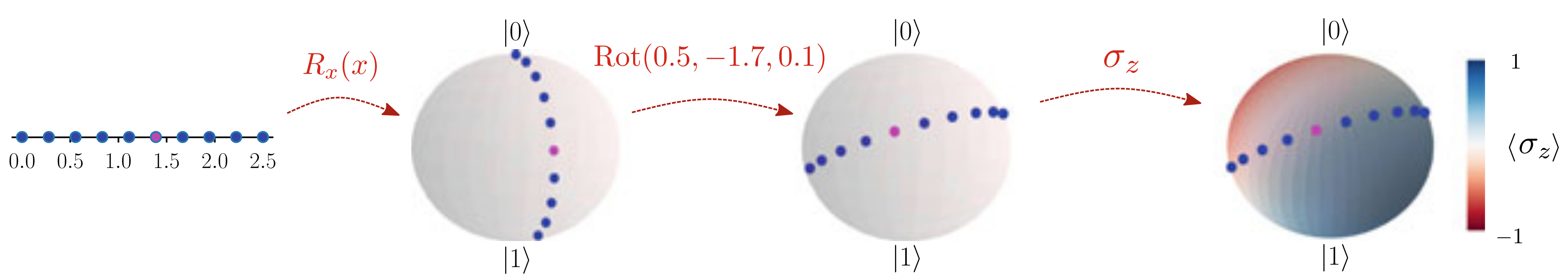
parametrised rotation

$$\begin{aligned} |\psi(x, \theta)\rangle &= \mathrm{Rot}(\theta_1, \theta_2, \theta_3)|\phi(x)\rangle \\ &= \begin{pmatrix} e^{i(-\frac{\theta_1}{2} - \frac{\theta_3}{2})} \cos(\frac{\theta_2}{2}) \cos(\frac{x}{2}) + i\, e^{i(-\frac{\theta_1}{2} + \frac{\theta_3}{2})} \sin(\frac{\theta_2}{2}) \sin(\frac{x}{2}) \\ e^{i(\frac{\theta_1}{2} - \frac{\theta_3}{2})} \sin(\frac{\theta_2}{2}) \cos(\frac{x}{2}) - i\, e^{i(\frac{\theta_1}{2} + \frac{\theta_3}{2})} \cos(\frac{\theta_2}{2}) \sin(\frac{x}{2}) \end{pmatrix} \end{aligned}$$

model output

$$f_\theta(x) = \langle \psi(x, \theta)|\sigma_z|\psi(x, \theta)\rangle = \cos(\theta_2)\cos(x) - \sin(\theta_1)\sin(\theta_2)\sin(x)$$

What happens on the Bloch-Sphere



Classifier from simple example:

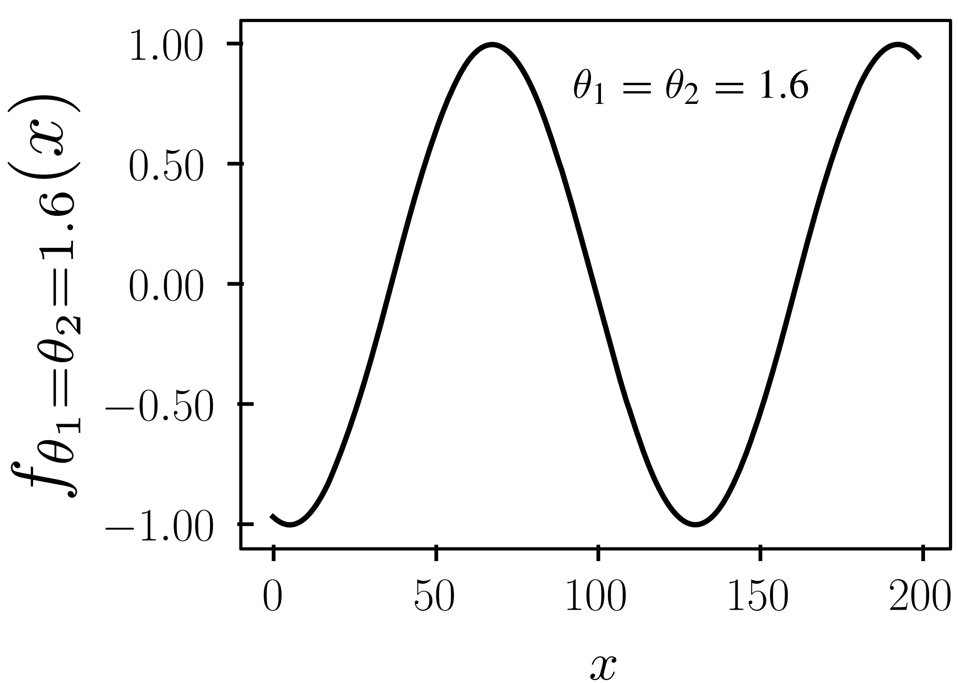
for binary classifier define e.g.

$$y = \begin{cases} 1 \text{ if } f_\theta(x) > 0 \\ -1 \text{ else} \end{cases}$$

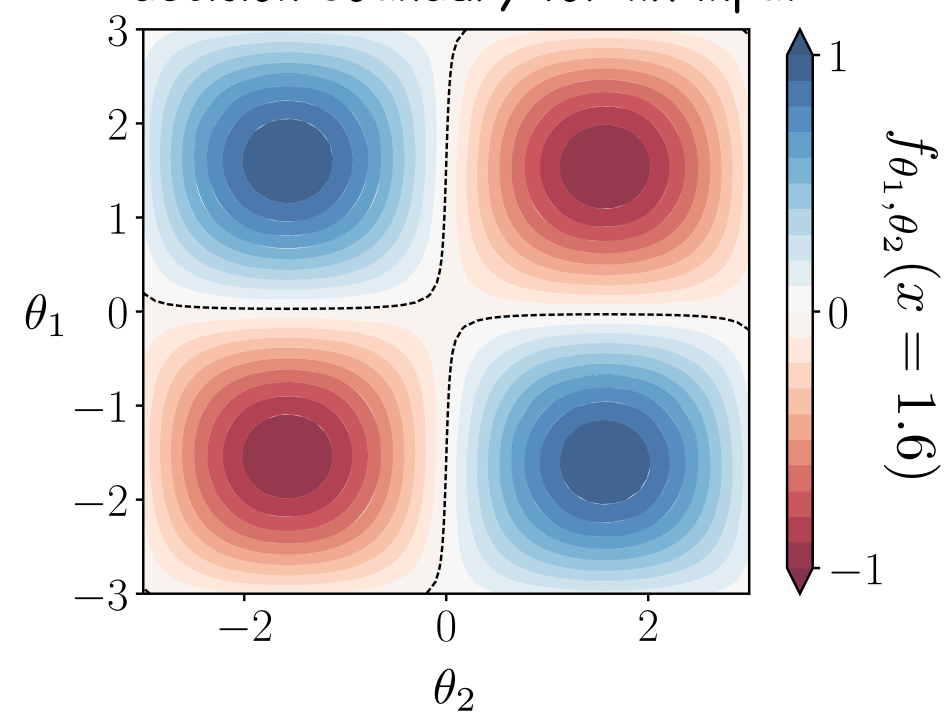
for probabilistic classifier
(density estimator)

$$p(1) = \frac{f_\theta(x) + 1}{2}, \qquad p_0 = 1 - p_1$$

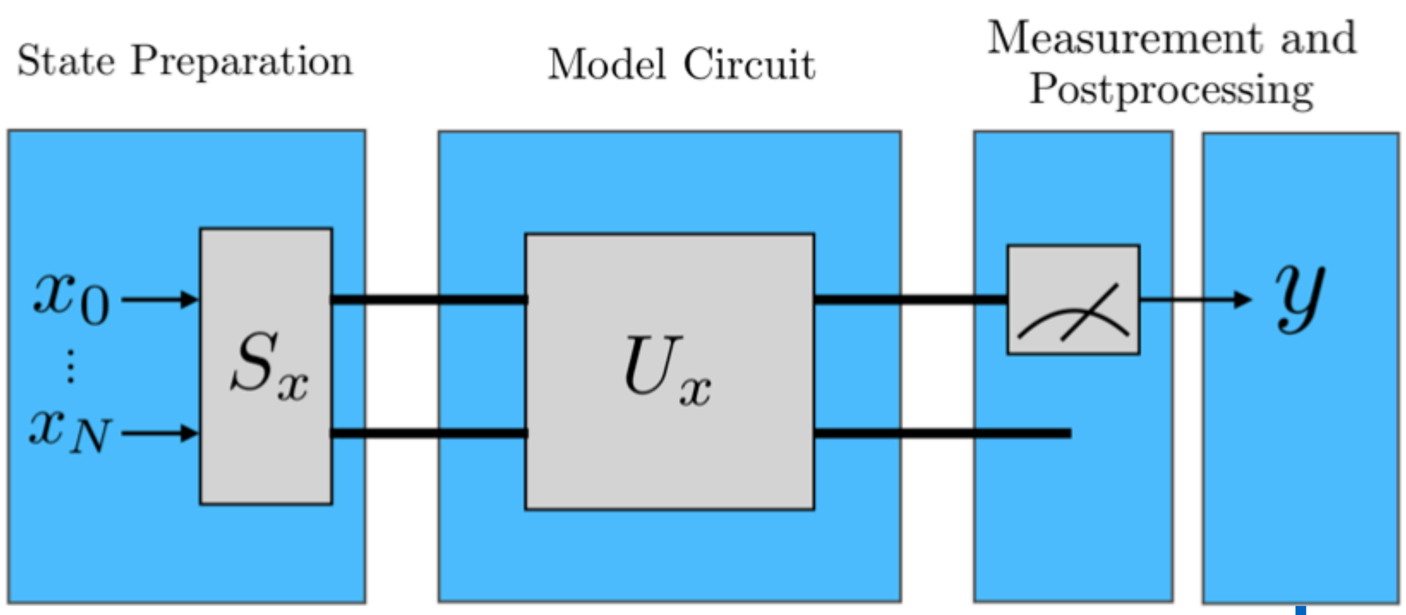
model output for fix angles



decision boundary for fix input



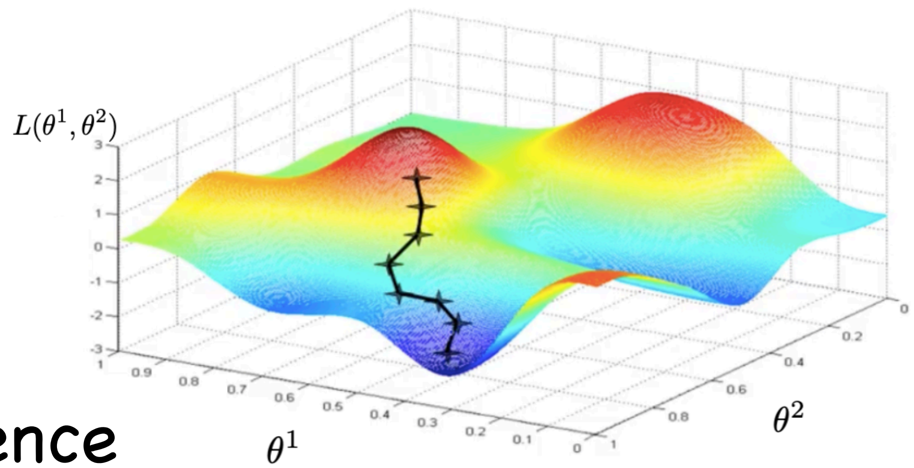
Quantum Machine Learning with a Variational Quantum Circuit



- Hybrid approach (QC to calculate exp. value, CC to optimise U operator)

- Loss function $$L = \frac{1}{n}\sum_{i=1}^{n}\left[y_i^{\text{truth}} - f(w,b,x_i)\right]^2$$

label (signal, bkg), supervised learning



- Quantum gradient descent - for fast convergence

Fubiny-Study metric underlies geometric structure of VQC parameter space: $\theta_{t+1} = \theta_t - \eta g^{+}\nabla L(\theta)$

[Cheng '10]
[Blance, MS '20]
[Abbas et al '20]

Barren Plateaus

- Area in loss landscape where gradients are close to zero

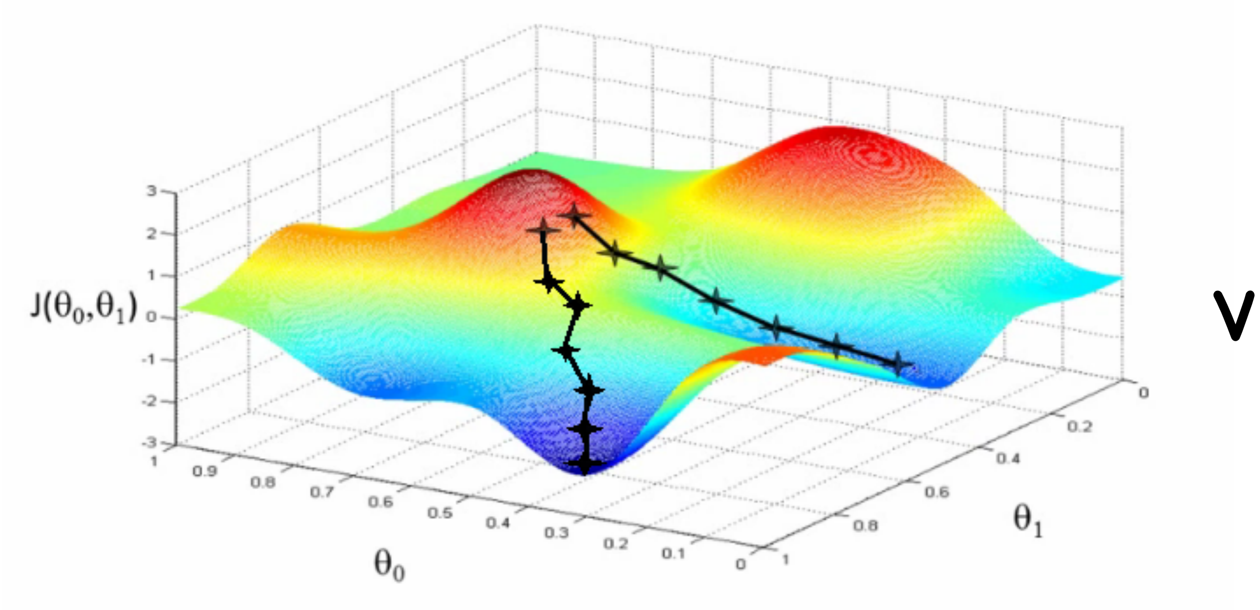
- Optimisation is slow and expensive, requiring high accuracy in evaluating gradient to avoid random walking

  ⟶ Barren Plateaus often arise if quantum model is overly expressive and Hilbert spaces are large

  ⟶ Individual gradient steps in exponentially large parameter and Hilbert space becomes less relevant

  ⟶ Important task for efficient learning is the choice of model, i.e. **as expressive as necessary while as small as possible**

Valley of mountain



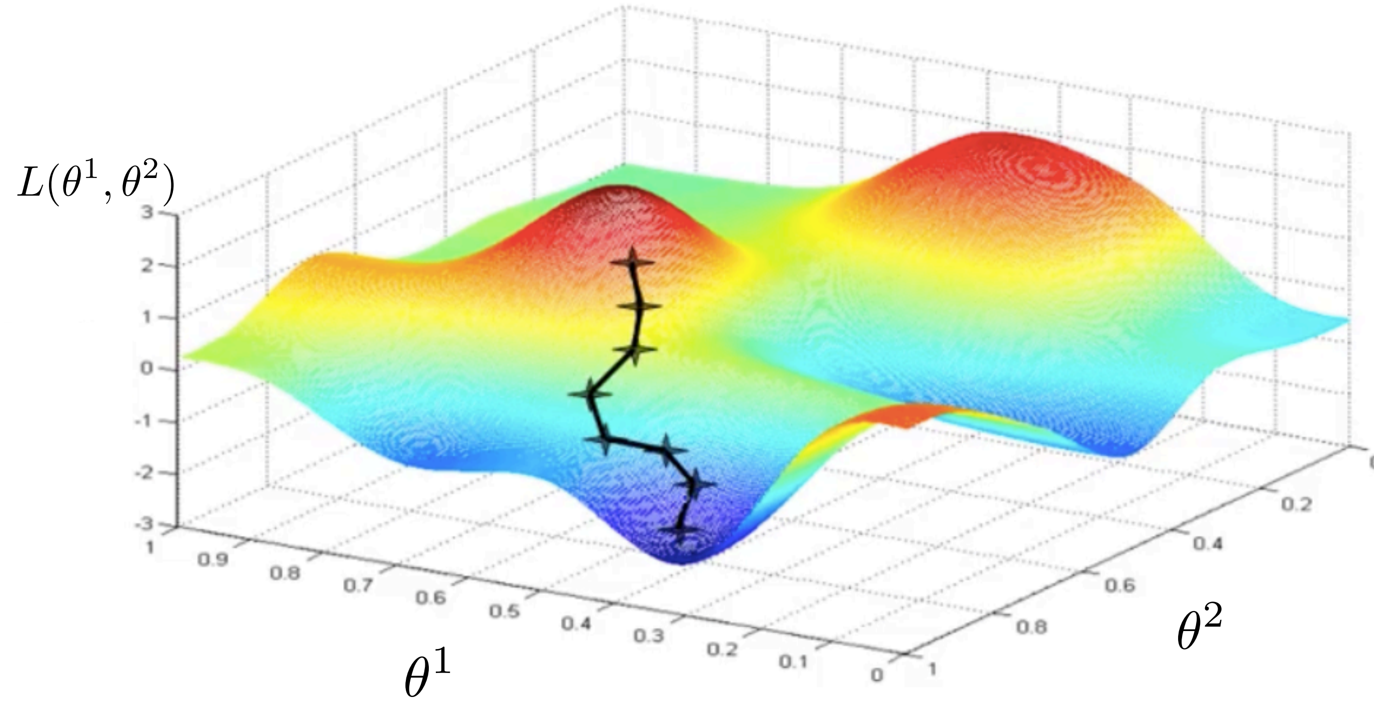
VS

Hole on golf course

Optimising the loss landscape

classical gradient descent (GD):

$$\theta_{t+1} = \theta_t - \eta \nabla L(\theta)$$



quantum gradient descent (QDC):

Fisher Information Matrix F promotes gradient descent to natural gradient descent (Riemannian geometry):

$$\theta_{t+1} = \theta_t - \eta F^{-1} \nabla L(\theta)$$

[Stokes, Izaac, Killoran, Carleo '20] [Blance, MS '20]

Fubiny-Study metric underlies geometric structure of VQC parameter space (complex projective Hilbert Spaces):

$$\theta_{t+1} = \theta_t - \eta g^{+} \nabla L(\theta)$$

VQC parameters

weights $\theta^w_{t+1} = \theta^w_t - \eta g^{+} \nabla^w L(\theta)$,

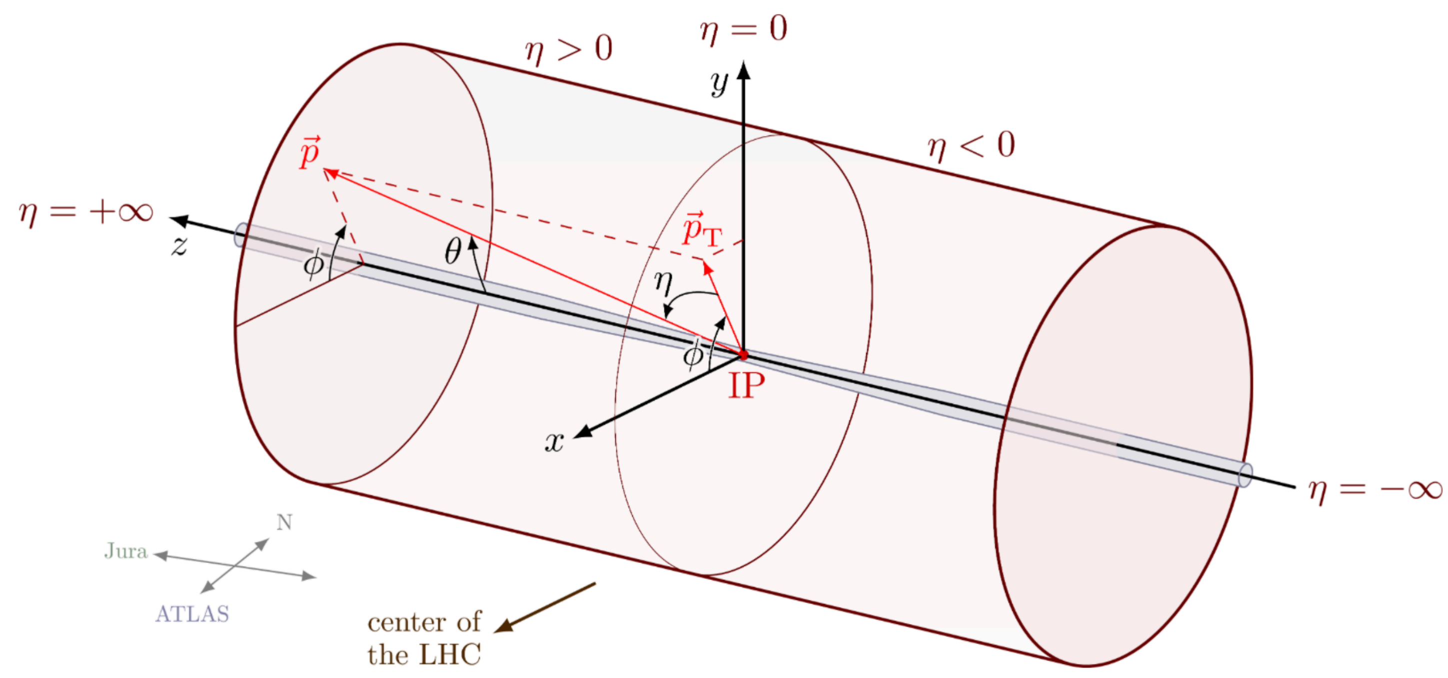
bias $\theta^b_{t+1} = \theta^b_t - \eta \nabla^b L(\theta)$,

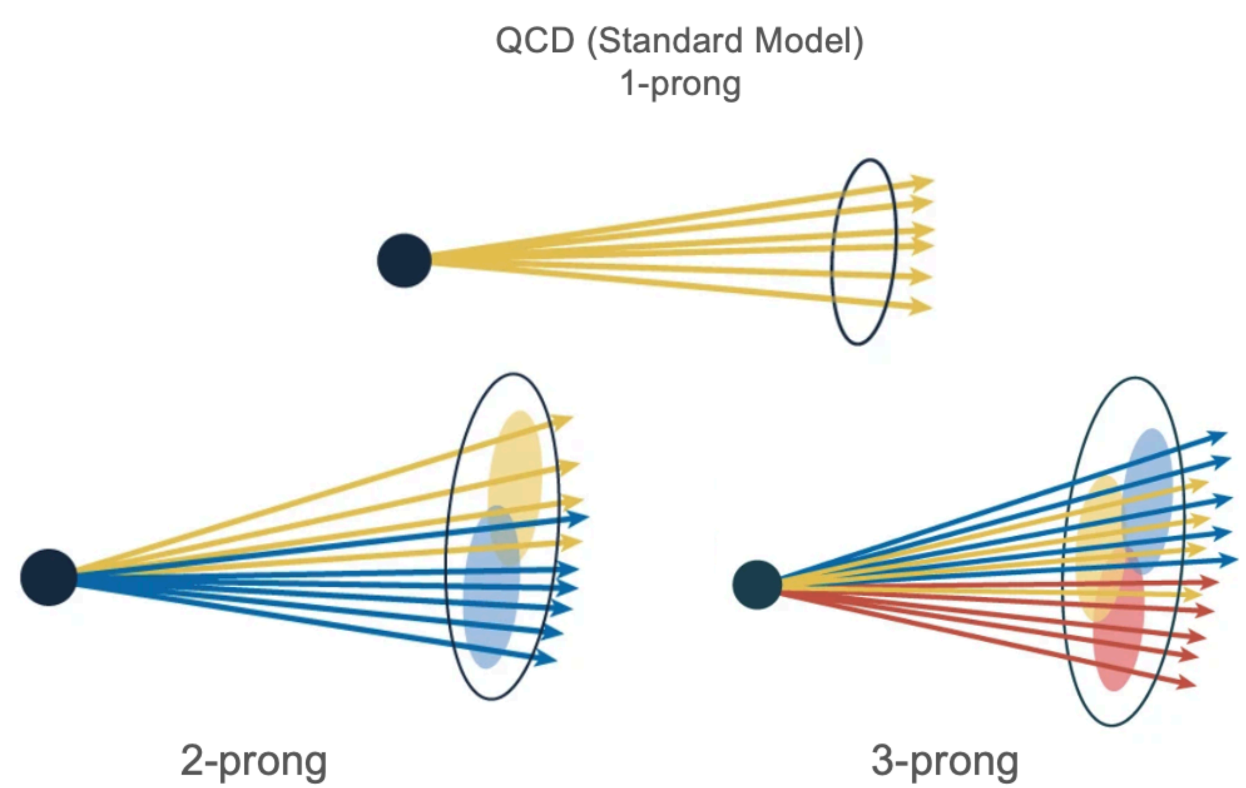
HEP Example

- Each particle defined by 3 features (ϕ, η, p_T)
- LHC events consist of $\mathcal{O}(500)$ particles
- Fat jets have $\mathcal{O}(10)$ subjets

↓

Efficient data encoding crucial for realistic data analysis on quantum device





1P1Q Encoding

[Bal, Klute, Maier, Oughton, Pezone, MS '25]

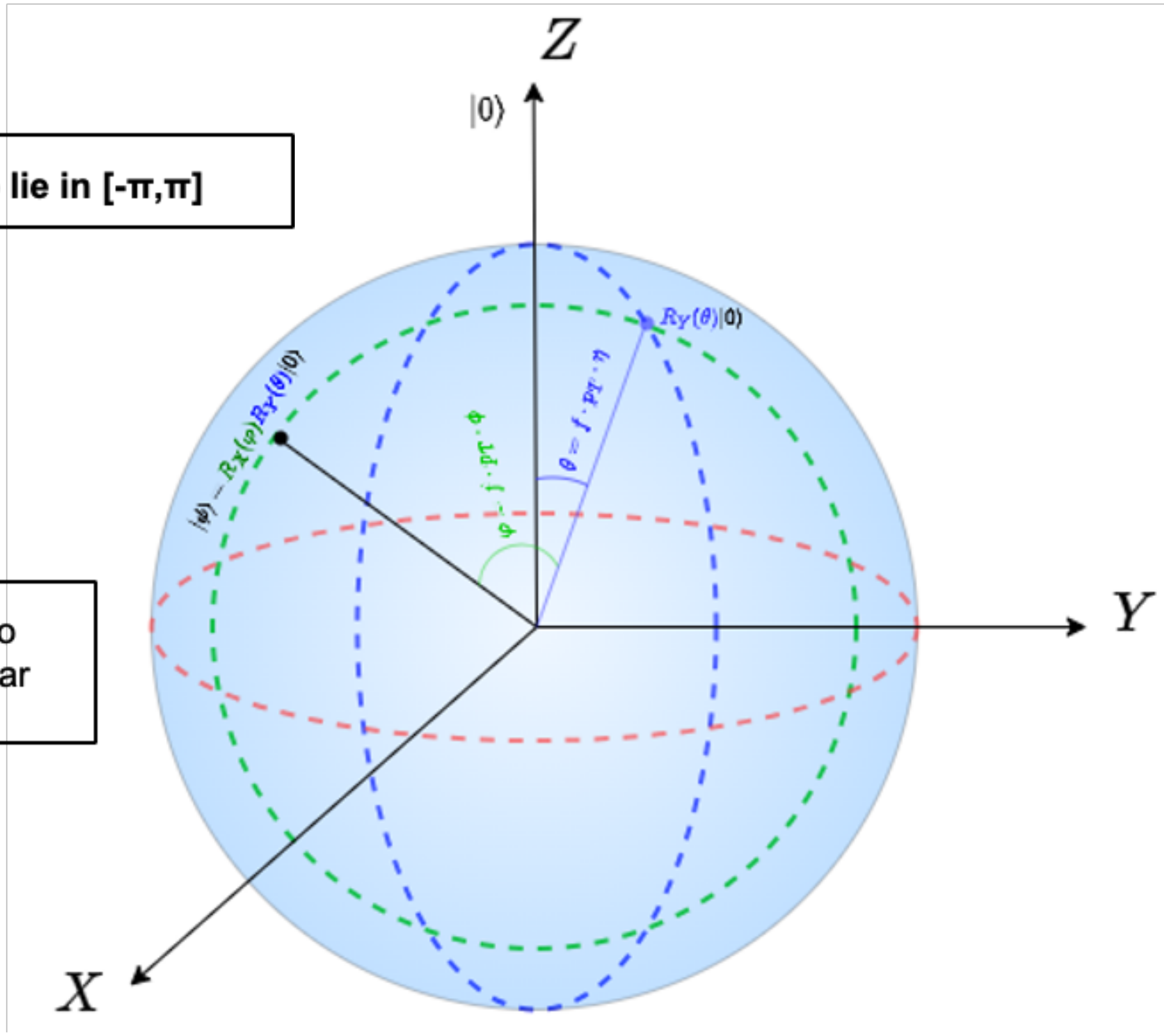
$$(p_T, \eta) \quad \rightarrow \theta = f \cdot \frac{p_T}{p_{T,jet}}(\eta - \eta_{jet})$$

Scaled to lie in [-π,π]

$$(p_T, \phi) \quad \rightarrow \varphi = f \cdot \frac{p_T}{p_{T,jet}}(\phi - \phi_{jet})$$

$$f = 1 + \frac{2\pi}{1 + e^{-w}}$$

trainable parameter to prevent clustering near North Pole

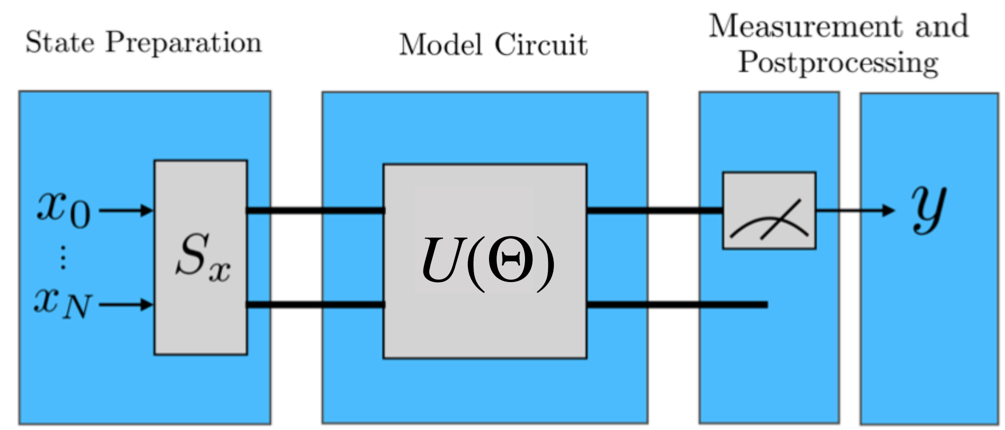


$$(p_\mathrm{T}, \eta, \phi) \quad \rightarrow \quad \big|\psi\rangle = R_X(\varphi) R_Y(\theta) \big|0\rangle$$

$$(p_\mathrm{T}, \eta, \phi) \quad \rightarrow \quad |\psi\rangle \quad = \left(\cos\frac{\theta}{2}\cos\frac{\varphi}{2} + i(\sin\frac{\theta}{2}\sin\frac{\varphi}{2})\right)|0\rangle + \left(\sin\frac{\theta}{2}\cos\frac{\varphi}{2} - i(\cos\frac{\theta}{2}\sin\frac{\varphi}{2})\right)|1\rangle$$

Supervised-Learning with Variational Quantum Circuit

- VQC supervised learning algorithm

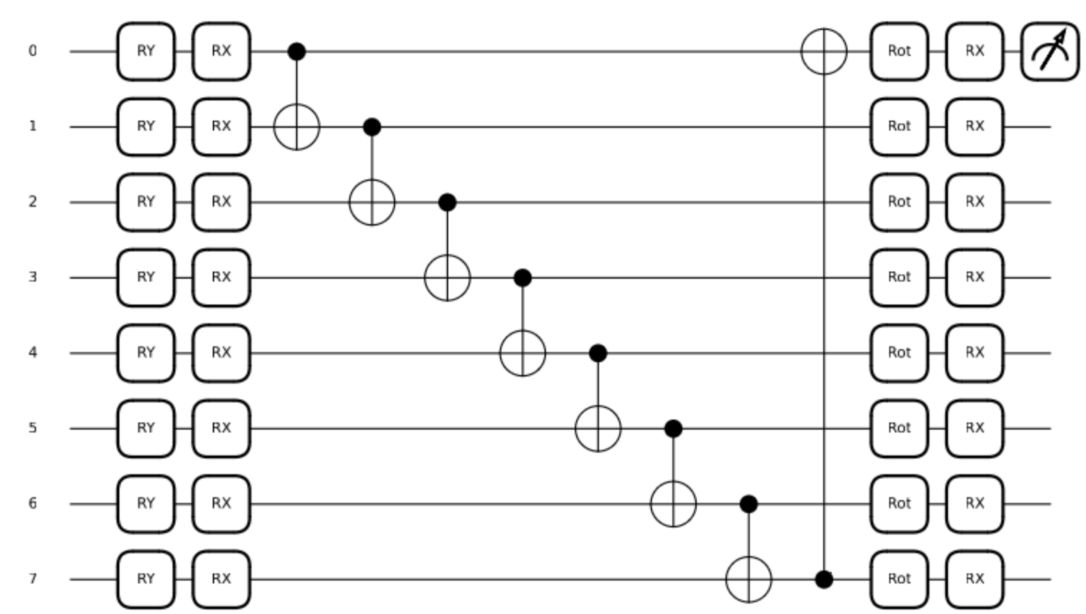


- Use the JetClass dataset first introduced by authors of Particle Transformer (ParT)

- Train on labelled data,
  signal = boosted top quarks
  bkg = QCD fat jets

$$U(\Theta) = \left(\bigotimes_{i=1}^{N-1} R_X(\phi_i) R_Y(\theta_i) R_Z(\omega_i) \right) \otimes \left(\bigotimes_{i=0}^{N} C_{i,(i+1) \bmod N} \right)$$

- Avoid jet bias
  -> flat p_T in [500,1000] GeV

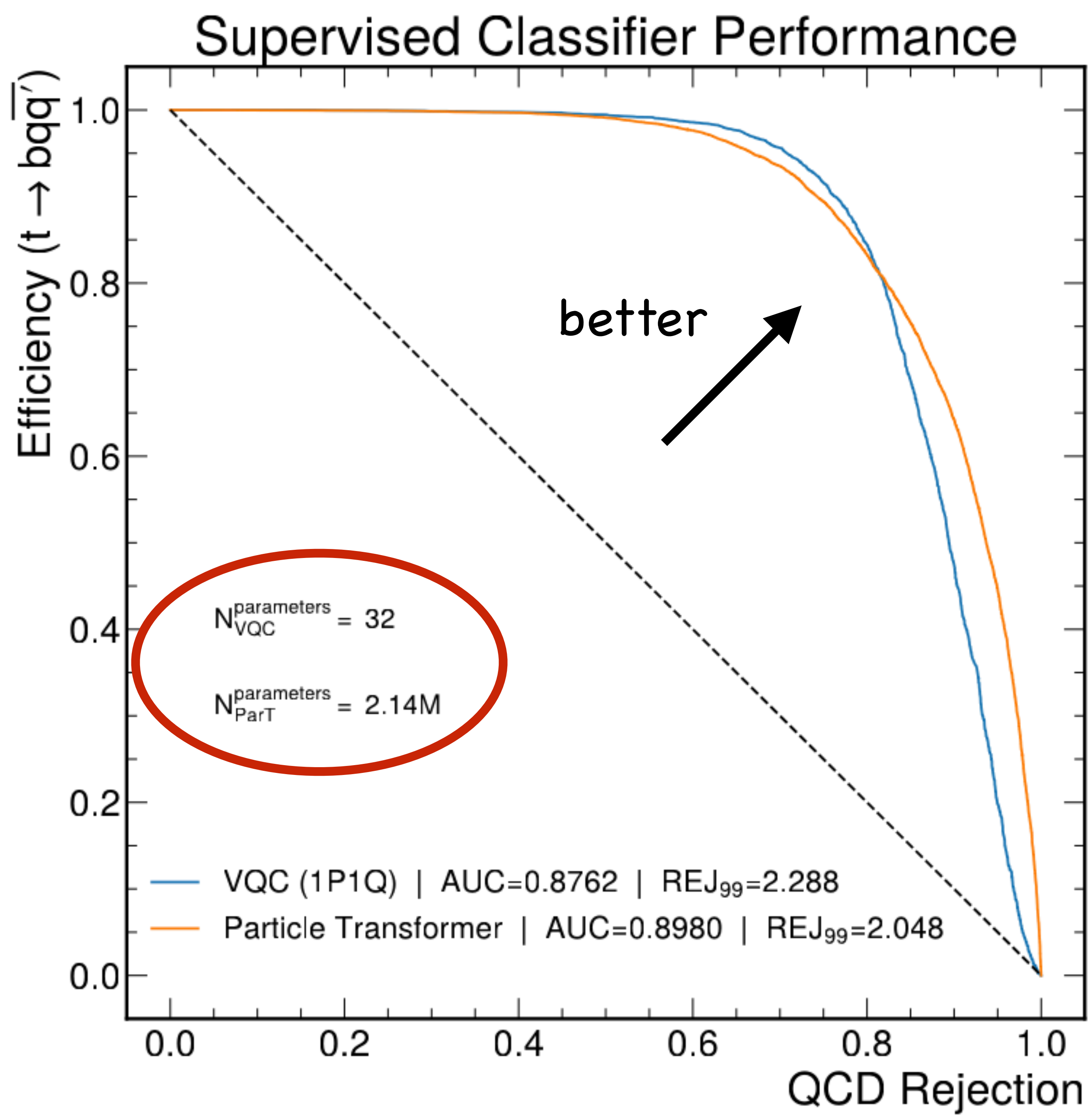


- Train only on 3 basic kinematic featrues (p_T, η, ϕ) with appropraite sclaing and normalization

  → Assign one particle to one qubit

  → Jet represented by N qubits (N hardest constits)

Results for VQC with 1P1Q Encoding



- Input are 10 hardest particles of fat jet

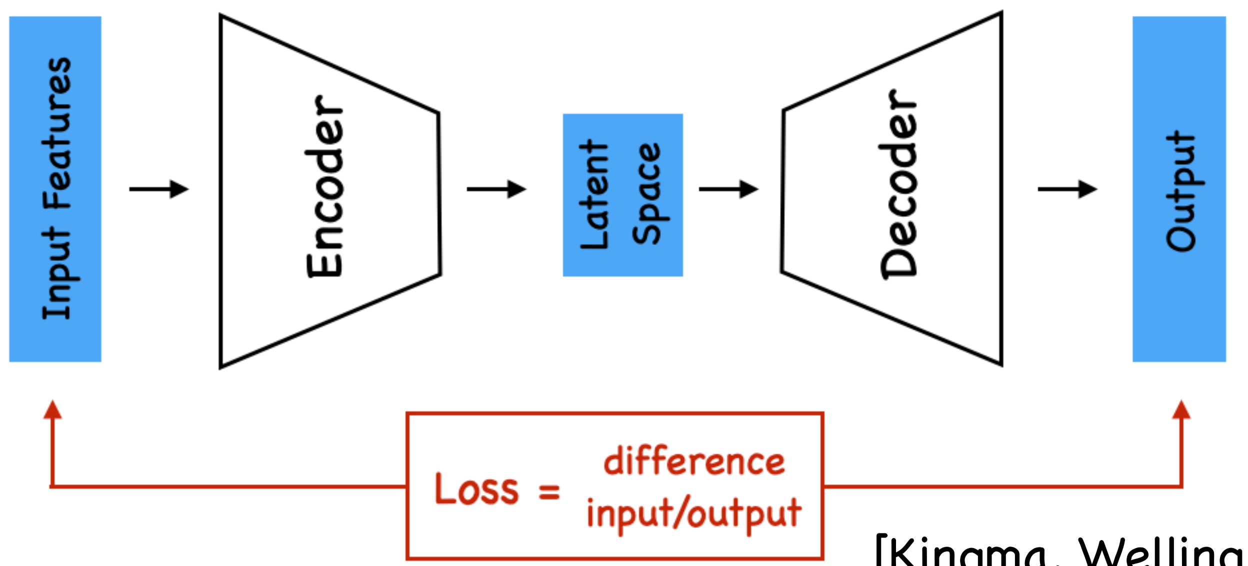
- Tiny VQC performs comparably to state-of-the-art Particle Transformer

- Drastically reduced model parameters: VQC (**32**) vs Transformer (**2.14 Millions**)

[Bal, Klute, Maier, Oughton, Pezone, MS '25]

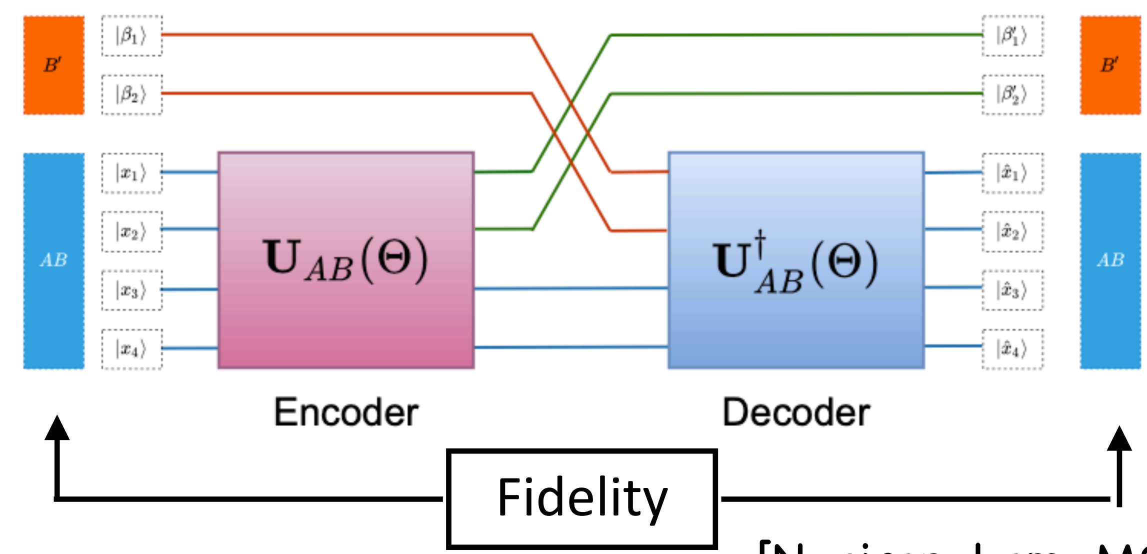
Unsupervised learning with Quantum Autoencoders

Classical Autoencoder (CAE)



[Kingma, Welling '13]

Quantum Autoencoder (QAE)



[Ngairangbam, MS, Takeuchi '21]

- Freature input is encoded into information bottleneck, i.e. latent space with smaller dimension that feature space

- Latent space decoded into reconstructed output, which is then compared with input via loss-function (often MSE)
  -> Encoder+Decoder trained together to produce output similar to input

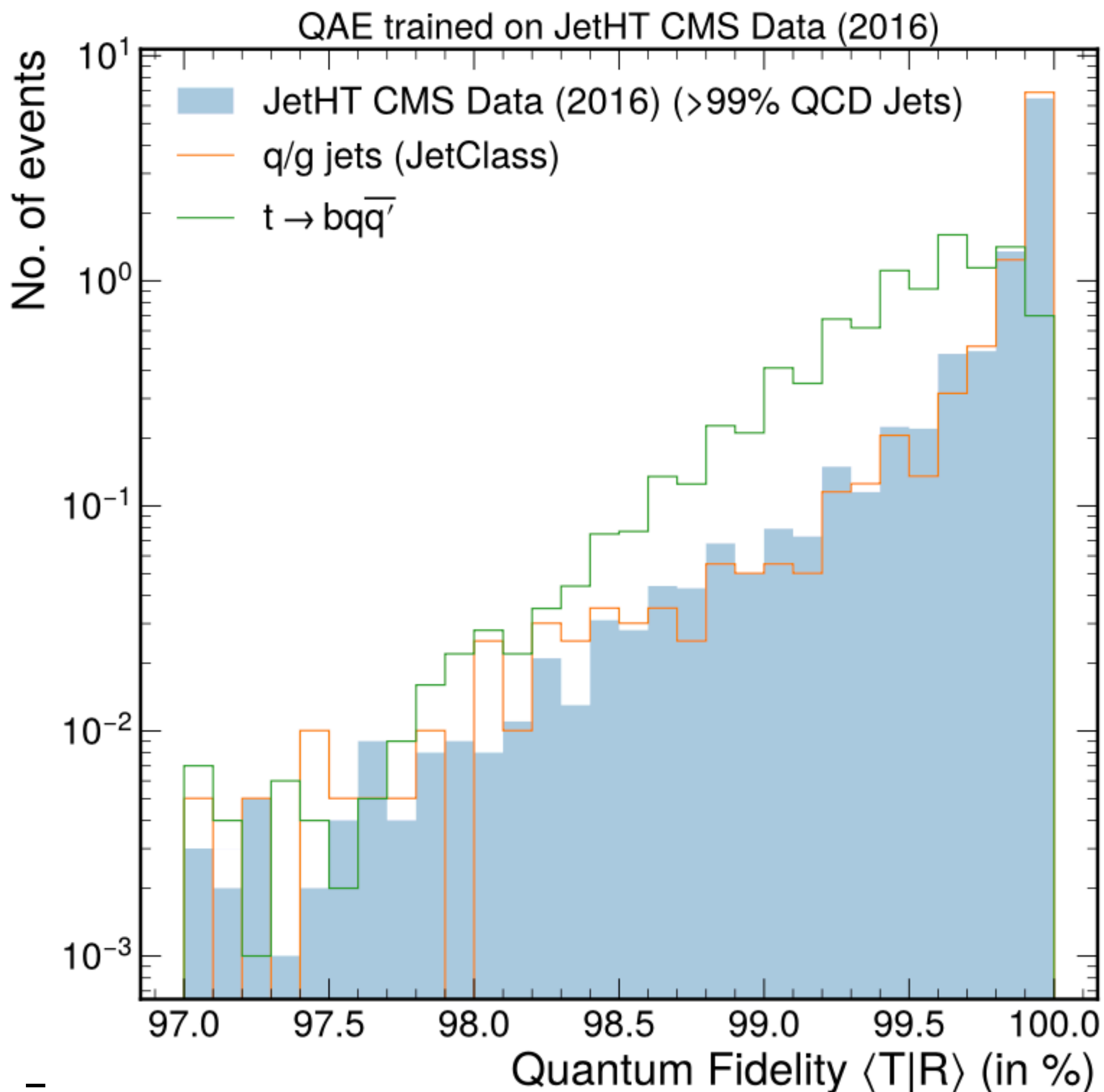
- Quantum AE needs to work with unitary gate operations.
  Thus, need trash states to realise information bottleneck

- Fidelity for QAE obtained from CMS data

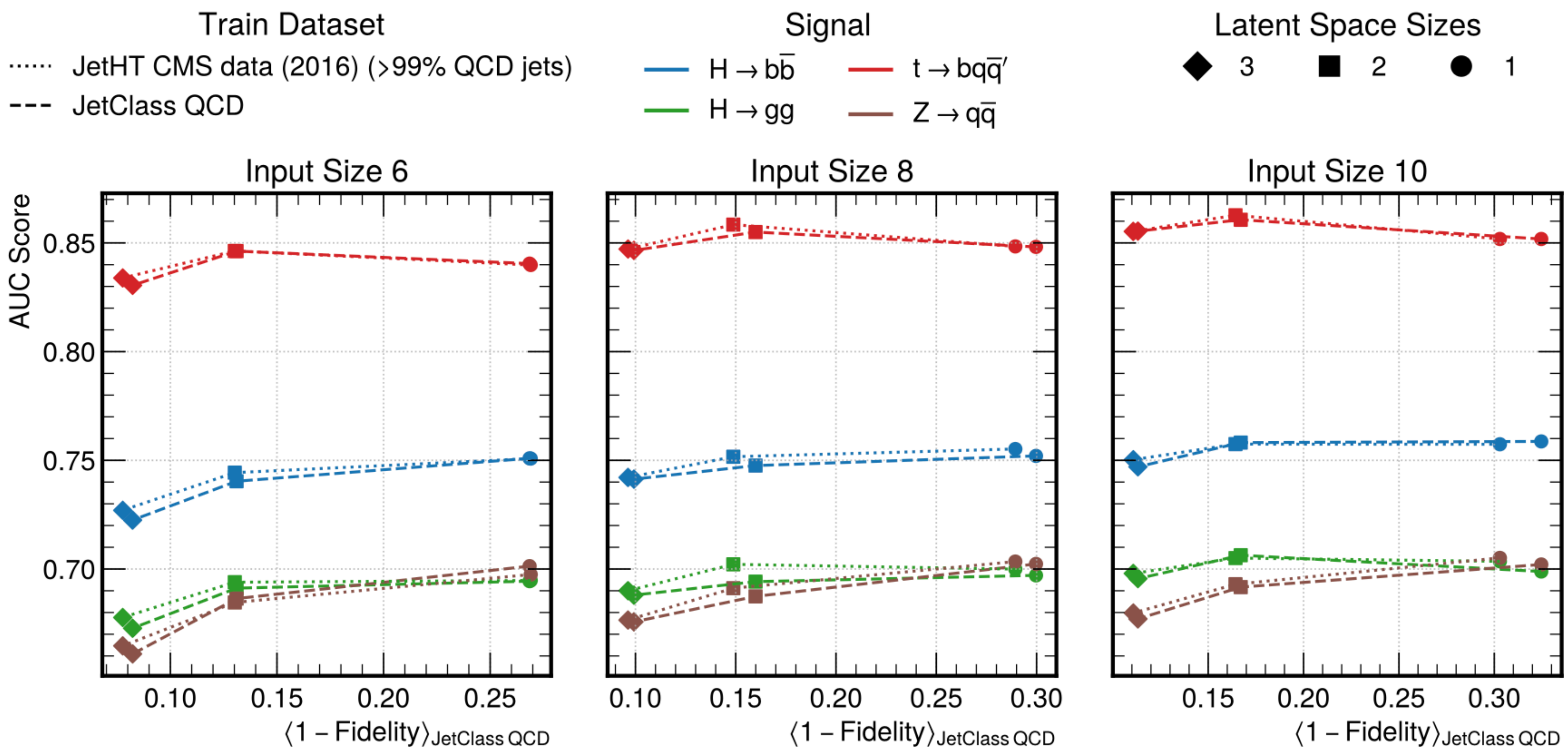
- Very good data - simulation agreement

- Signals and bkgs flat in pT in [500, 1000] GeV

- Signals:
  $H \to c\bar{c}$, $H \to gg$
  $W \to q\bar{q}$, $Z \to q\bar{q}$, $t \to bq\bar{q}$



Performance QAE for different signals



Comparison with CAE

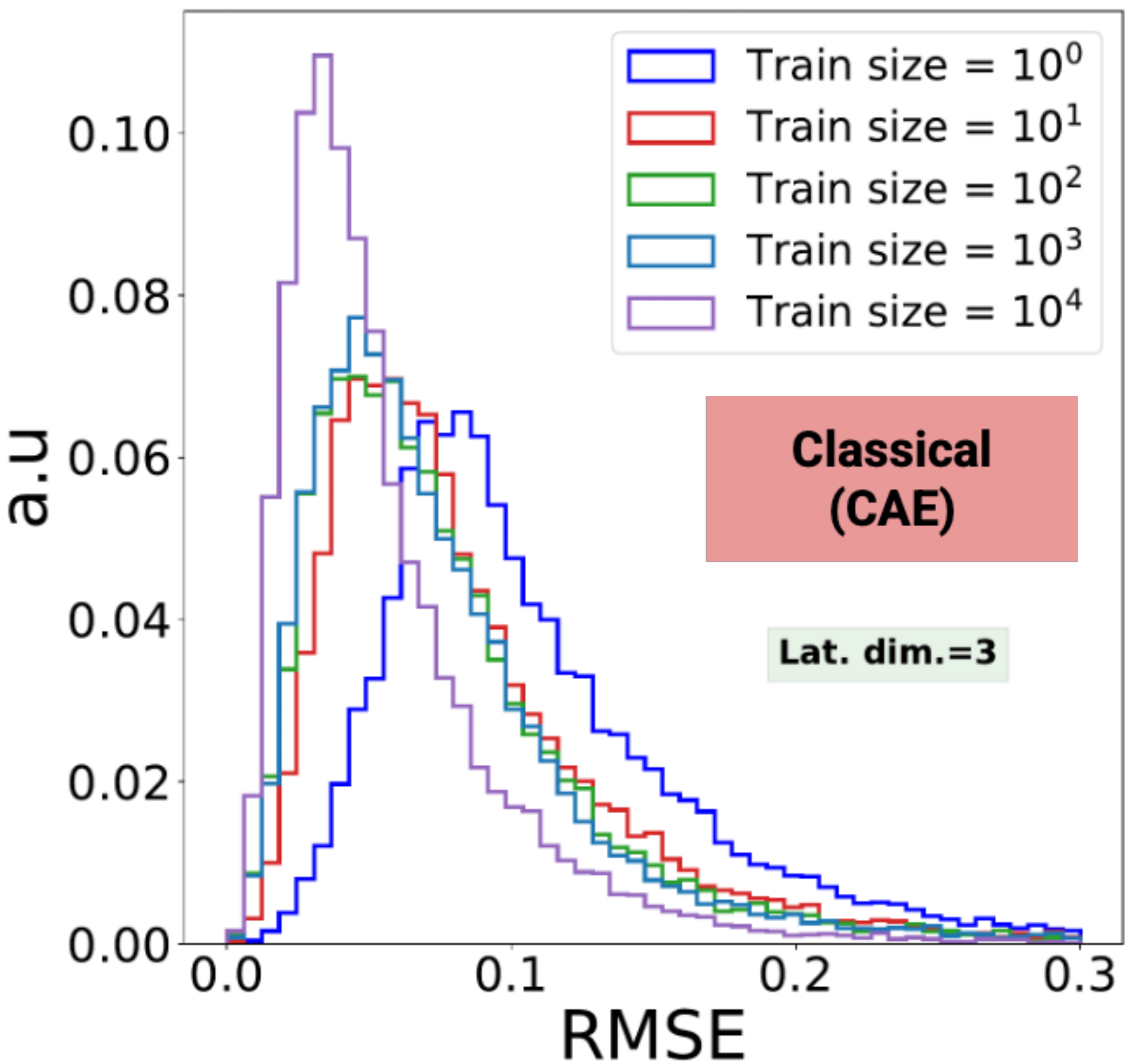
CAE: 30-20-16-12-6

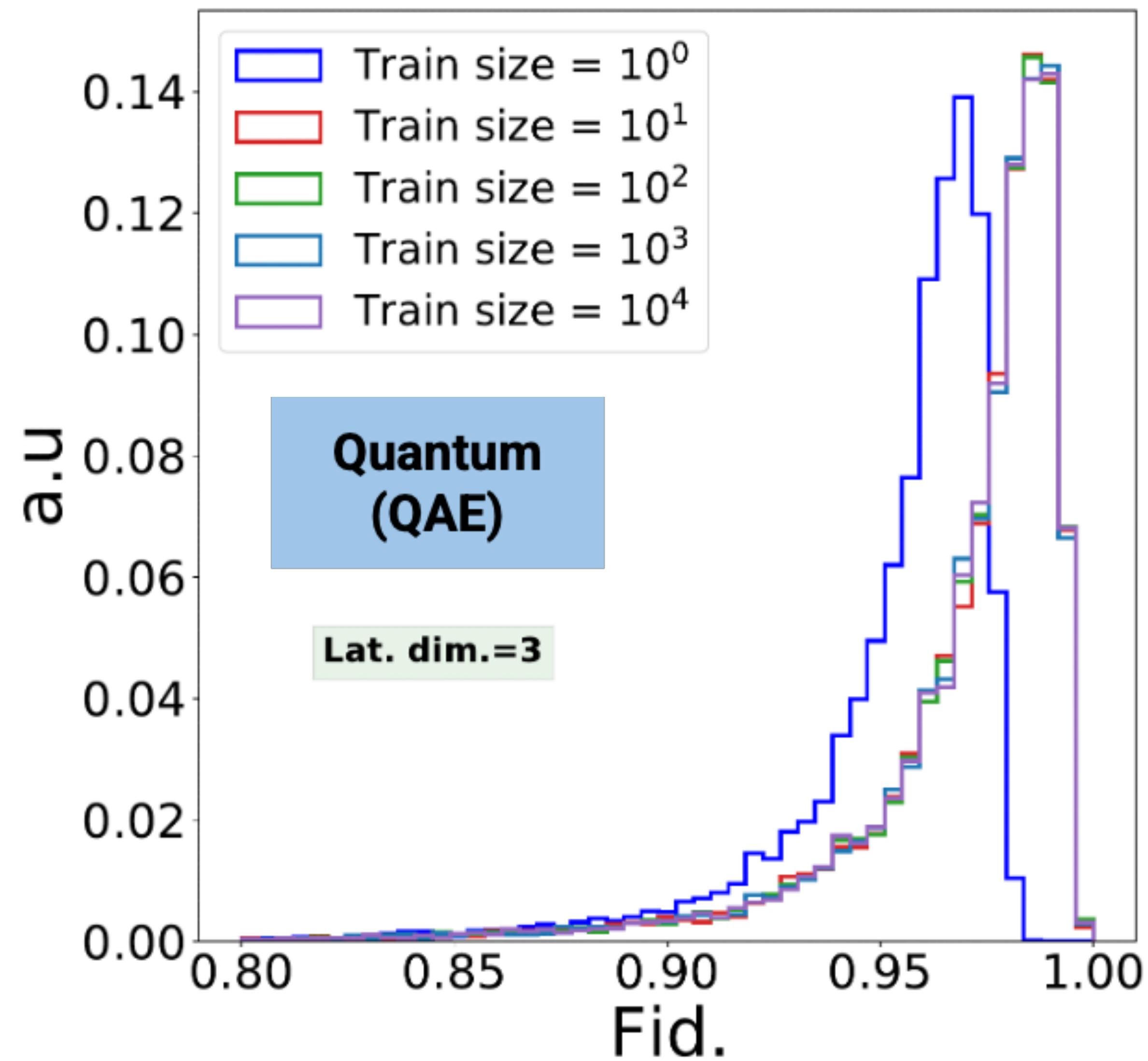
latent space

Algorithm \ Signals	$W \to q\bar{q}$	$H \to b\bar{b}$	$t \to bq\overline{q'}$
QAE	0.693	0.757	0.861
CAE	0.671	0.739	0.858

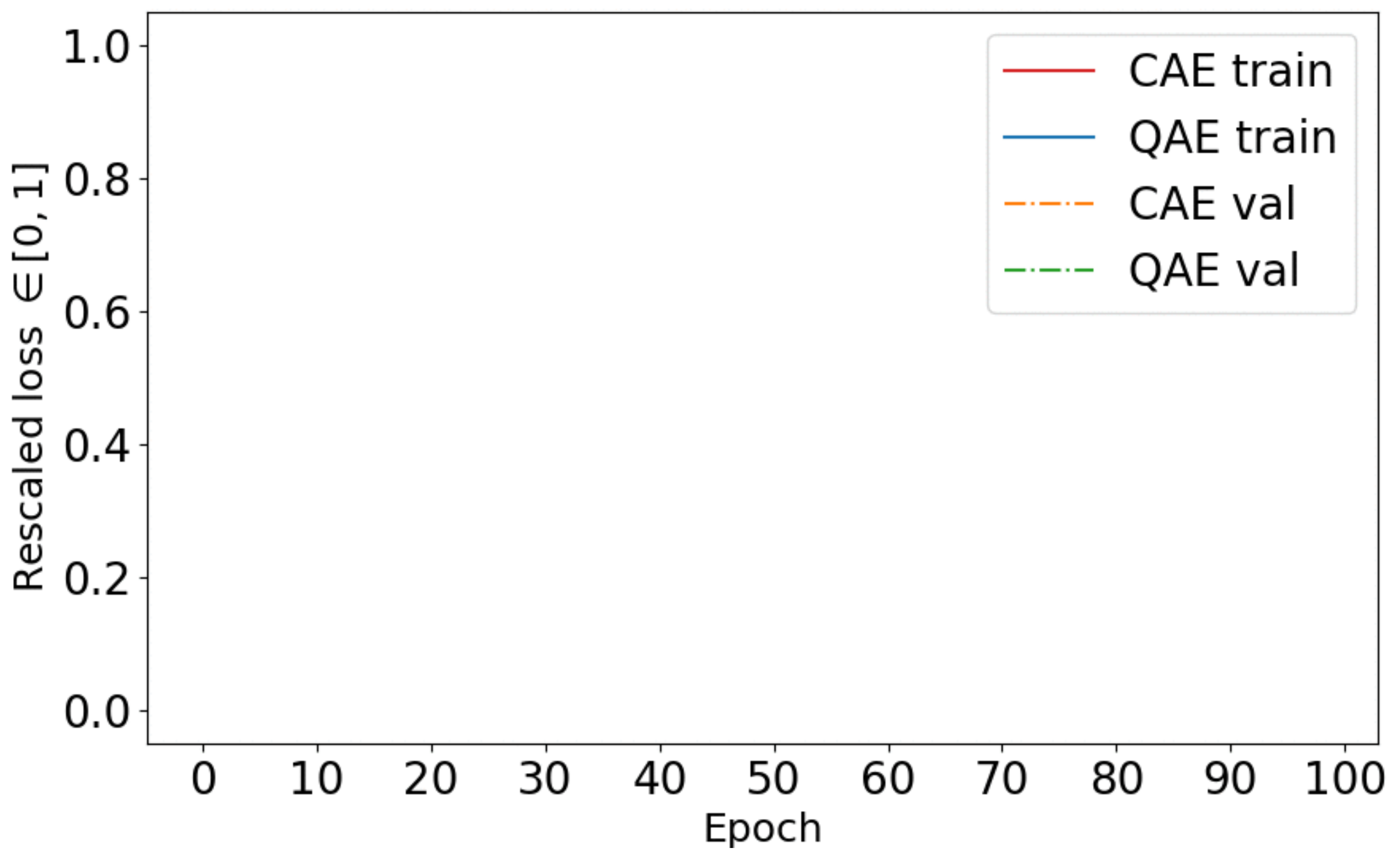
Results: Training size dependence

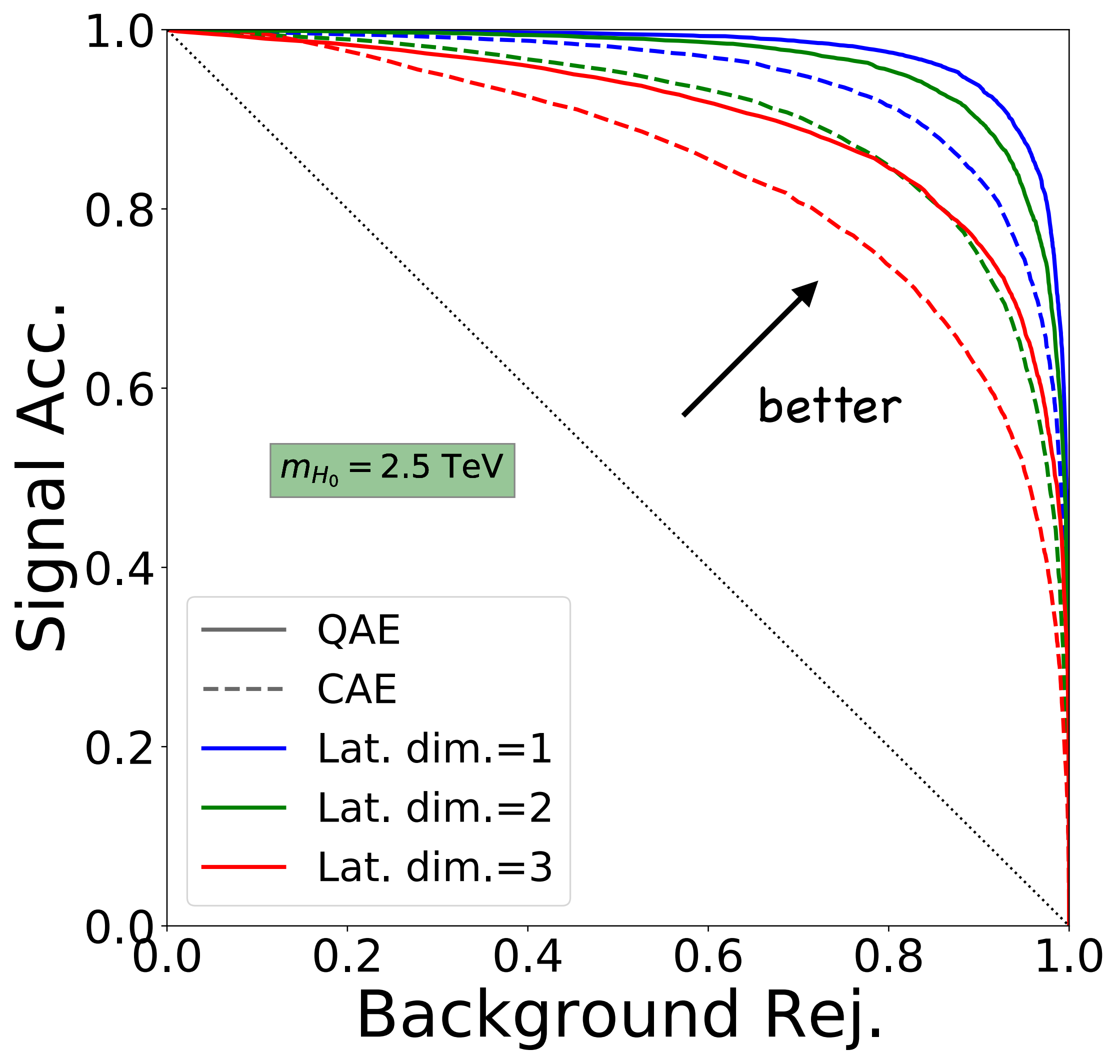
Dependence of (BG) test loss on training size





[Ngairangbam, MS, Takeuchi '21]

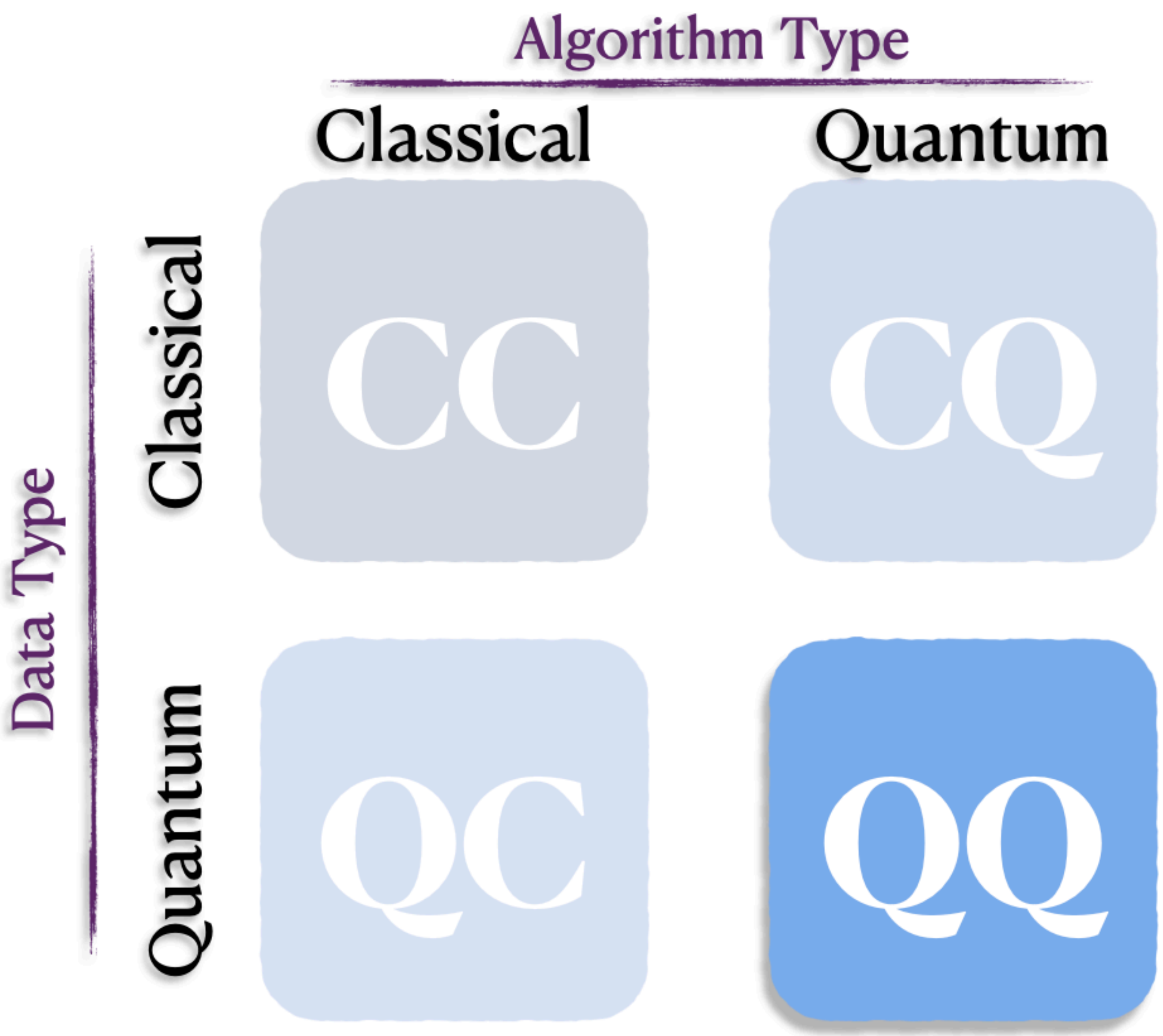




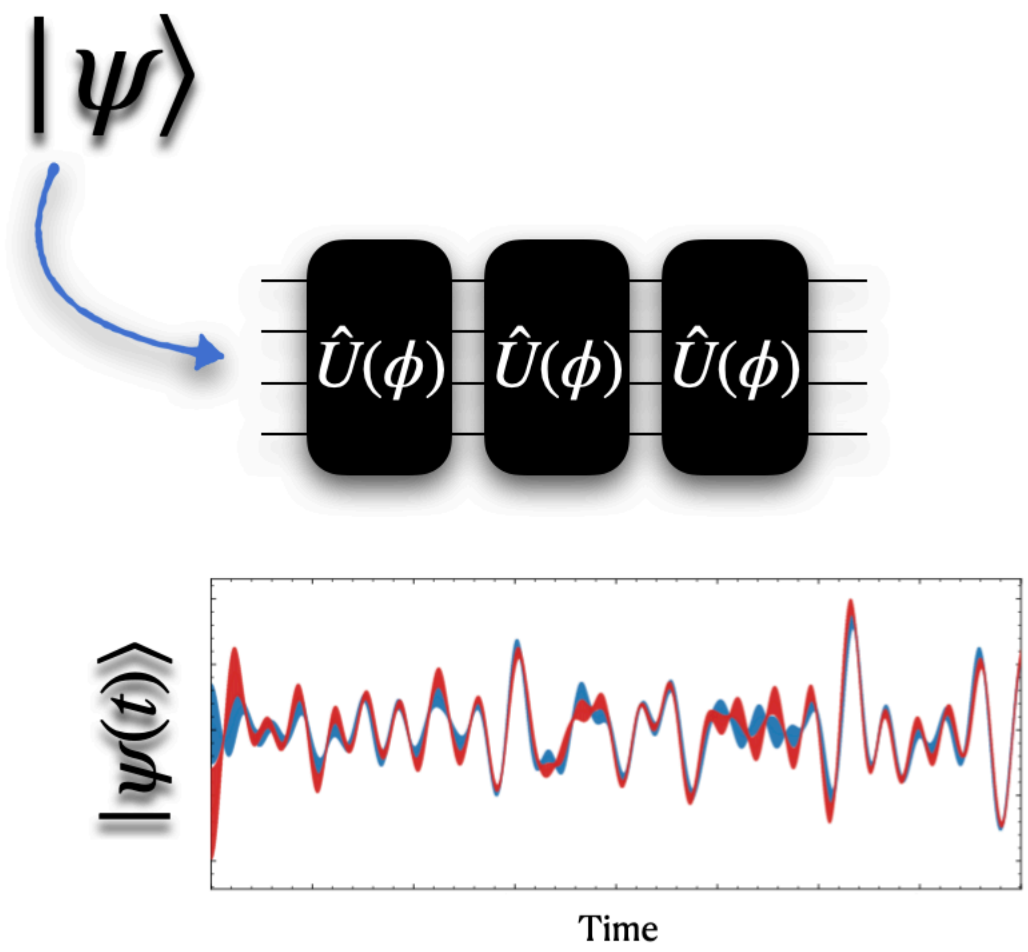
[Ngairangbam, MS, Takeuchi '21]

→ Much faster training and better performance for Quantum autoencoder

→ In our test cases QAE > CAE for much larger classical networks

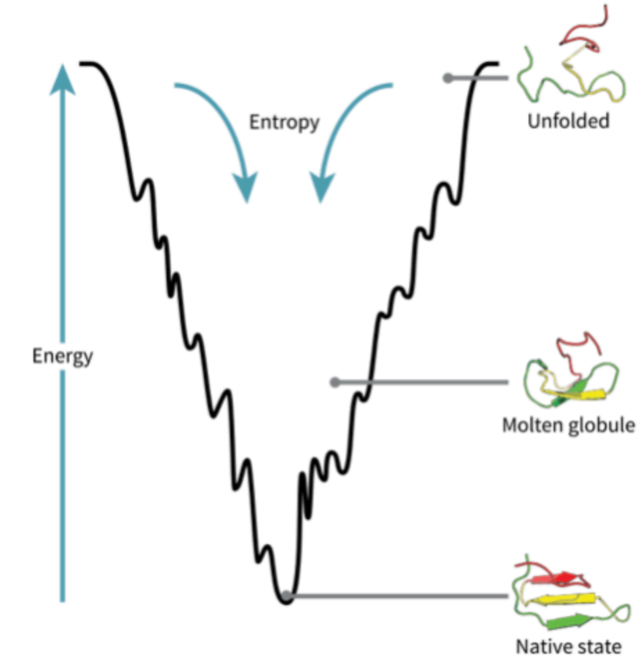


Algorithm Type
Classical
Quantum
Data Type
Classical
Quantum
CC
CQ
QC
QQ



$|\psi\rangle$
$\hat{U}(\phi)$
$\hat{U}(\phi)$
$\hat{U}(\phi)$
$|\psi(t)\rangle$
Time

Quantum annealing: Non-universal but universally powerful?



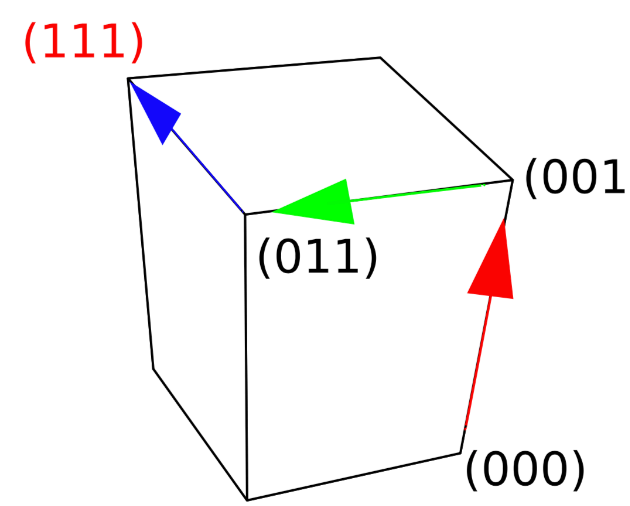
- Specific Hamiltonian. What does the "anneal" mean?

$$\mathcal{H}_{\mathrm{QA}}(t) = \sum_i \sum_j J_{ij}\sigma_i^Z \sigma_j^Z + \sum_i h_i \sigma_i^Z + \Delta(t) \sum_i \sigma_i^X$$

final Hamiltonian (encodes actual problem)

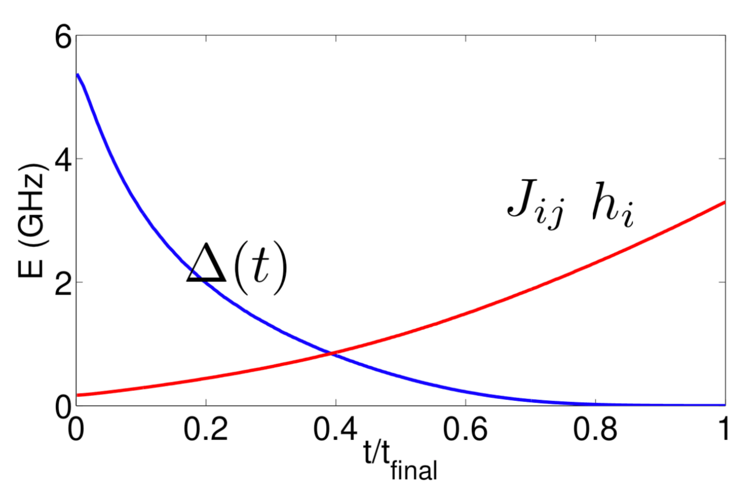
initial Hamiltonian (ground state = superposition of qubits with 0 and 1)

$\Delta(t)$ induces bit-hopping in the Hamming/Hilbert space



- Anneal idea: transition from ground state of initial Hamiltonian into ground state of problem Hamiltonian

- The idea is to dial this parameter to land in the global minimum (i.e. the solution) of some "problem space" described by J, h:

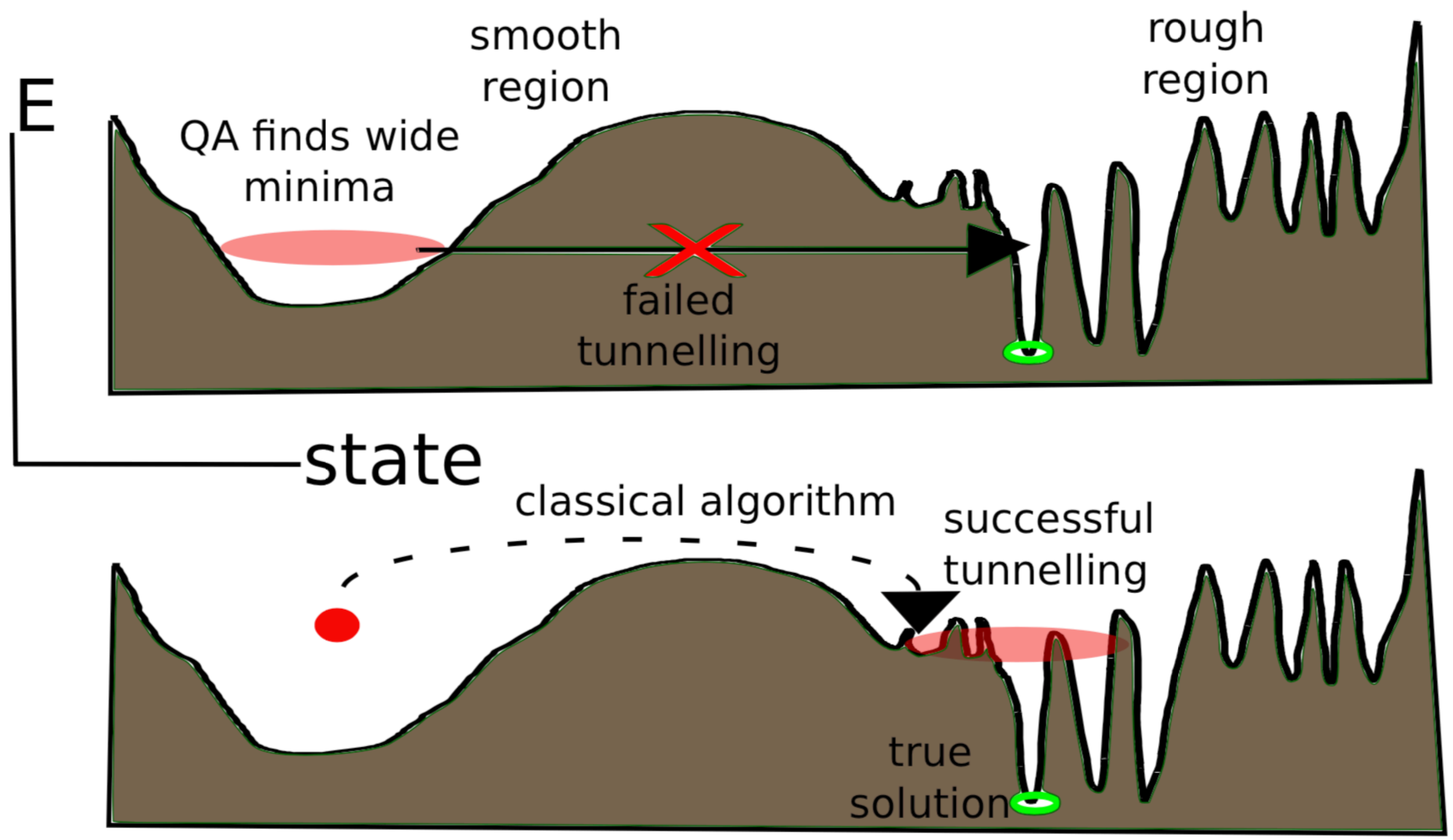


Thermal (classical) and Quantum Annealing are complementary:

- Thermal tunnelling is fast over broad shallow potentials $\sim e^{-\text{height}/T}$
  (Quantum "tunnelling" is exponentially slow)

- Quantum tunnelling is fast through tall thin potentials $\sim e^{-\sqrt{\text{height}}\times\text{width}/\hbar}$
  (Thermal "tunnelling" is exponentially slow - Boltzmann suppression)

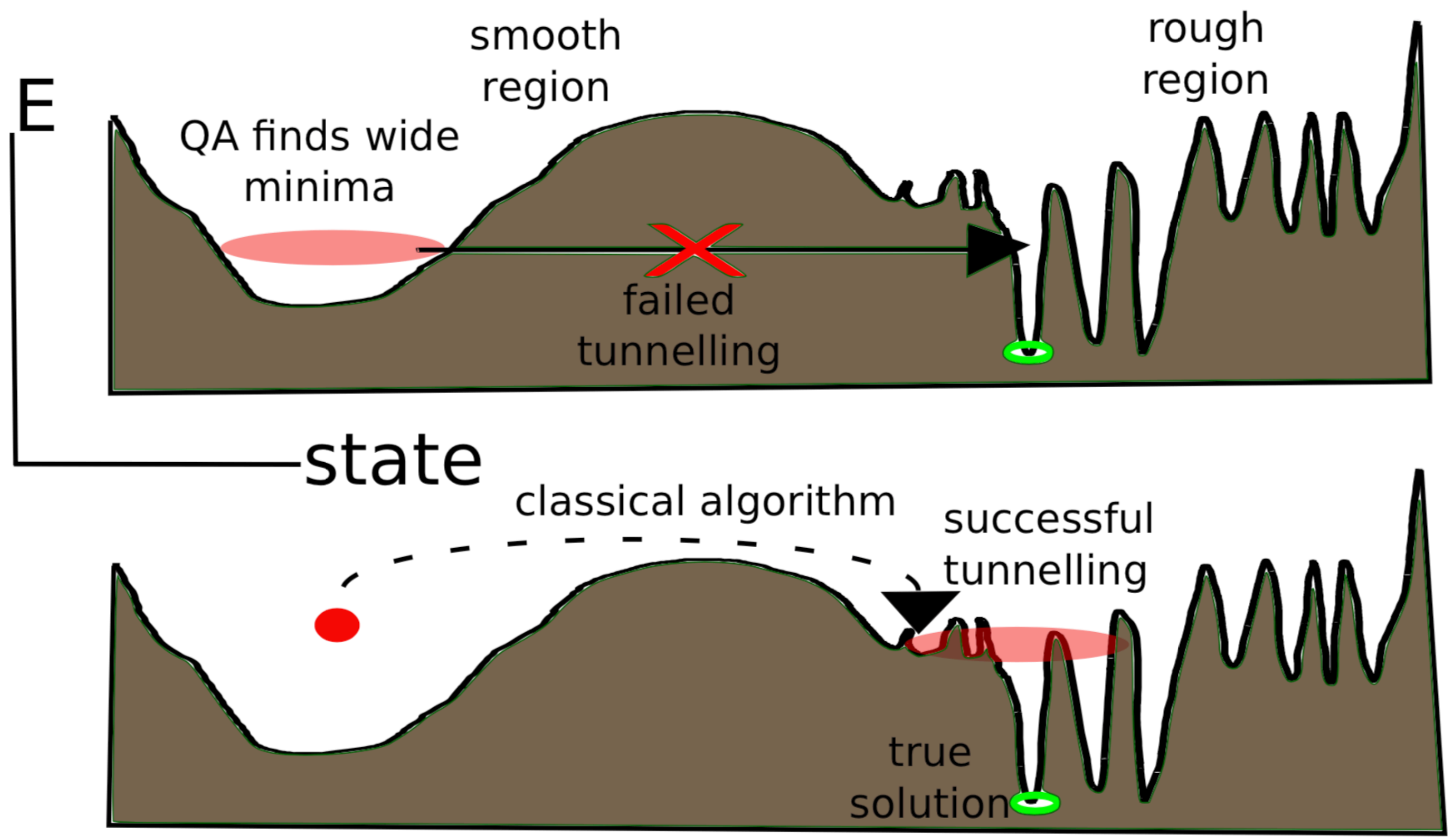
- Hybrid approach can be useful depending on solution landscape



- More specifically - thermal annealing uses Metropolis algorithm: accept random σ_i^Z flips with probability

$$P = \begin{cases} 1 & \Delta H \leq 0 \\ e^{-\Delta H/KT} & \Delta H > 0 \end{cases}$$

- Quantum tunnelling in QFT happens with probability $P \sim e^{-w\sqrt{2m\Delta H}/\hbar}$ so by contrast it can be operative for tall barriers if they are thin



A quantum laboratory for QFT and QML

- going beyond the reach of classical computers -

- Using the spin-chain approach for field theories discussed before, we can encode a QFT on a quantum annealer and study its dynamics directly.

[Abel, MS '20]

- To show that the system is a true and genuine quantum system we investigate if the state can tunnel from a meta-stable vacuum into a the true vacuum.

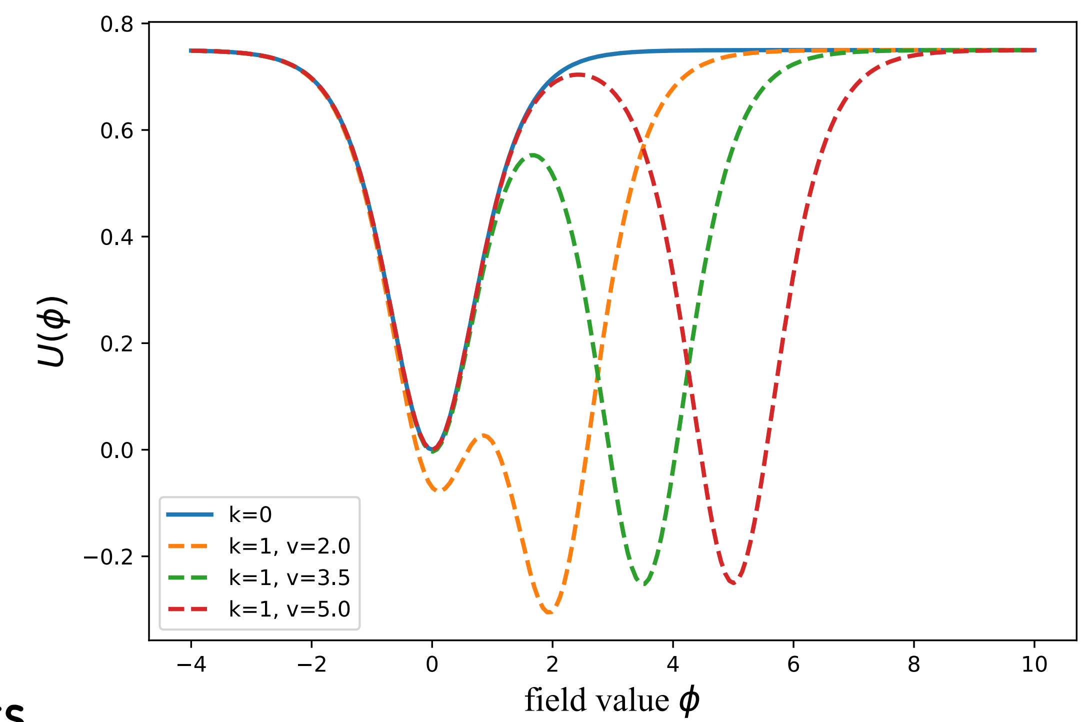
- Choose a potential of interest:

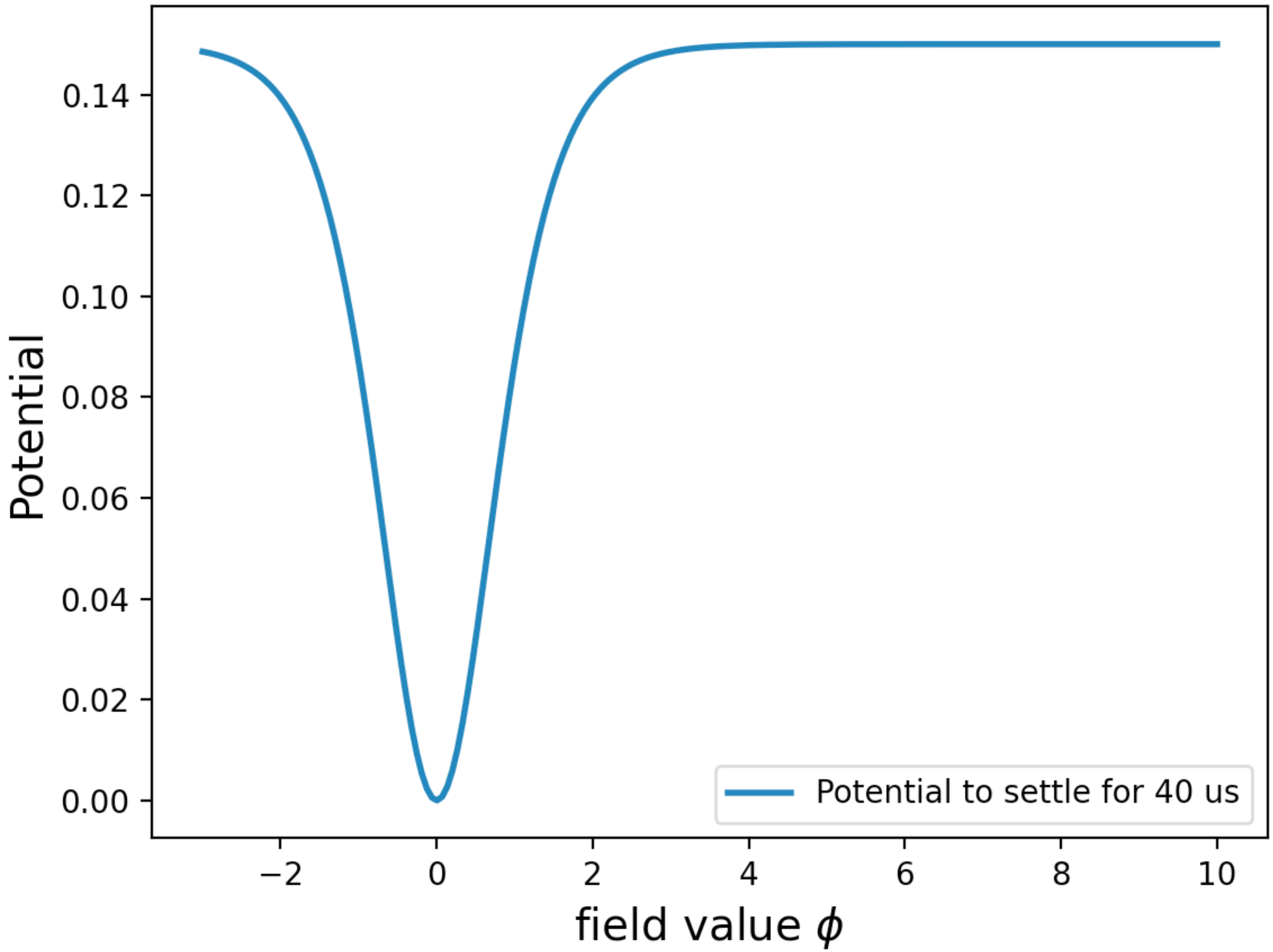
$$U(\phi) = \frac{3}{4}\tanh^2\phi - k(t)\,\mathrm{sech}^2\left(c(\phi - v)\right)$$

time dependent

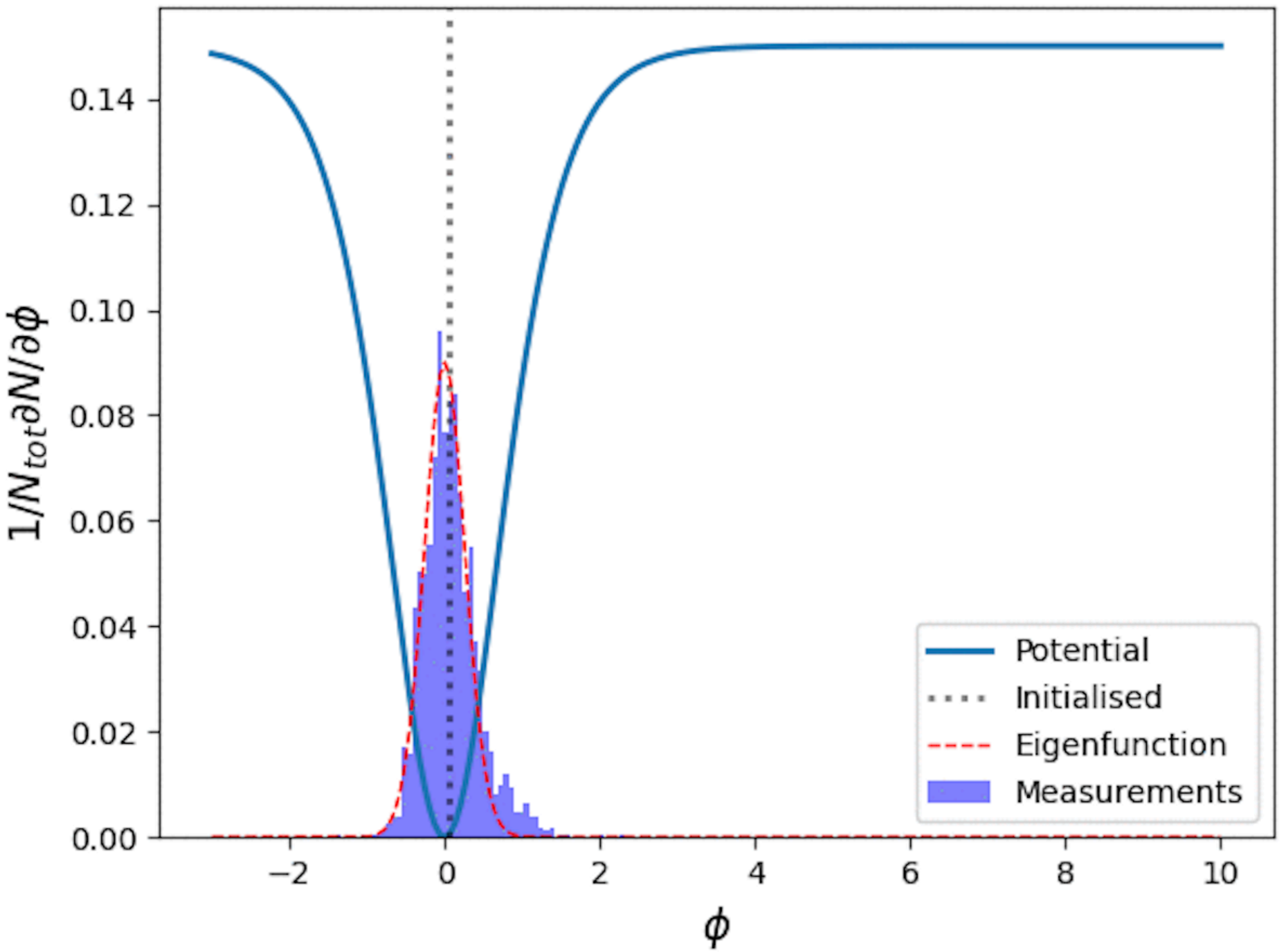
where $\phi = \eta/\eta_0$

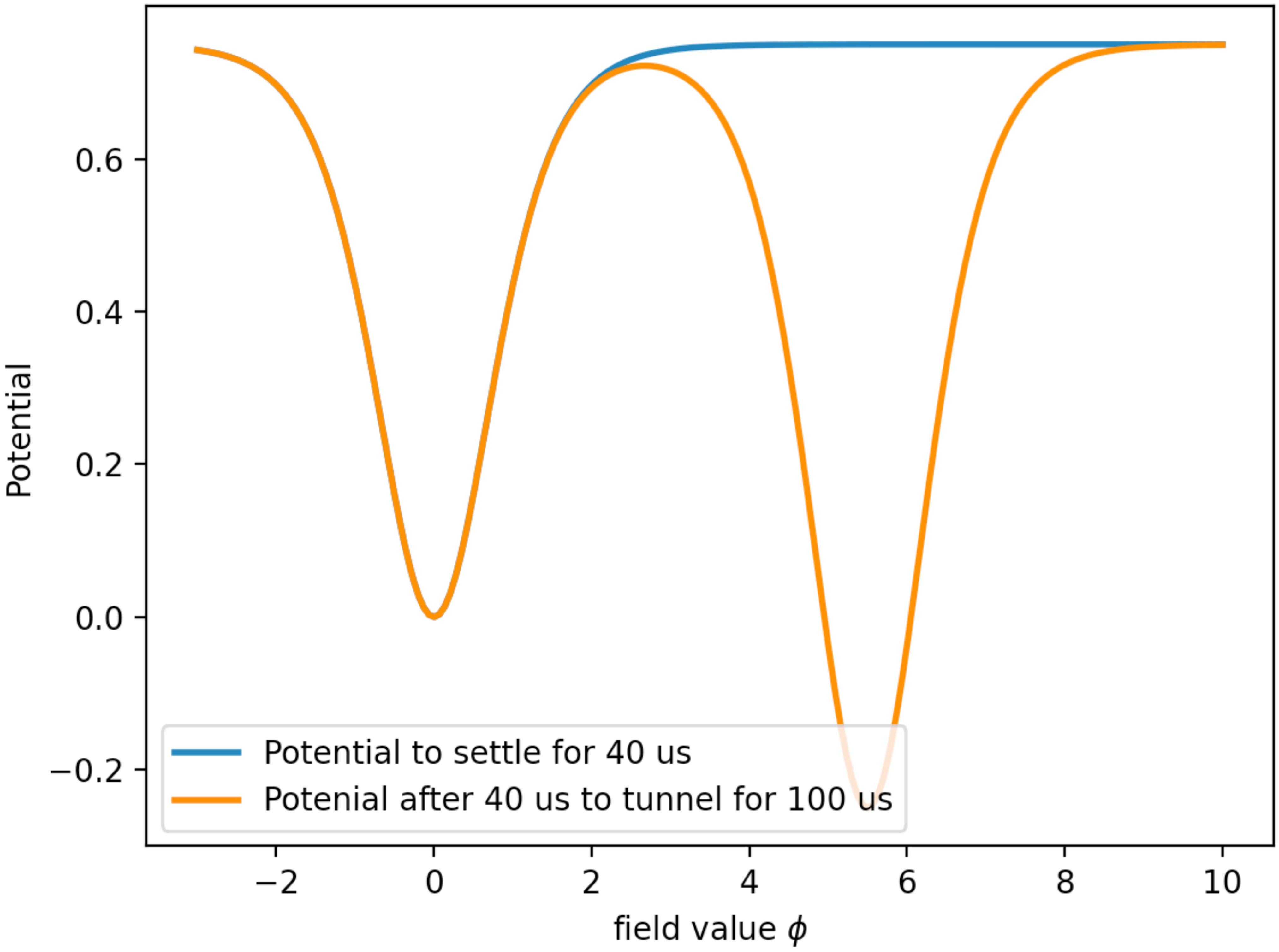
$\phi(t)$ is the field and c, v are dimless constants




0.14
0.12
0.10
0.08
0.06
0.04
0.02
0.00
Potential
−2
0
2
4
6
8
10
field value ϕ
Potential to settle for 40 us

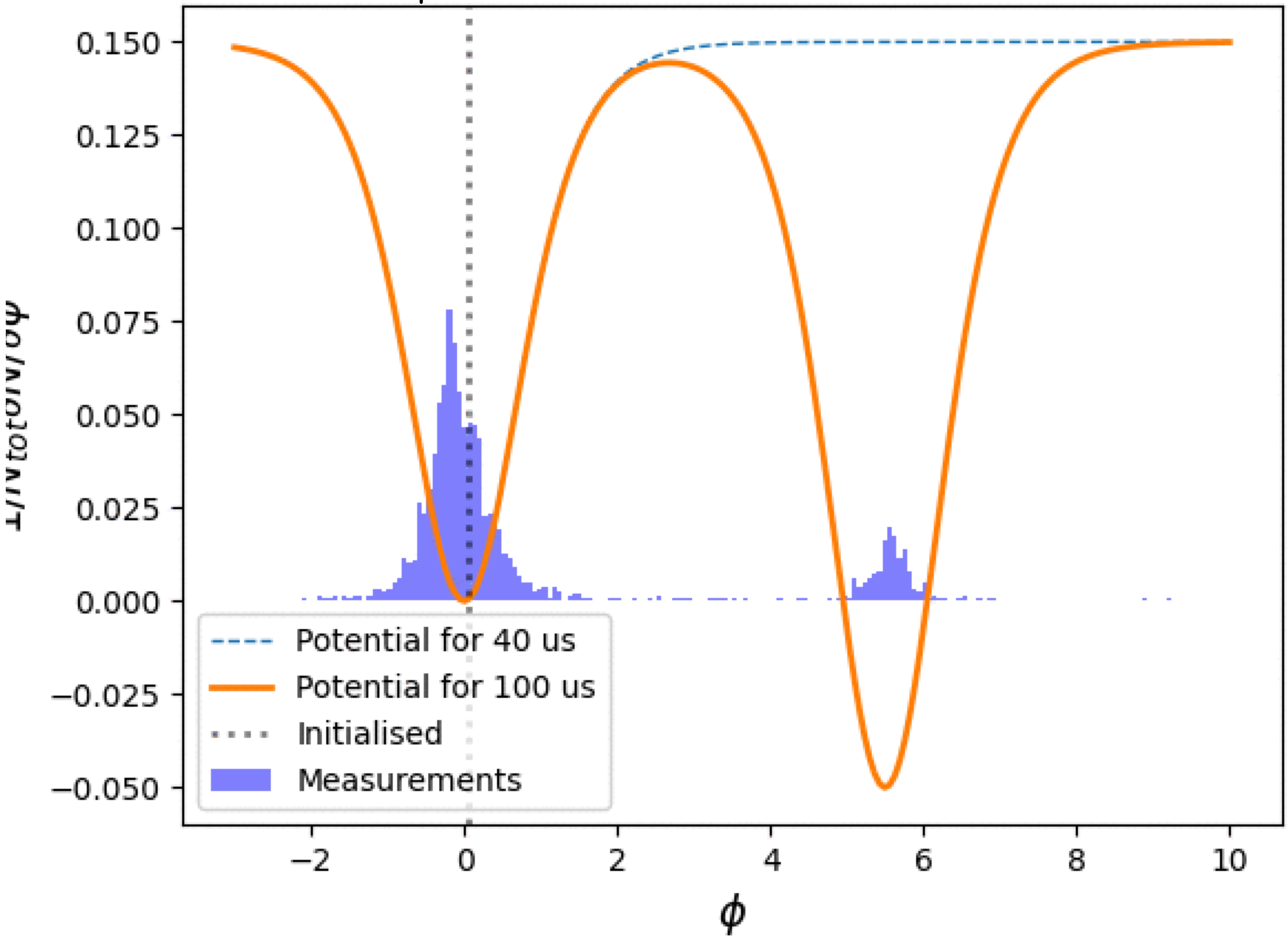
Implemented and executed on D-Wave Q2000 machine





0.6
0.4
0.2
0.0
−0.2
Potential
−2
0
2
4
6
8
10
field value ϕ
Potential to settle for 40 us
Potenial after 40 us to tunnel for 100 us

Implemented and executed on D-Wave Q2000 machine



Calibrating the system with a simple harmonic oscillator

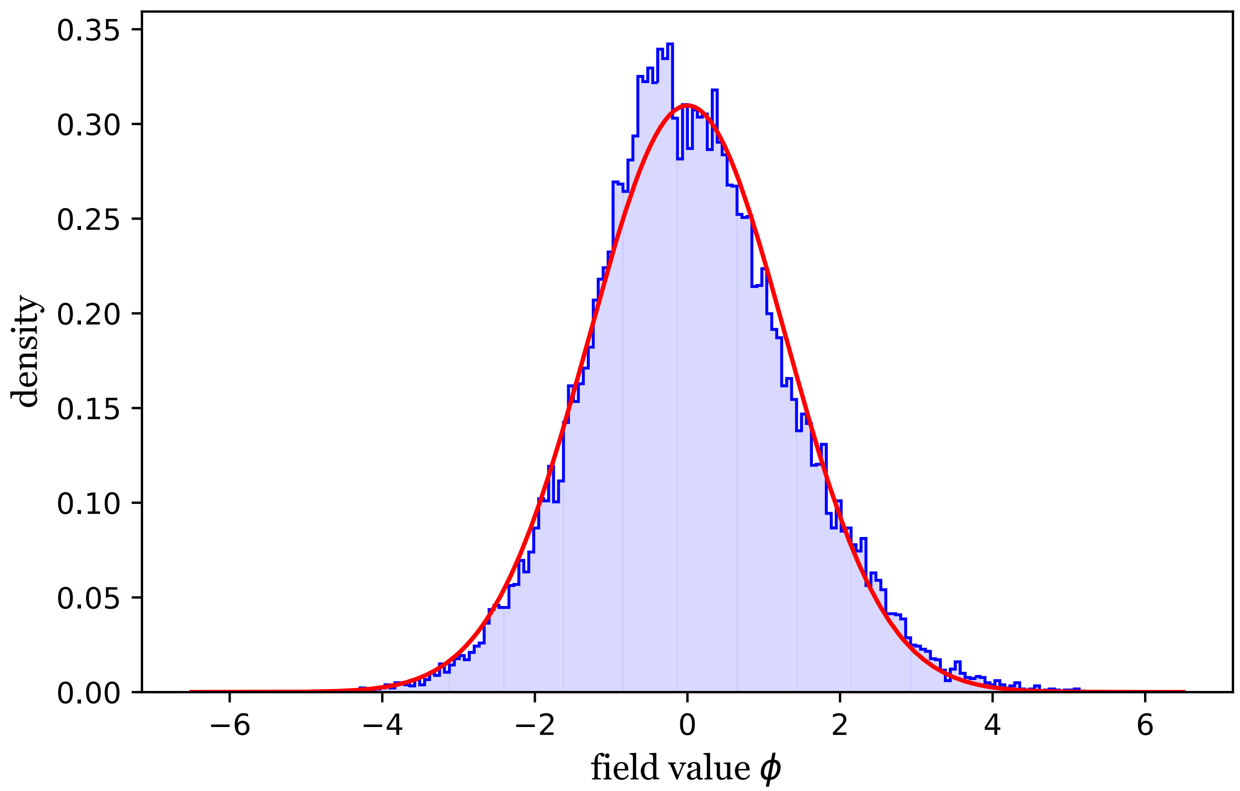
we assume for the potential $U_0(\phi) = \frac{\kappa}{2}\phi^2$ with $\kappa = 0.06$

we initialise classically at $\phi = 0$ and let it settle for $75\ \mu\mathrm{s}$ with $s_q = 0.7$

Fitted with Gaussian

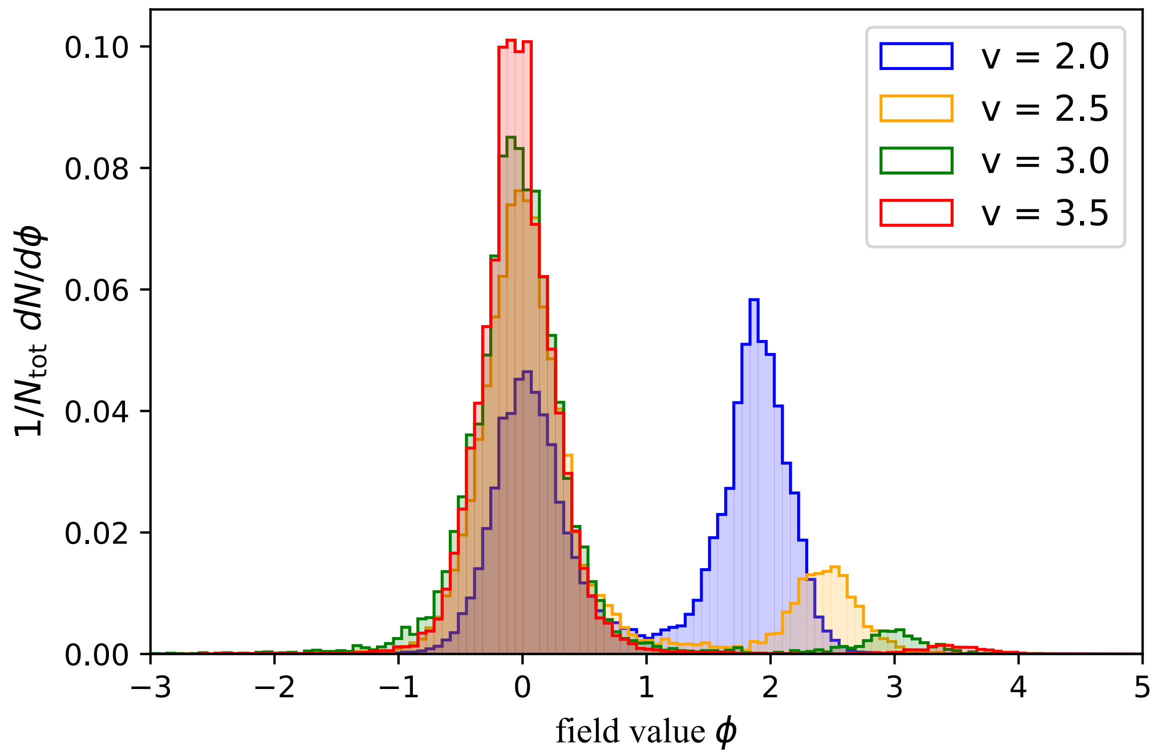
$$|\psi|^2 = \frac{(\kappa/2\gamma)^{\frac{1}{4}}}{\pi^{\frac{1}{2}}} e^{-\sqrt{\kappa/2\gamma}\,\phi^2} \longrightarrow \gamma = 0.33$$

No dynamics was induced by hand on annealer. Thus, constitutes a genuine measurement of the ground state wave function of a quantum mechanical system



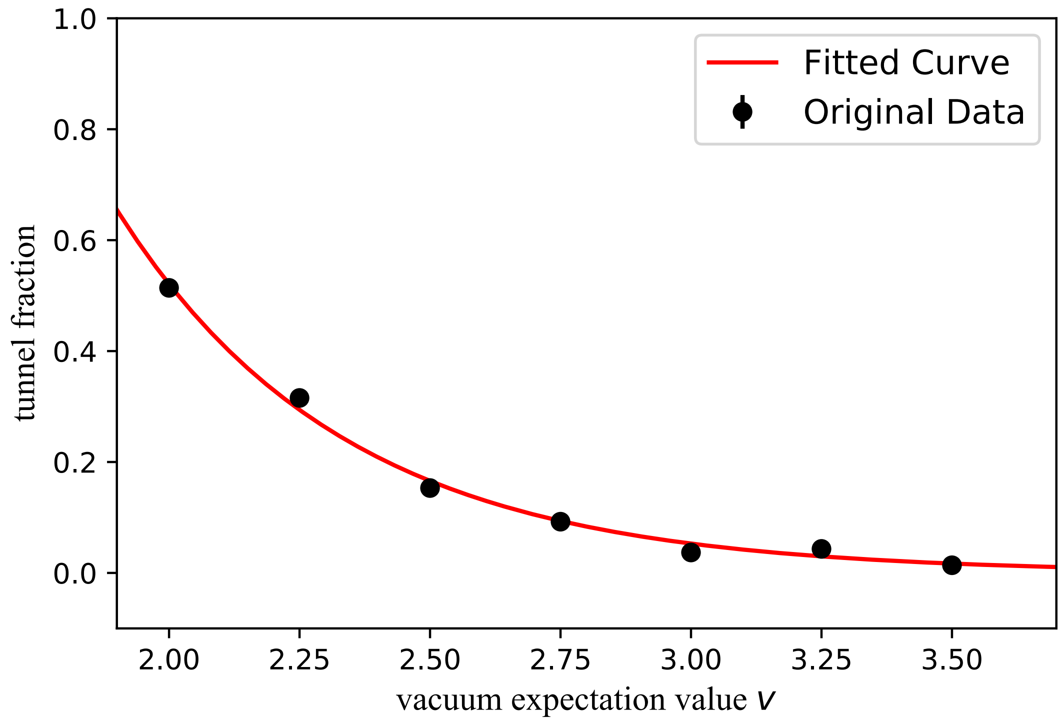
30k shots on D-Wave machine

Results: it decays with v as expected



Perform tunnelling for

$t_{\mathrm{tunnel}} = 100\mu s$ at $s_q = 0.7$

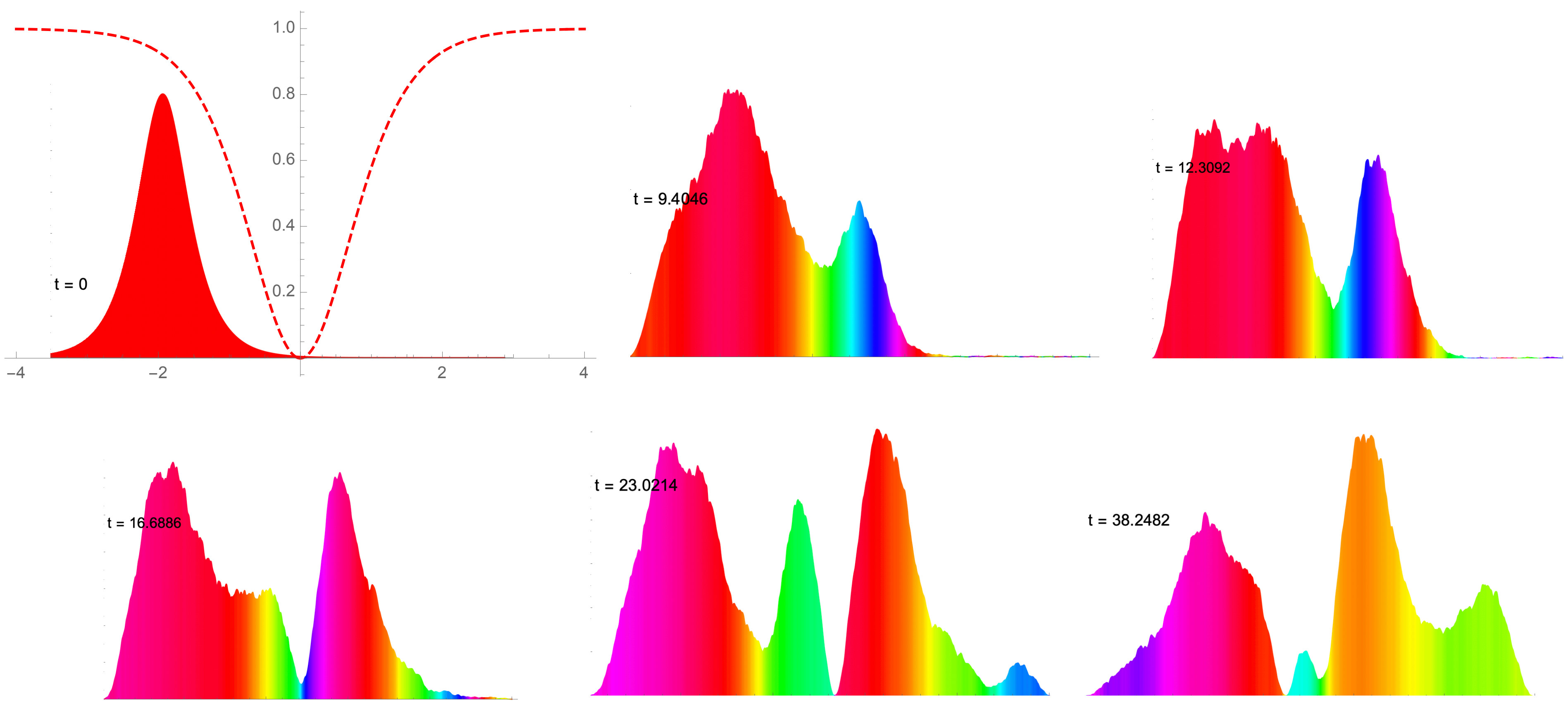


Theory: $\log\Gamma = 3.0 \times (1.66 - v)$

Exp: $\log\Gamma = 2.29 \times (1.71 - v)$

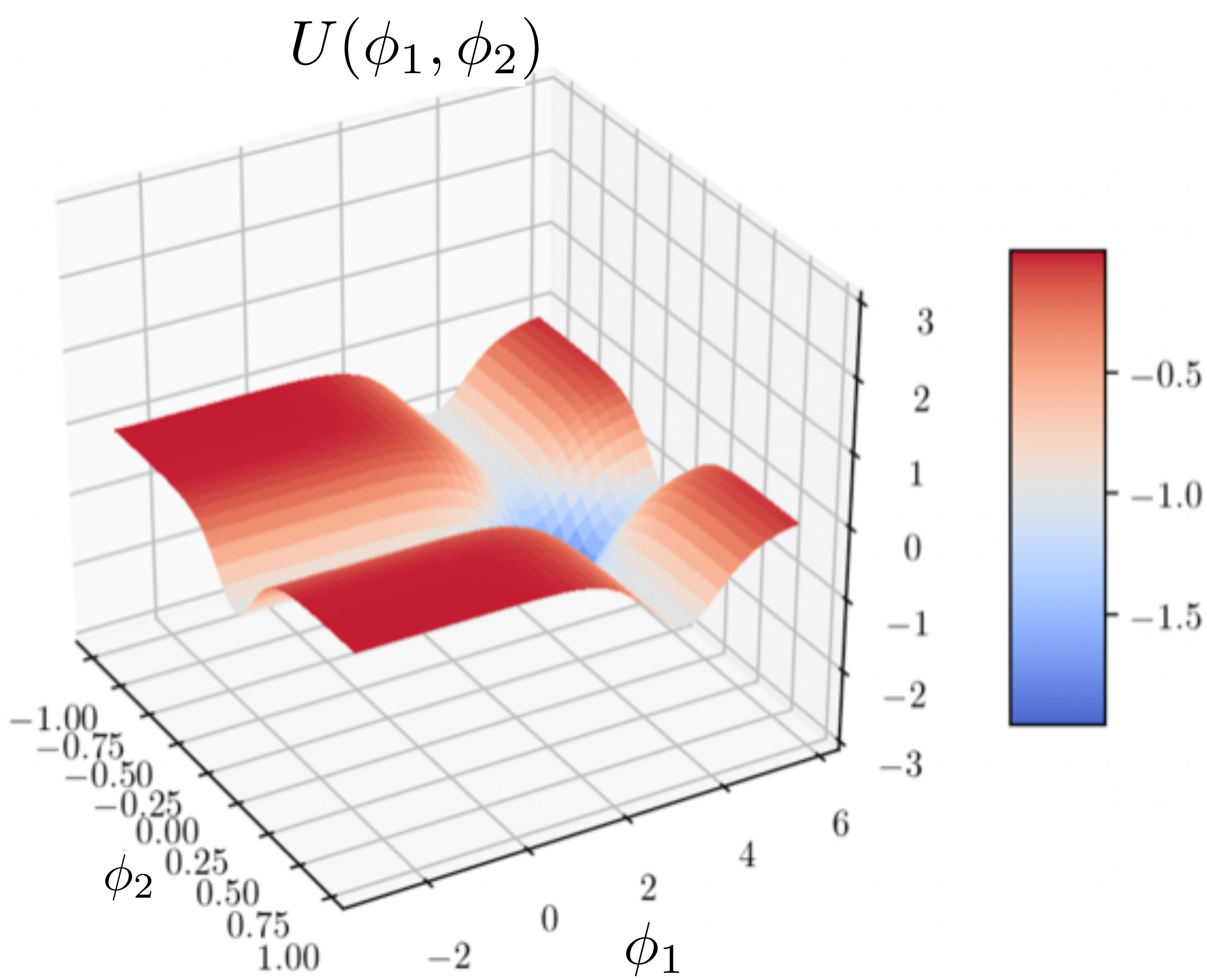
Also dynamics has characteristic behaviour. For example it still "tunnels" to the bottom of a potential even if there is no barrier: i.e. the wave function leaks across, rather than rolling as a lump —

Numerically solving S.E. we find (this takes an hour!)



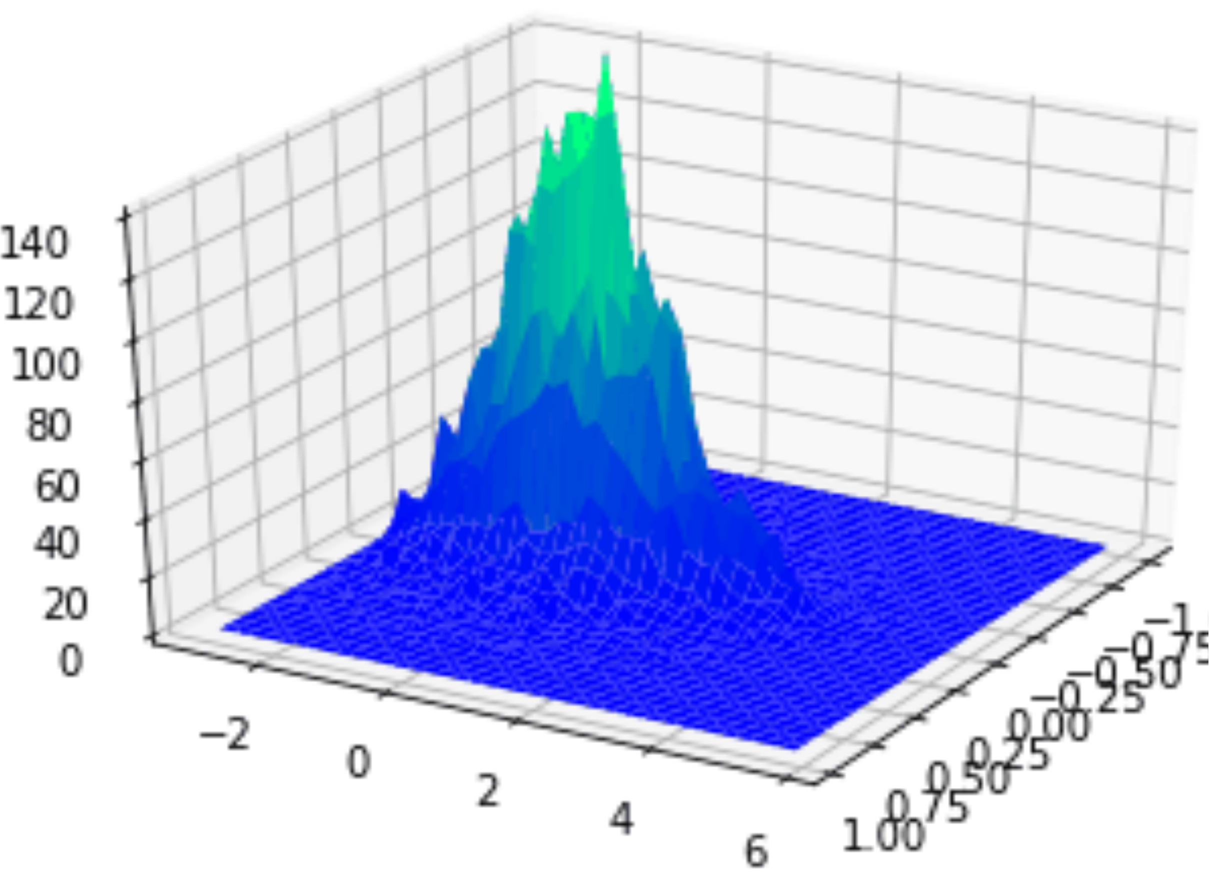
Also dynamics has characteristic behaviour. For example it still “tunnels” to the bottom of a potential even if there is no barrier: i.e. the wave function leaks across, rather than rolling as a lump —

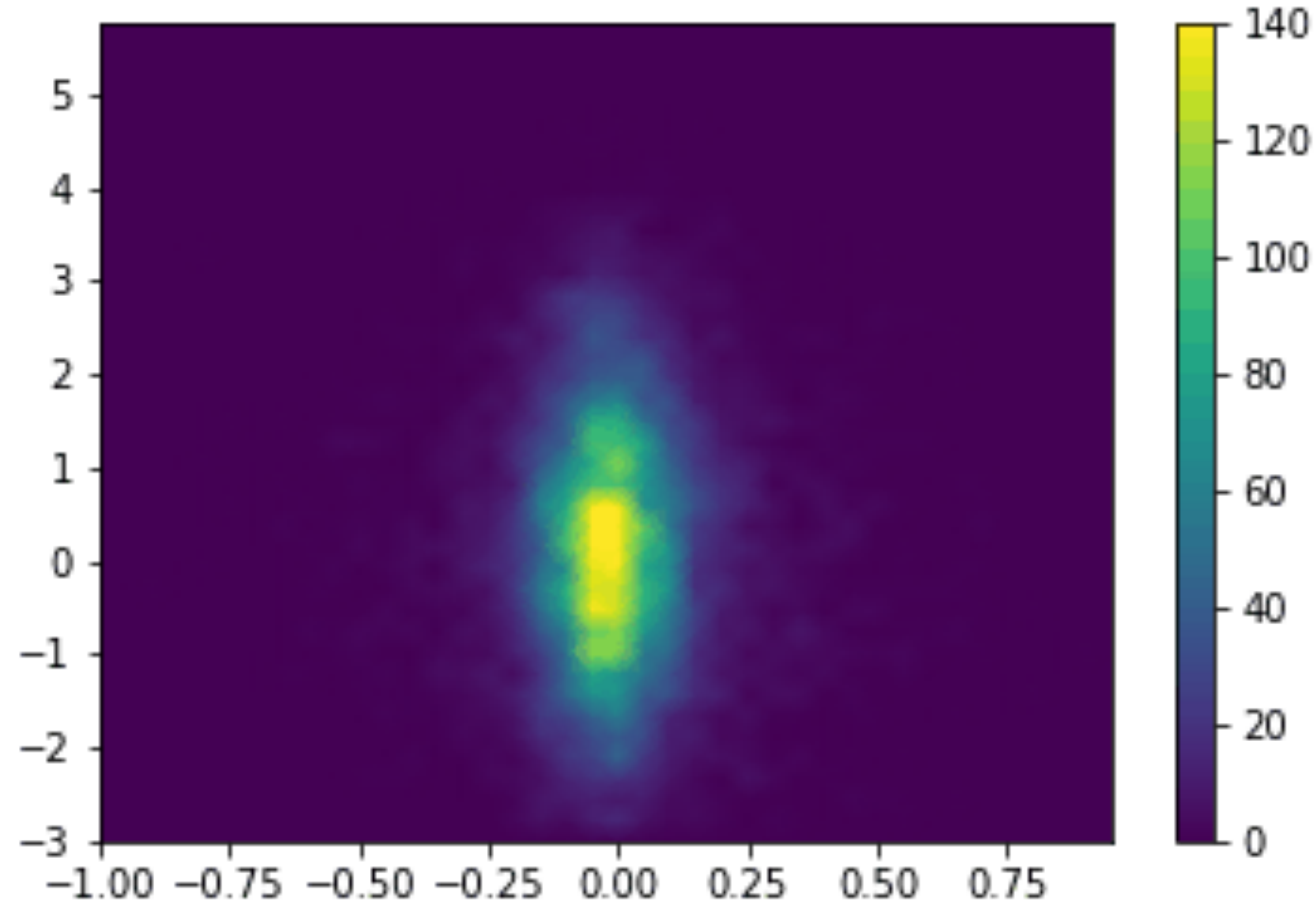
Multiple measurements on the quantum annealer:



Also dynamics has characteristic behaviour. For example it still "tunnels" to the bottom of a potential even if there is no barrier: i.e. the wave function leaks across, rather than rolling as a lump —

Multiple measurements on the quantum annealer:



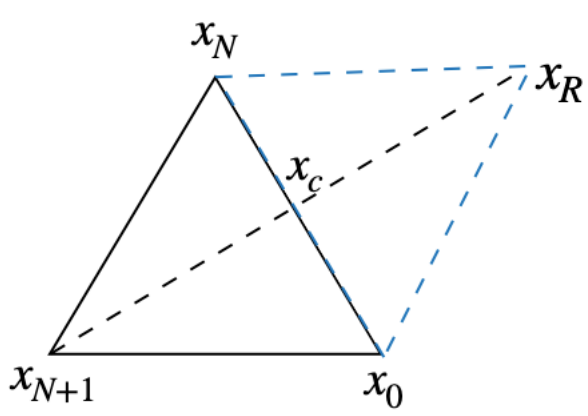


Example 2: Optimisation comparison quantum vs classical

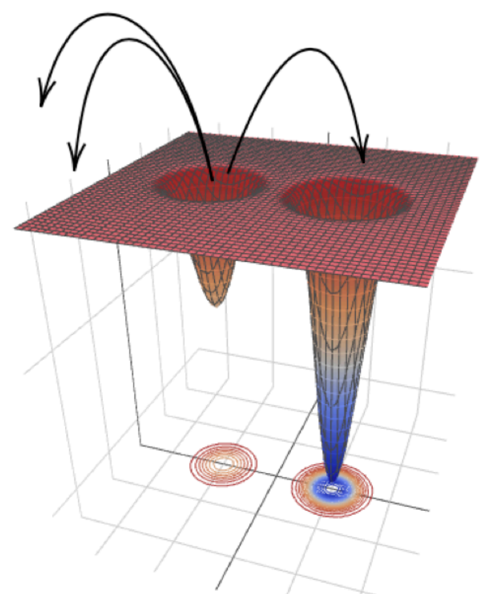
gradient descent

$$x_{i+1} = x_i - \nabla f(x_i)$$

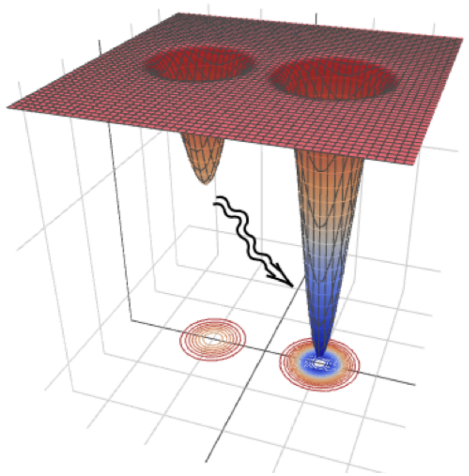
Nelder-Mead



Thermal Annealing

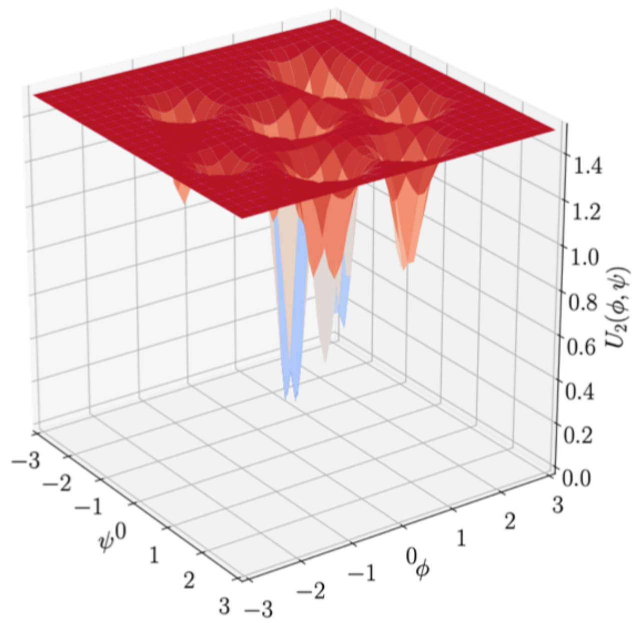


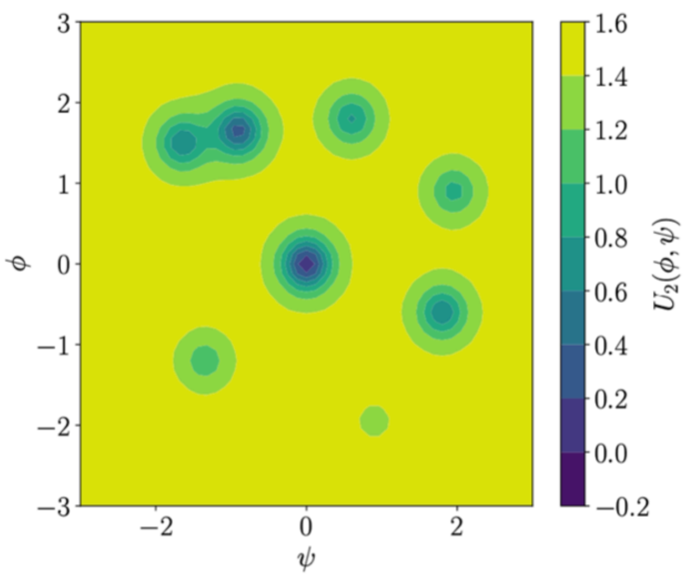
Quantum Annealing



Applied to several examples in [Abel, Blance, MS '21], let's show one here:

Multi-well potential



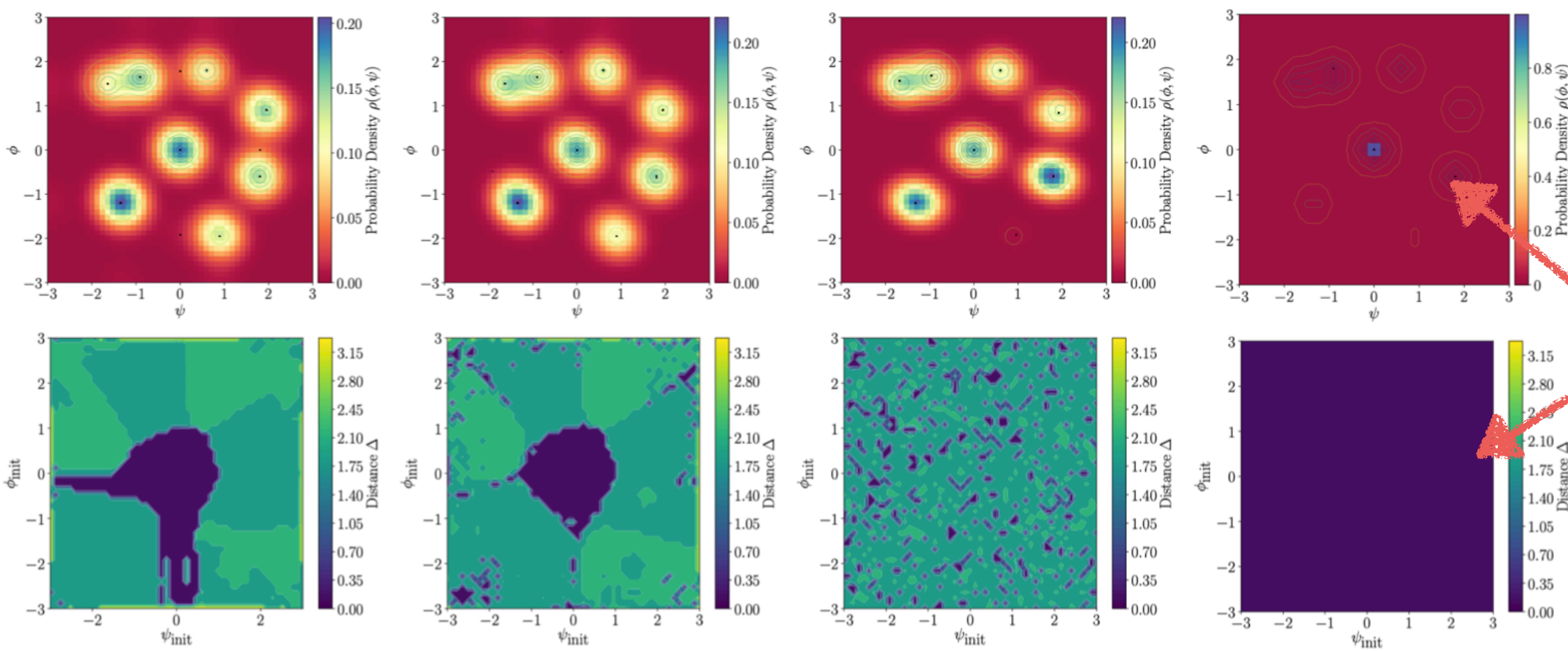


Results for Multi-well potential

[Abel, Blance, MS '21]

- Quantum algorithms finds global minimum of potential reliably and fast!

Method	Time/run (μs)
Nelder-Mead	4900
Gradient Descent	2900
Thermal Annealing	5×10^5
Quantum Annealing	115



(a) Nelder-Mead (b) Gradient descent (c) Thermal annealing (d) Quantum annealing

Quantum annealer almost never gets stuck in wrong minimum

QA is depth savvy, i.e. works qualitatively different

→ Clear quantum advantage

Completely Quantum Neural Networks

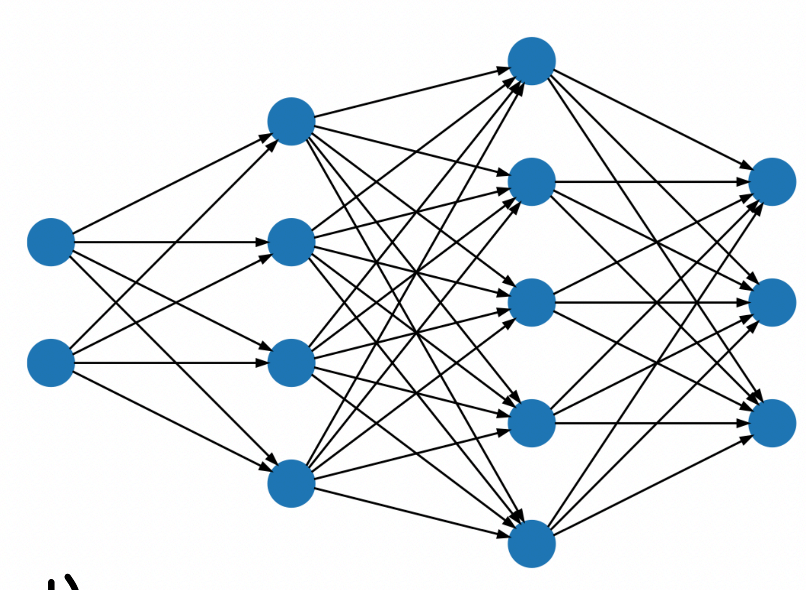
Structure of node i, in layer L $$L_i(x) = g\left(\sum_i w_{ij}x_i + b_i\right)$$

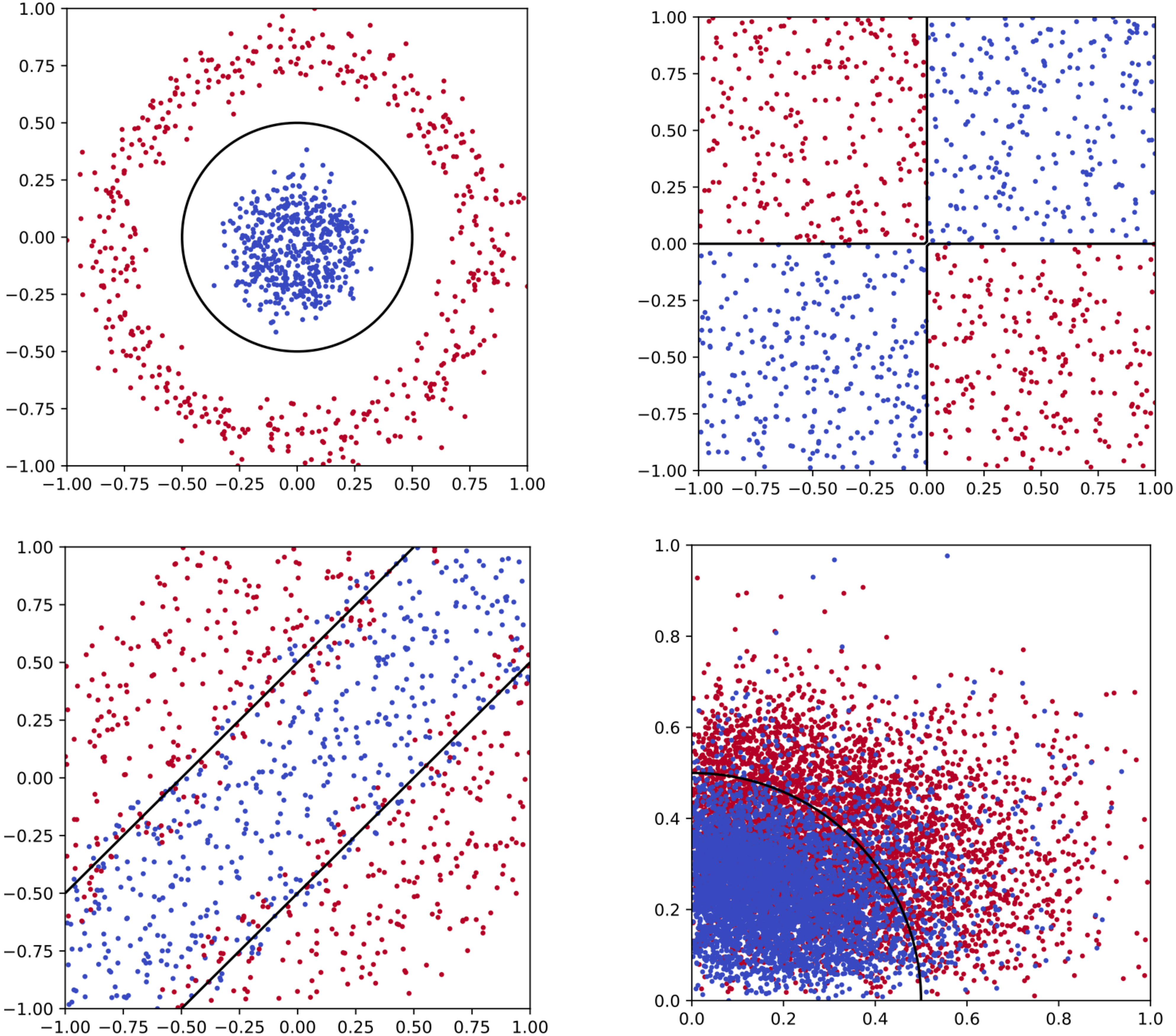
Network output in final layer $$Y = L^{(n)} \circ \ldots \circ L^{(0)}$$

Loss function $$\mathcal{L}(Y) = \frac{1}{N_d}\sum_a |y_a - Y(x_a)|^2$$

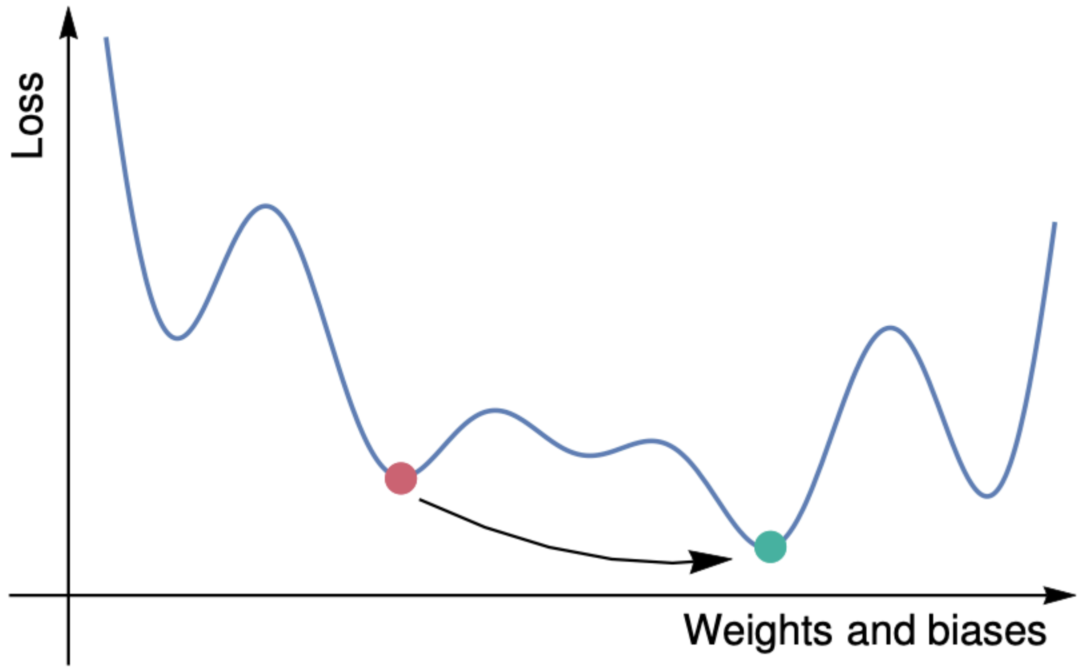
[Abel, Criado, MS '22]

- Developed binary encoding of weights (discretised)
- Polynomial approximation of activation function
- Reduction of binary higher-order polynomials into quadratic ones (Ising model)

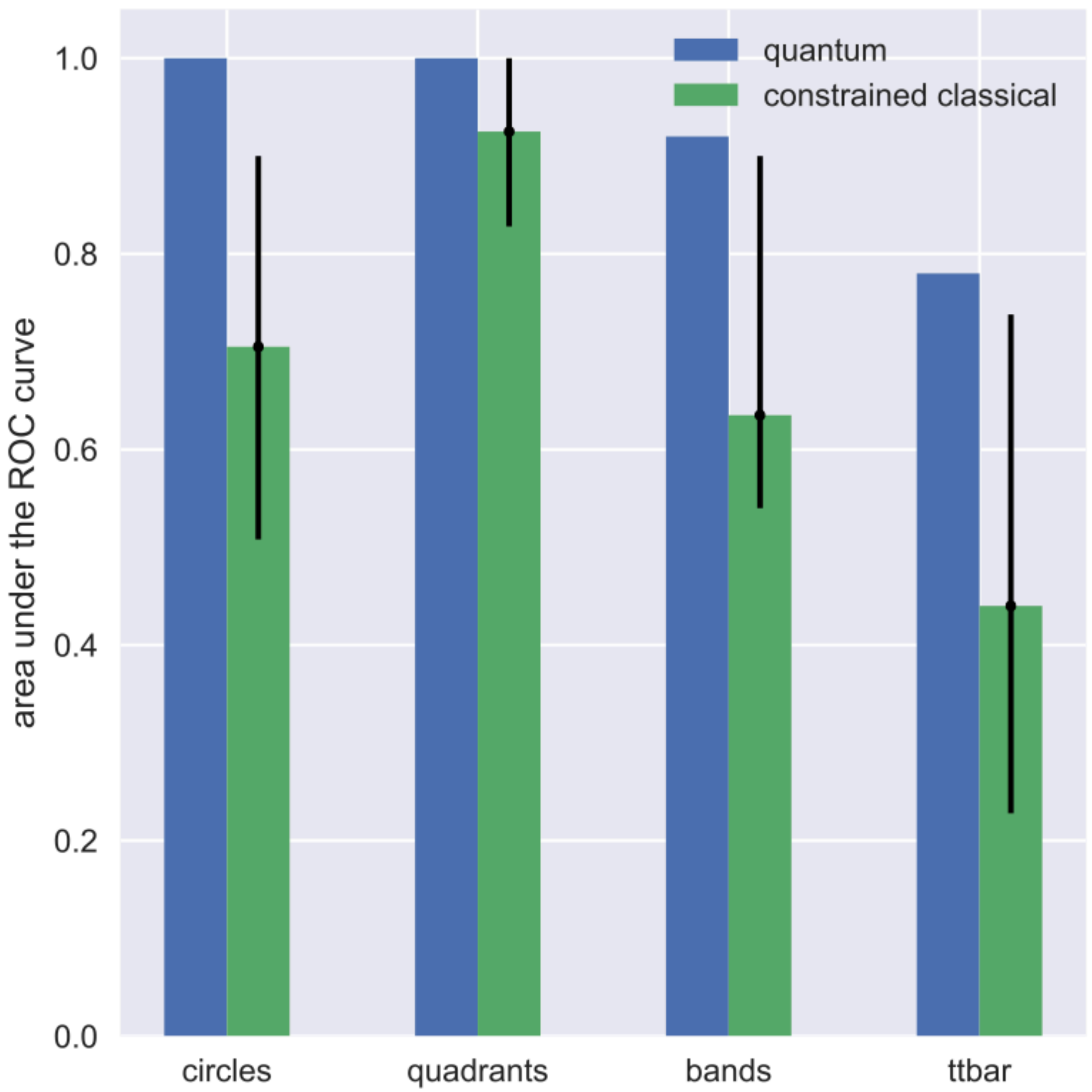



1.00
0.75
0.50
0.25
0.00
−0.25
−0.50
−0.75
−1.00
−1.00 −0.75 −0.50 −0.25 0.00 0.25 0.50 0.75 1.00
1.0
0.8
0.6
0.4
0.2
0.0
0.0 0.2 0.4 0.6 0.8 1.0

Completely Quantum Neural Networks



Reliable and very fast ground-state finder of loss function



Optimal network training

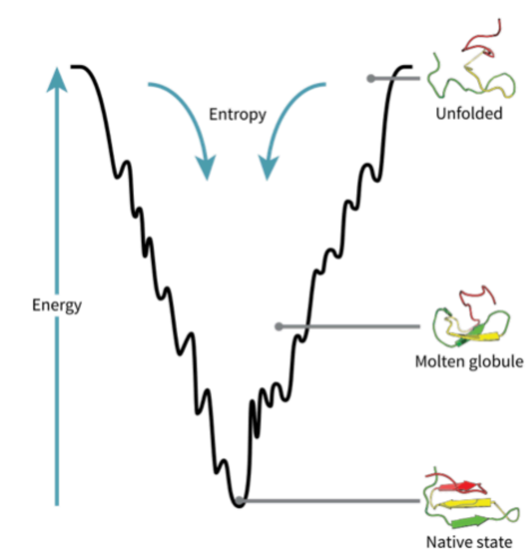
Application to differential equations and variational methods

Define your mathematical task as an **optimisation problem**

$$\mathcal{F}_m(\vec{x}, \phi_m(\vec{x}), \nabla \phi_m(\vec{x}), \cdots, \nabla^j \phi_m(\vec{x})) = 0$$

Build the full function, here a DE into the loss function, incl boundary conditions

$$\mathcal{L}(\{w, \vec{b}\}) = \frac{1}{i_{\max}} \sum_{i,m} \hat{\mathcal{F}}_m(\vec{x}^i, \hat{\phi}_m(\vec{x}^i), \cdots, \nabla^j \hat{\phi}_m(\vec{x}^i))^2 + \sum_{\text{B.C.}} (\nabla^p \hat{\phi}_m(\vec{x}_b) - K(\vec{x}_b))^2 \ ,$$



[Piscopo, MS, Waite '19]

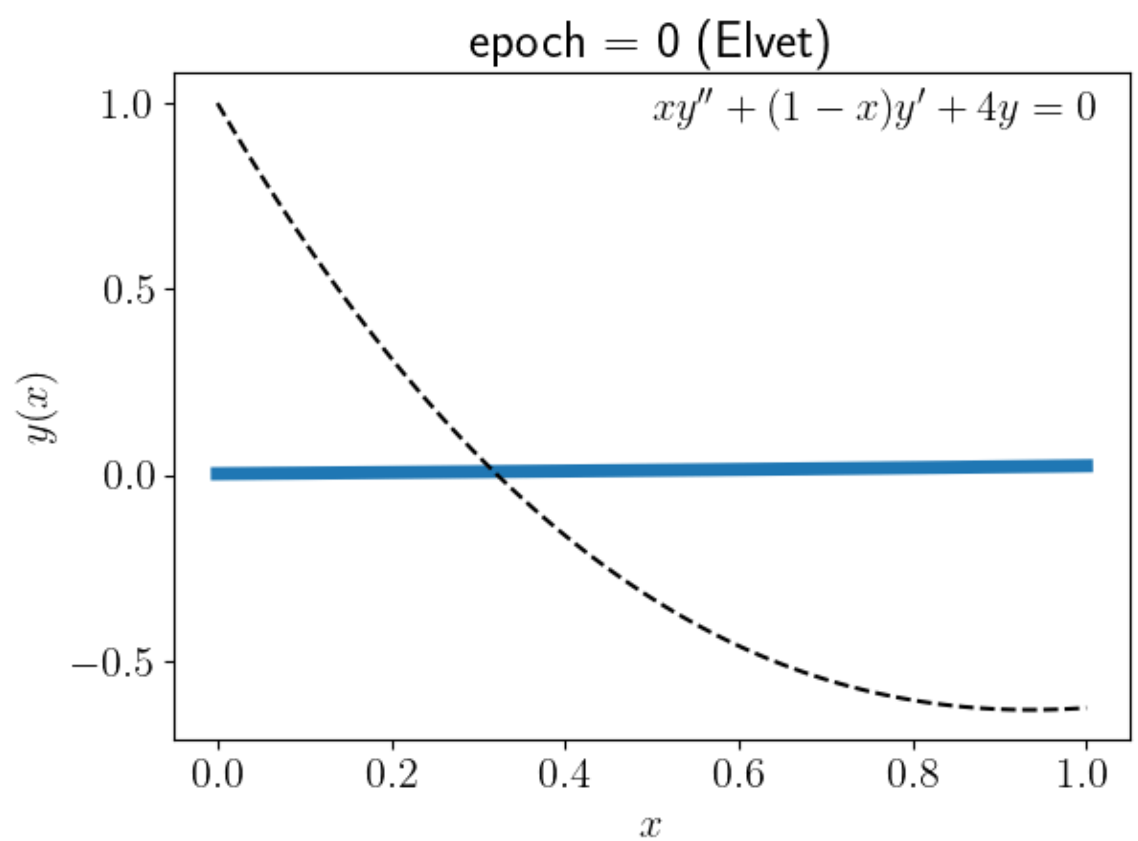
identify trial solution with network output $\hat{\phi}_m(\vec{x}) \equiv \check{N}_m(\vec{x}, \{w, \vec{b}\})$

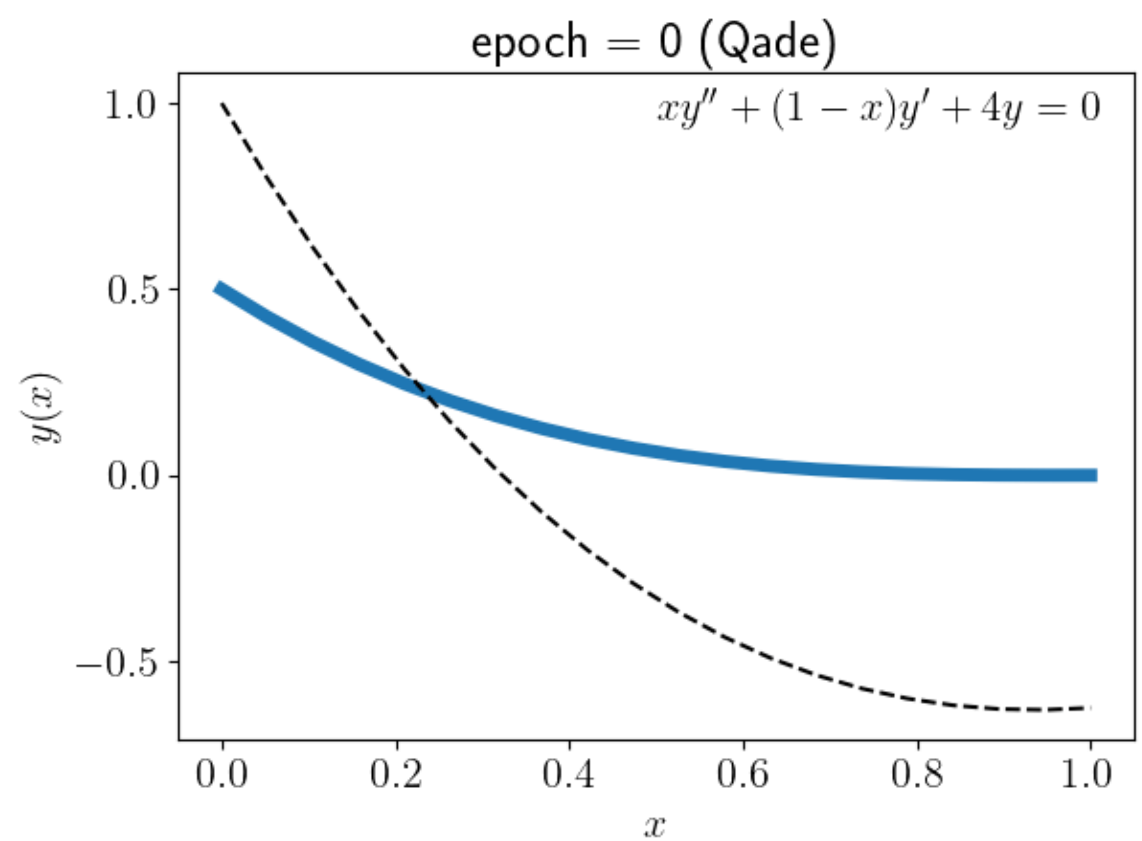
QADE: Solving differential equations with a quantum annealer

[Criado, MS '22]

Example Laguerre differential equation:

$$xy'' + (1-x)y' + 4y = 0 \text{ with } y(0) = 1 \text{ and } y(1) = L_4(1)$$





Classical Neural Network

[Piscopo, MS, Waite '19] [Araz, Criado, MS '21]

https://gitlab.com/elvet/elvet

Quantum algorithm

http://gitlab.com/jccriado/qade

QFitter

Example Higgs EFT fit:

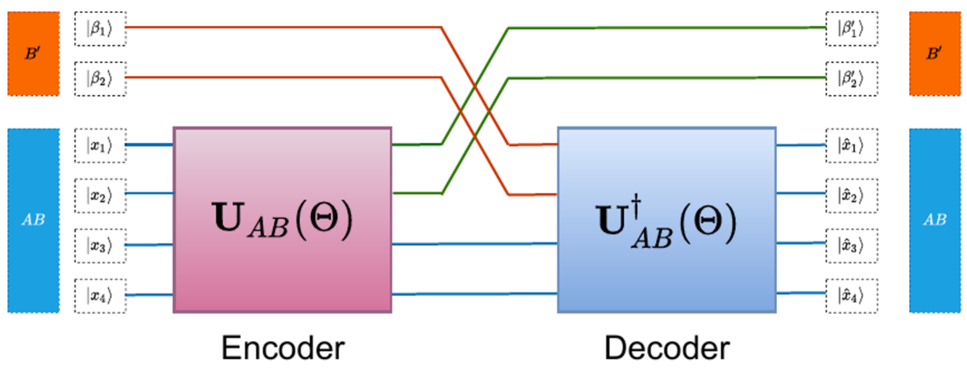
[Criado, Kogler, MS '22]

$$\mathcal{L} = \frac{c_{u3} y_t}{v^2}(\phi^\dagger\phi)(\bar{q}_L\tilde{\phi}u_R) + \frac{c_{d3} y_b}{v^2}(\phi^\dagger\phi)(\bar{q}_L\phi d_R) + \frac{i c_W g}{2m_W^2}(\phi^\dagger\sigma^a D^\mu\phi)D^\nu W^a_{\mu\nu} + \frac{c_H}{4v^2}\left(\partial_\mu(\phi^\dagger\phi)\right)^2 + \frac{c_\gamma (g')^2}{2m_W^2}(\phi^\dagger\phi)B_{\mu\nu}B^{\mu\nu} + \frac{c_g g_S^2}{2m_W^2}(\phi^\dagger\phi)G^a_{\mu\nu}G^{a\mu\nu} + \frac{i c_{HW} g}{4m_W^2}(\phi^\dagger\sigma^a D^\mu\phi)D^\nu W^a_{\mu\nu} + \frac{i c_{HB} g'}{4m_W^2}(\phi^\dagger D^\mu\phi)D^\nu B_{\mu\nu} + \text{h.c.}$$

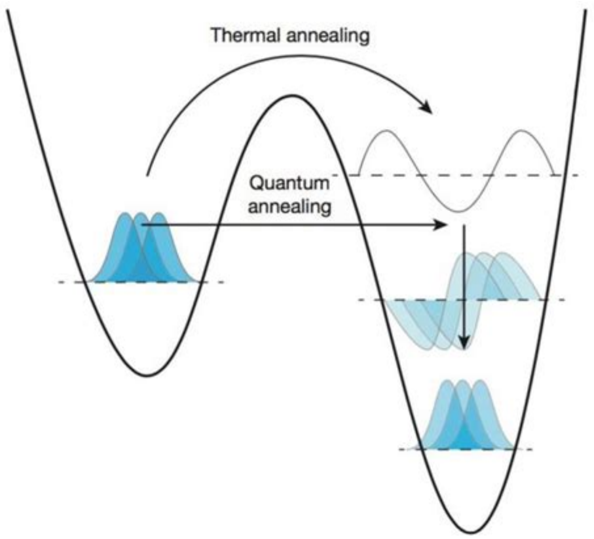
$$\chi^2 = \sum_{ij} V_a C^{-1}_{ab} V_b \qquad V_a = O_a^{(\text{exp})} - O_a^{(\text{th})}(c)$$

- Fast and reliable state-of-the-art Higgs, ELW, ... fits
- Convergence no problem for non-convex $\Delta\chi^2 = \chi^2 - \chi^2_{\text{min}}$ functions

Formulation	Method	Fit time	c_{HW}	c_H	c_g	c_γ	χ^2
Standard	Minuit (initial $c_{HW} = 0$)	2.0 s	-0.009	0.100	1.4×10^{-5}	3.2×10^{-6}	4110
	Minuit (initial $c_{HW} = -0.05$)	2.4 s	-0.050	0.039	-9.7×10^{-6}	-1.0×10^{-4}	135
	Simulated annealing (initial $c_{HW} = 0$)	642 s	-0.009	0.100	1.4×10^{-5}	3.7×10^{-6}	4110
	Simulated annealing (initial $c_{HW} = -0.05$)	644 s	-0.009	0.100	1.4×10^{-5}	3.7×10^{-6}	4110
QUBO	Simulated annealing (Class A)	6.4 s	-0.012	-0.054	-3.0×10^{-5}	3.9×10^{-5}	3910
	Simulated annealing (Class B)	6.4 s	-0.045	-0.175	-3.7×10^{-5}	1.8×10^{-4}	228
	Quantum annealing	0.2 s	-0.047	-0.050	1.9×10^{-5}	7.5×10^{-7}	68



Summary



- Quantum Machine Learning is exciting research area that rapidly expands, supported through private and public sector. Many algorithms to be invented.

- Quantum Machine Learning often shows an improved performance over classical Machine Learning, when limiting to a similar complexity of the model. Can exploit QM prop: entanglement, superposition principle and tunnelling

- For more exciting applications (quantum advantage), need development of technical realisation of quantum computers (size, fault tolerance, coherence, operations,...)