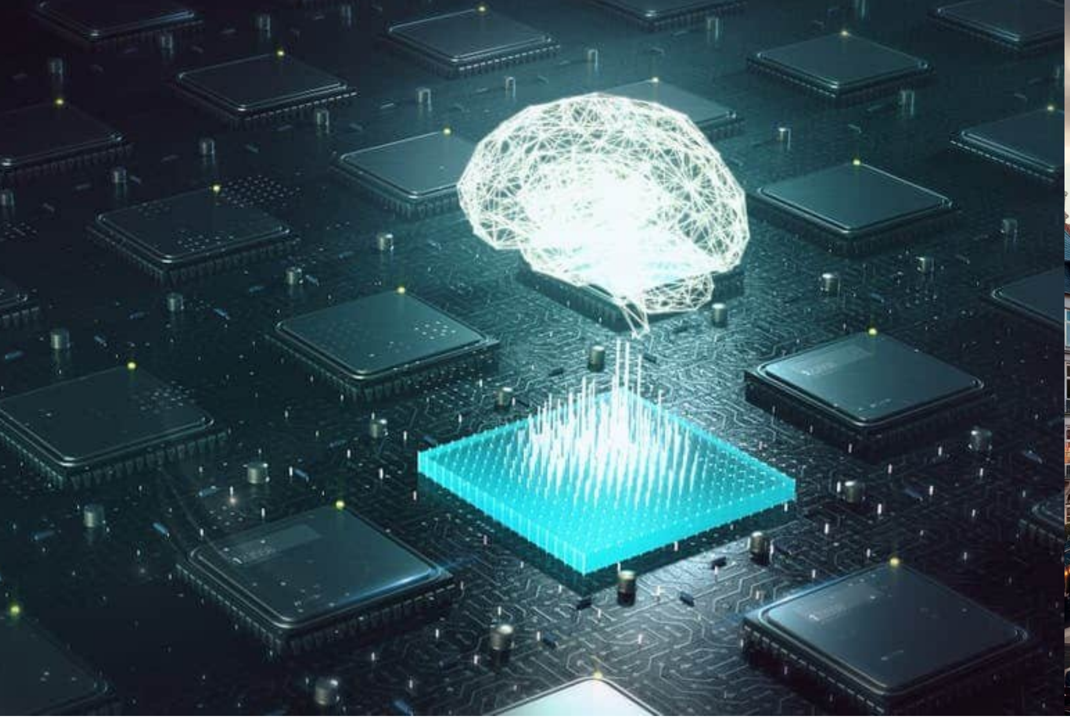
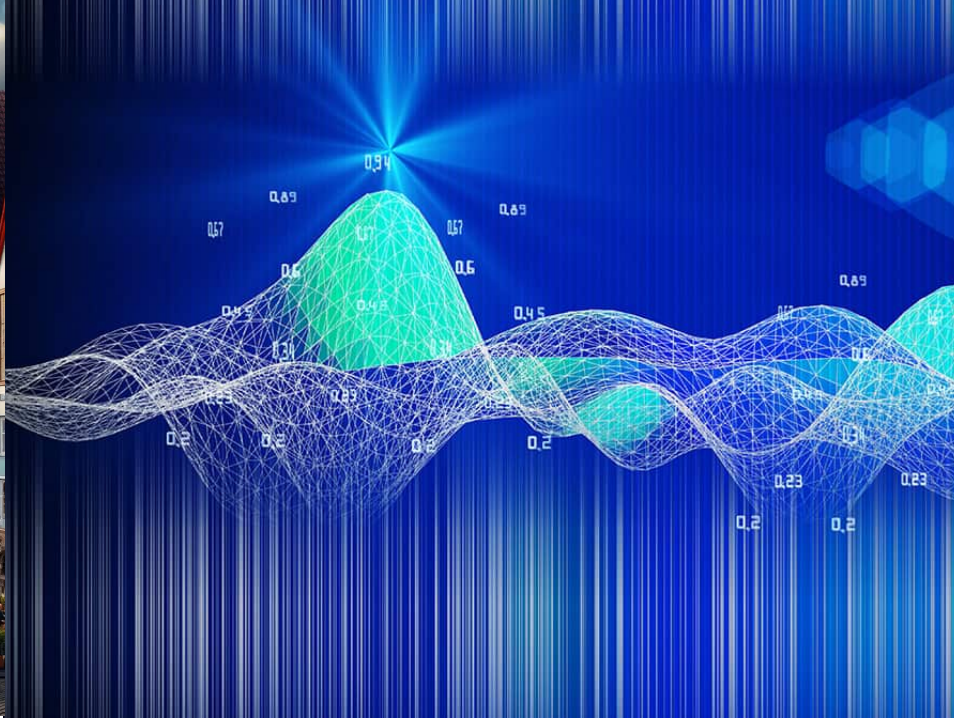


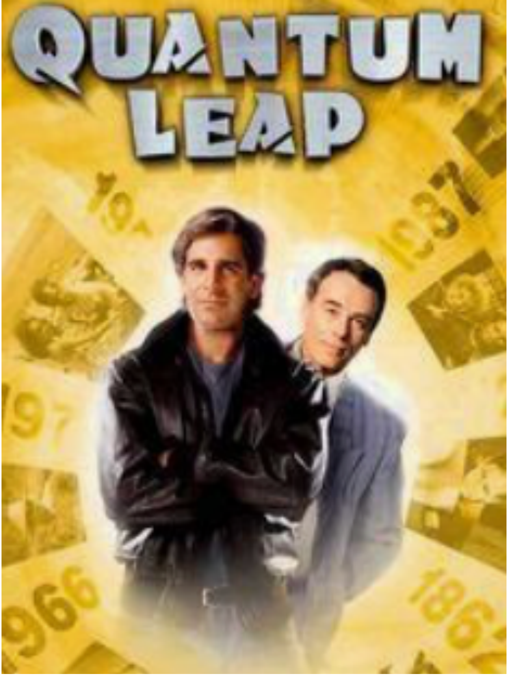


- Train by Tunnelling -

Quantum Annealing for AI Optimisation

Michael Spannowsky

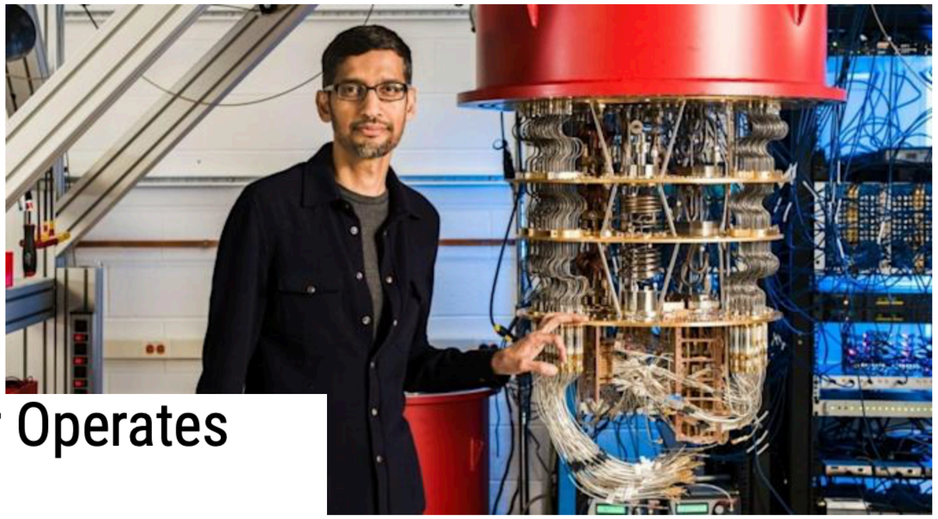
IPPP, Durham University



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

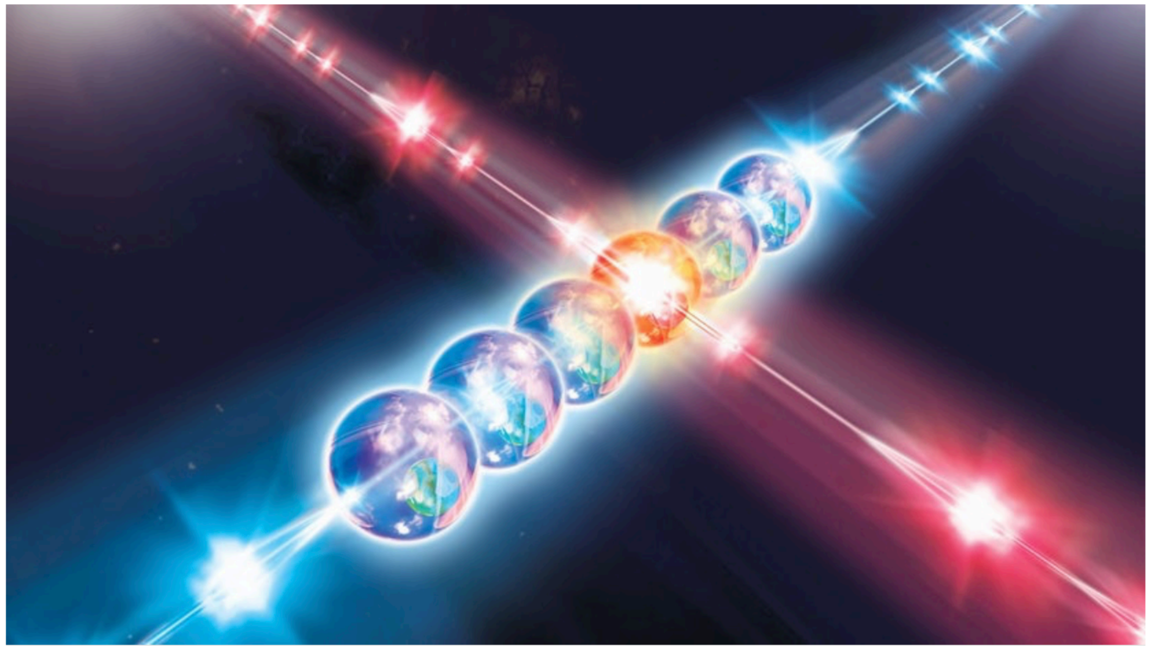
Master Thesis: Design of digitally enhanced power management circuits for Future Quantum Computers
Forschungszentrum Jülich · Jülich, North Rhine-Westphalia, Germany (On-site)
1 company alum

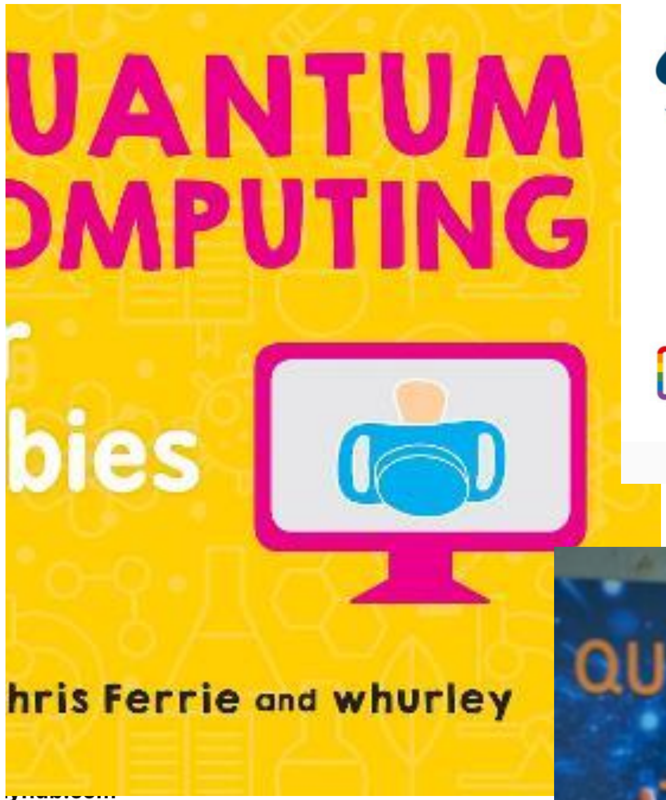
Expertin für Quantenkommunikation (w/m/d)
Deutsche Bahn · Frankfurt, Hesse, Germany (On-site)
Actively recruiting

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

Tunnelling is an immanently quantum mechanical phenomenon

Present in nature in various forms, eg:

- Possible first-order phase transitions in the early Universe -> stochastic Grav Waves
- Nuclear fusion in stars
- Radioactive decays, e.g. alpha decay
- Scanning Tunnelling Microscopy (STM)
- Proton tunnelling in biological systems

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

Easier said than done ... so how do we do that?

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

Analog vs Digital Quantum Computing

Analog and digital quantum computing are two different approaches to quantum computing, each with its own advantages and disadvantages.

Analog Quantum Computing (AQC):

- Based on the principle of quantum evolution of a quantum system, e.g. quantum annealing
- The system uses its intrinsic quantum dynamics, following the Schroedinger Equation
- Ground state represents the solution to the problem at hand
- Not always universal, but often well-suited for optimisation problems

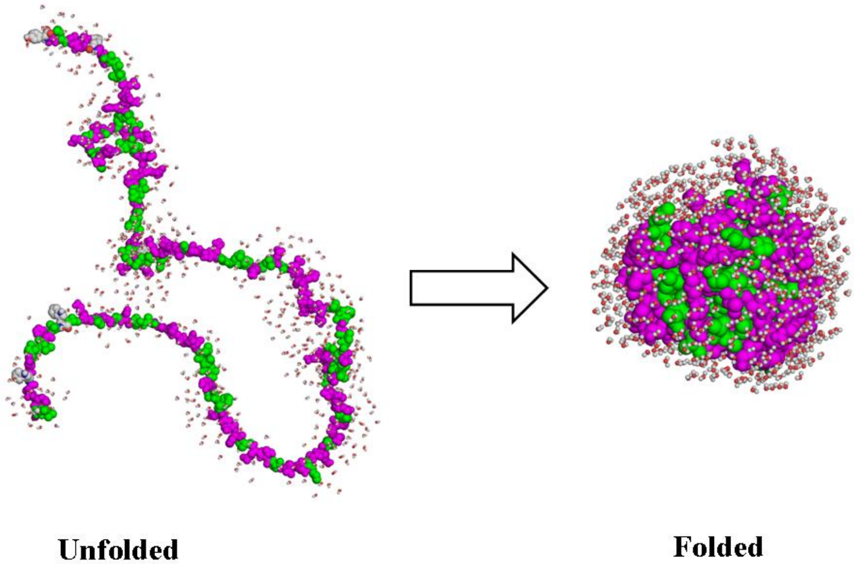
Example: D-Wave Systems. The D-Wave quantum annealer uses a network of qubits that can collectively tunnel through the solution space to find the global minimum of a given function.

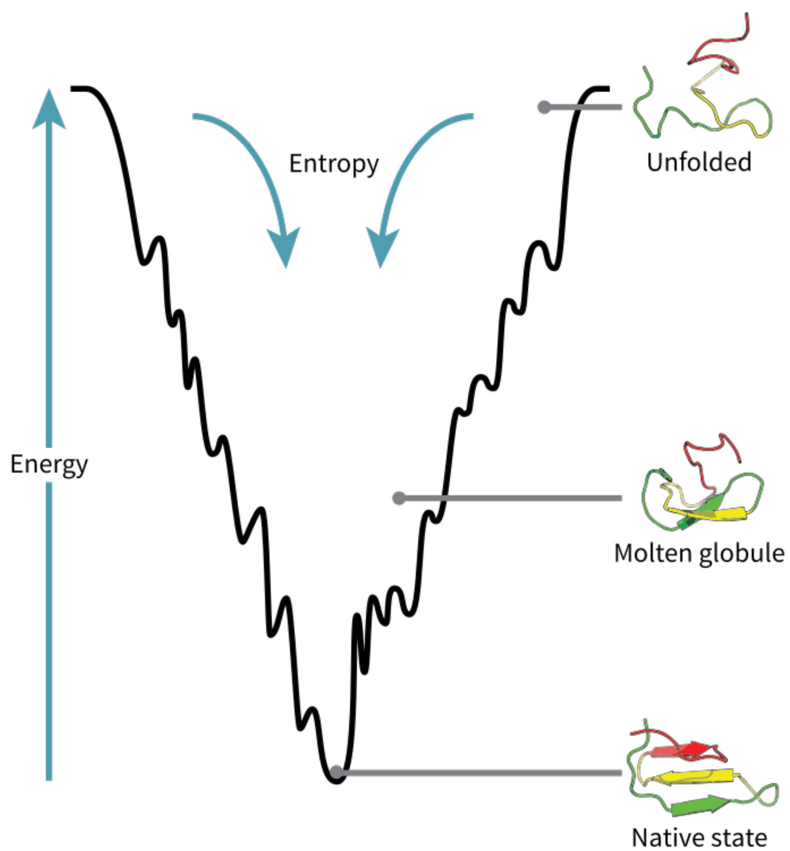
Digital Quantum Computing (DQC):

- Digital quantum computing, also known as gate-based quantum computing
- Uses quantum logic gates to perform operations on qubits
- Considered to be more versatile than analog computing.
- However, might require higher level of control over the quantum system, which can be challenging

Example: IBM's and Google's quantum computers use the gate-based model of quantum computing.

Protein-folding and Levinthal's Paradox





- Elongated proteins fold to same state within microseconds
- Some proteins have 3^{300} conformations
- Levinthal's Paradox (1969):
  Sequential sampling of states would take longer than lifetime of Universe (even if only nanoseconds per state spent)
- Solution: No sequential sampling, but rapid descend into the potential minimum.

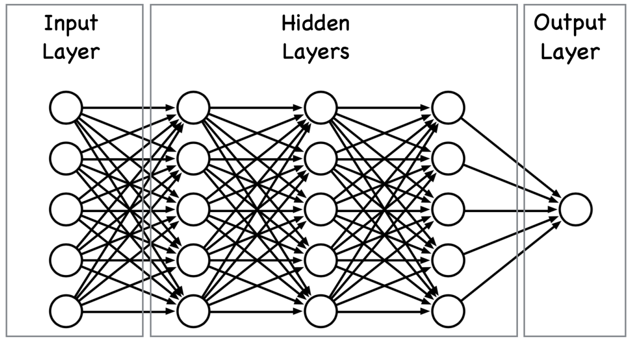
→ **Optimisation = Life**

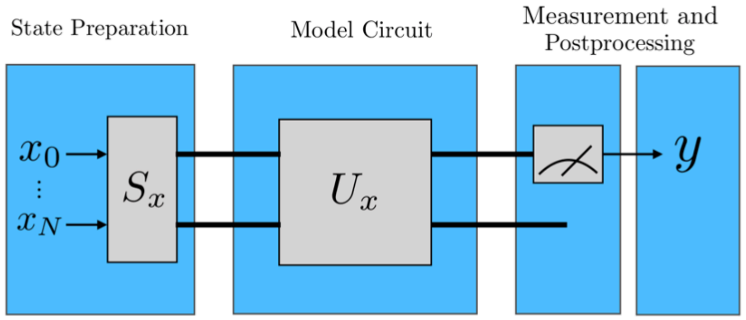
→ Solution of mathematical problem can be found quickly if encoded in ground state of complex system

Classical
ML Algorithms

Quantum
Computing

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





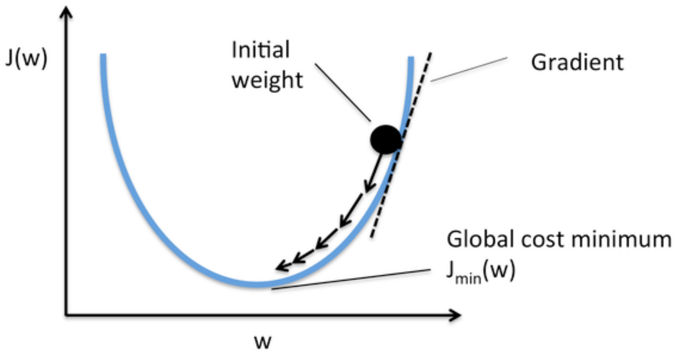
2. the calculation of a loss function used to define the task the method

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

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.

optimisation →
- Data Analysis (Classification, anomaly, regression, fitting, …)
- Simulation of field theories (Groundstate, tunnelling, Real-time…)
- Calculation of differential equations, etc etc

Analog quantum computing

Adiabatic quantum computing (AQC): If Hamiltonian changed smoothly and slowly enough system remains in ground state

[Farhi, Goldstone, Gutmann '00]

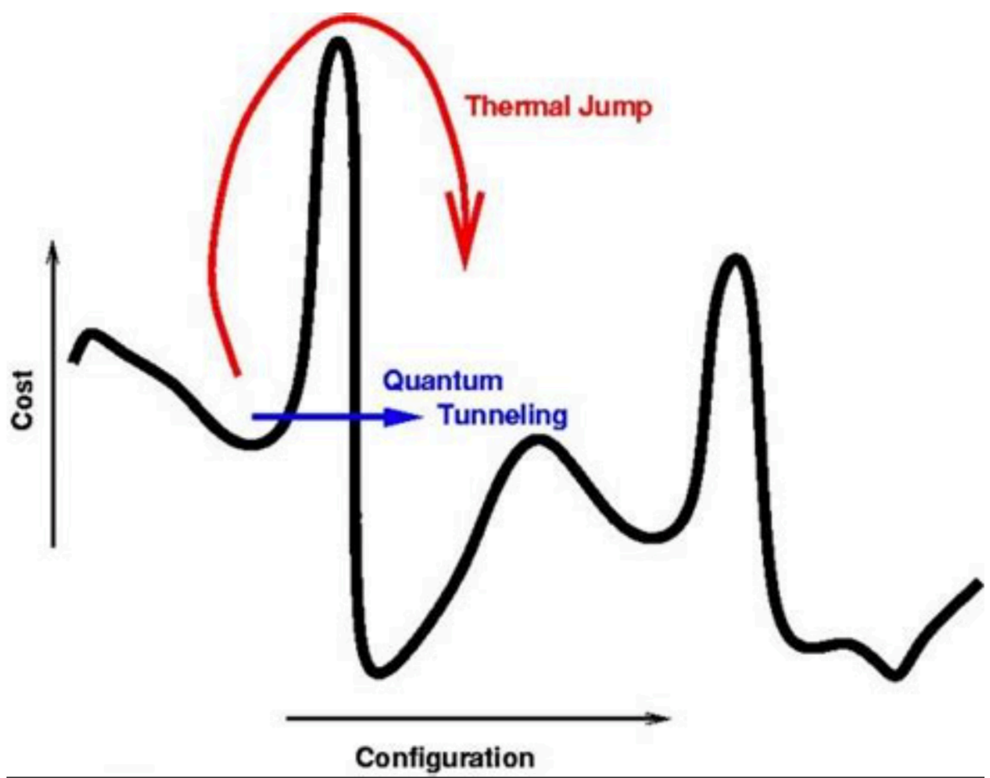
[Aharonov, et al '07]

Quantum annealing: transition from ground state of initial Hamiltonian into ground state of problem Hamiltonian, potentially via tunnelling

➡ D-wave:

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

[King et al '22]



- Thermal transitions are fast over broad shallow potentials $\sim e^{-\mathrm{height}/T}$
- Quantum tunnelling is fast through tall thin potentials $\sim e^{-\sqrt{\mathrm{height}}\times \mathrm{width}/\hbar}$

➡ Quantum Annealer = Laboratory for QFT

[Abel, MS '20]

[Abel, Chancelor, MS '20]

A quantum laboratory for QFT and QML

- aiming to go 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]

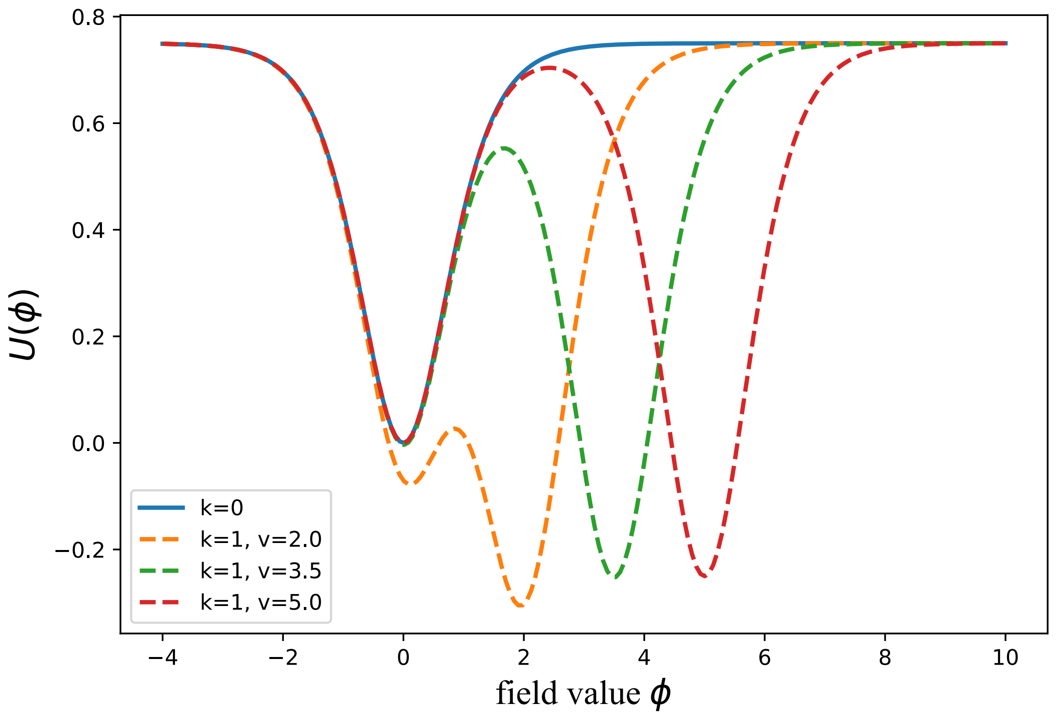
- To show that the system is a true and genuine quantum system we investigate if the state can tunnel from a meta-stable vacuum into a the true vacuum.

- Choose a potential of interest:

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

where $\phi = \eta/\eta_0$ ← time dependent ($k(t)$)

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



- For real-time evolution of field theory on QA see [Fromm, Philipsen, Winterowd '22]

The tunnelling probability in a QFT is calculated by evaluating the path-integral in Euclidean space around the action's critical points using the steepest gradient-descent method

$$\langle \eta_i | \eta_f \rangle_E = \int \mathcal{D}\delta\eta \, e^{-\hbar^{-1} \int dt \left(\frac{m(\dot{\eta}_{cl} + \delta\dot{\eta})^2}{2} + U(\eta_{cl} + \delta\eta) - E_0 \right)} = A e^{-\hbar^{-1} S_{E,cl}}$$

↖ quantum annealer

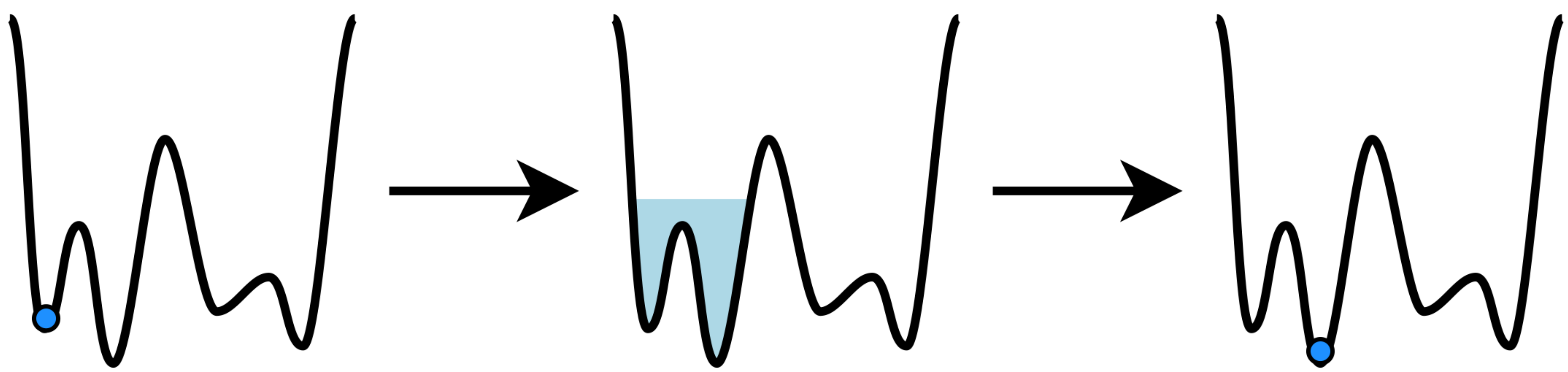
For the tunnelling rate $\Gamma = |\langle \eta_i | \eta_f \rangle_E|^2 \approx e^{-2\hbar^{-1} S_{E,cl}}$ with $S_{E,cl} = \int_{\eta_+}^{\eta_e} d\eta \sqrt{2m(U - E_0)}$

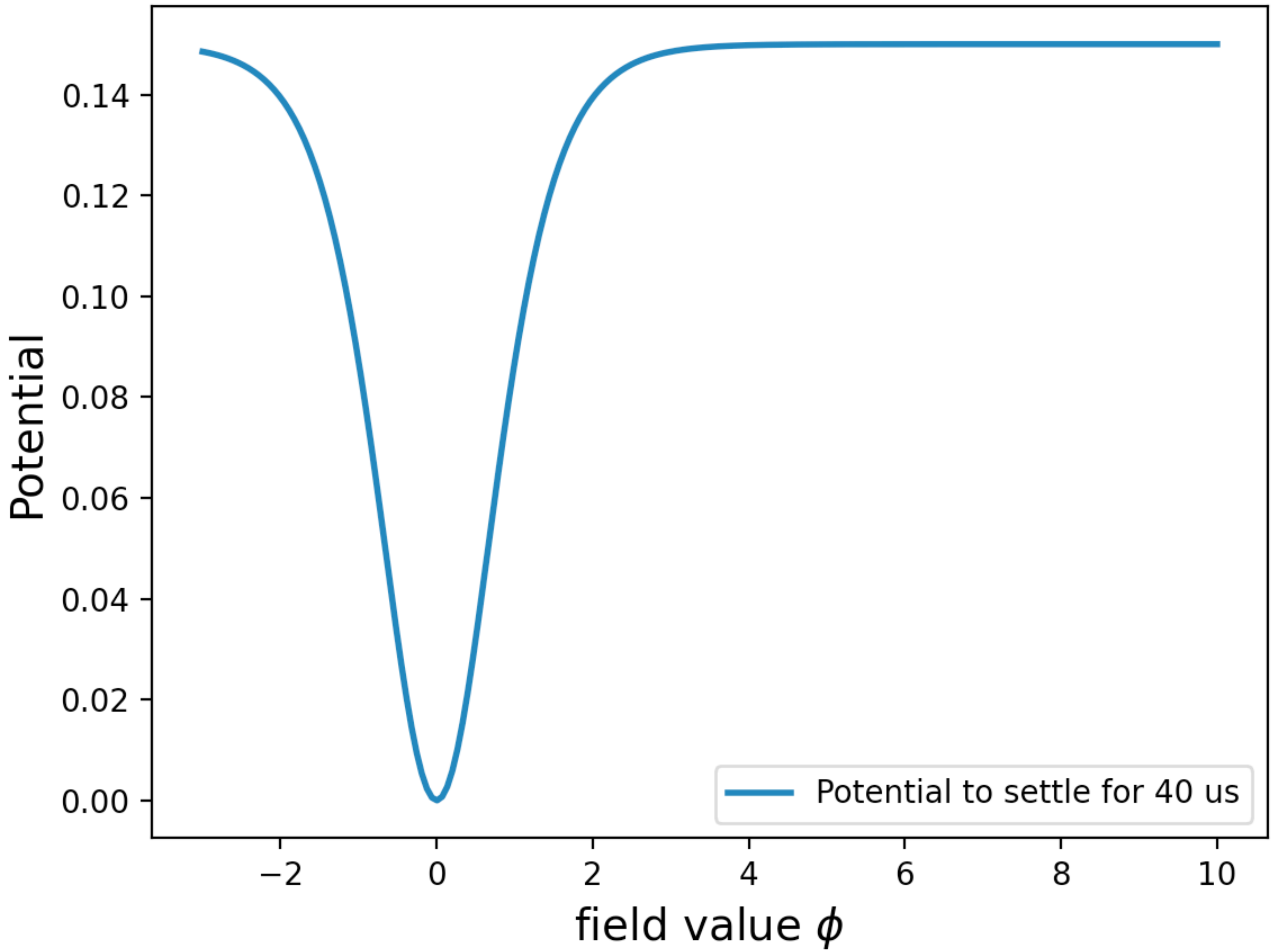
Exponent is object of interest: $\hbar^{-1} S_E \approx \gamma^{-\frac{1}{2}} \int_{\phi_+}^{\phi_e} \sqrt{\frac{3}{4} \tanh^2 \phi - \mathrm{sech}^2(\phi - v)} \, d\phi$ with $\gamma \stackrel{\mathrm{def}}{=} \hbar^2 / 2m\eta_0^2$

$$\log \Gamma \approx -2\hbar^{-1} S_E \approx \sqrt{\frac{3}{\gamma}} \left(\frac{5}{3} - v \right)$$

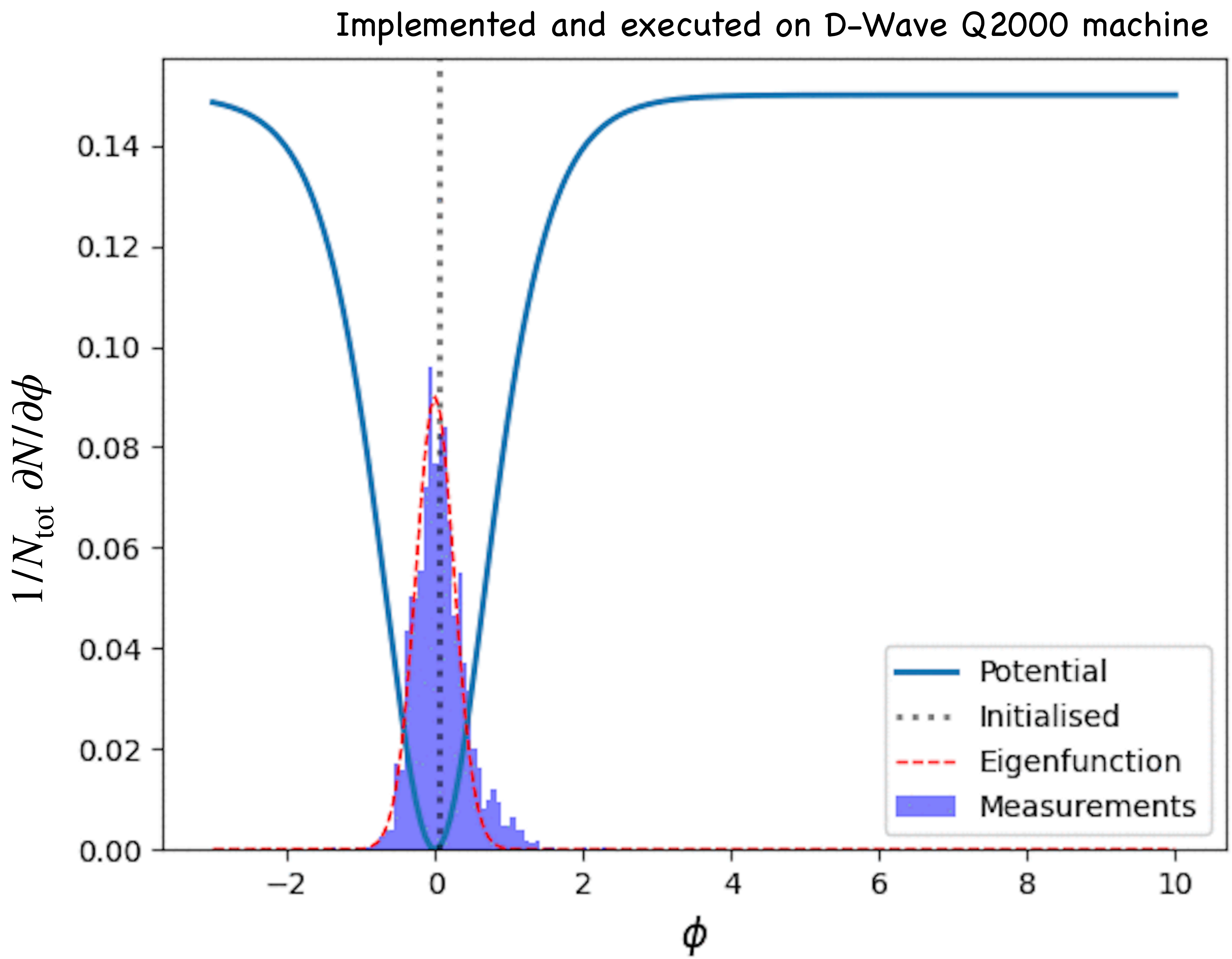
D-Wave reverse annealing

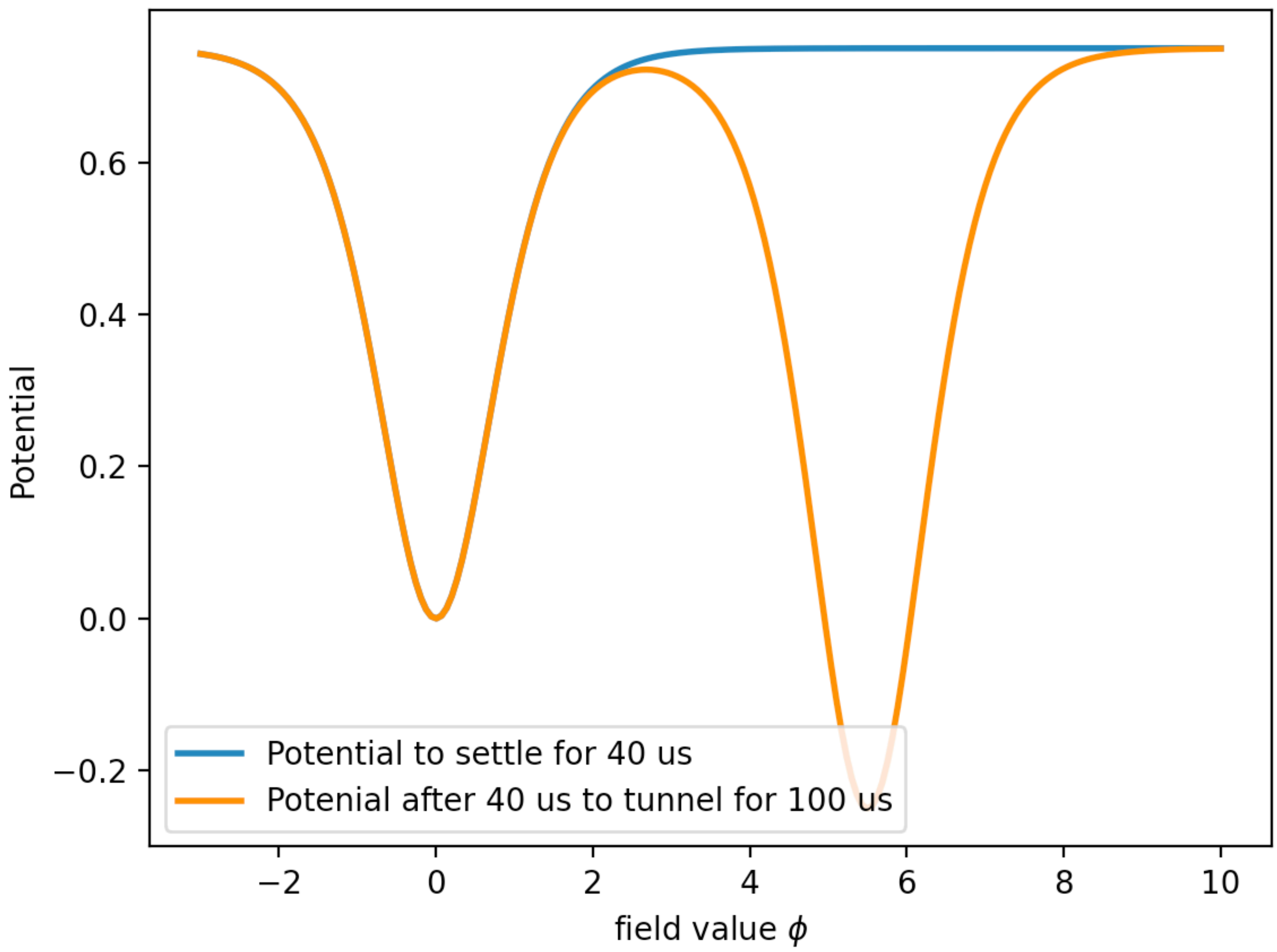
starts at sq=1 (classical) -> sq < 1 (quantum) -> measurement in sq=1 (classical)



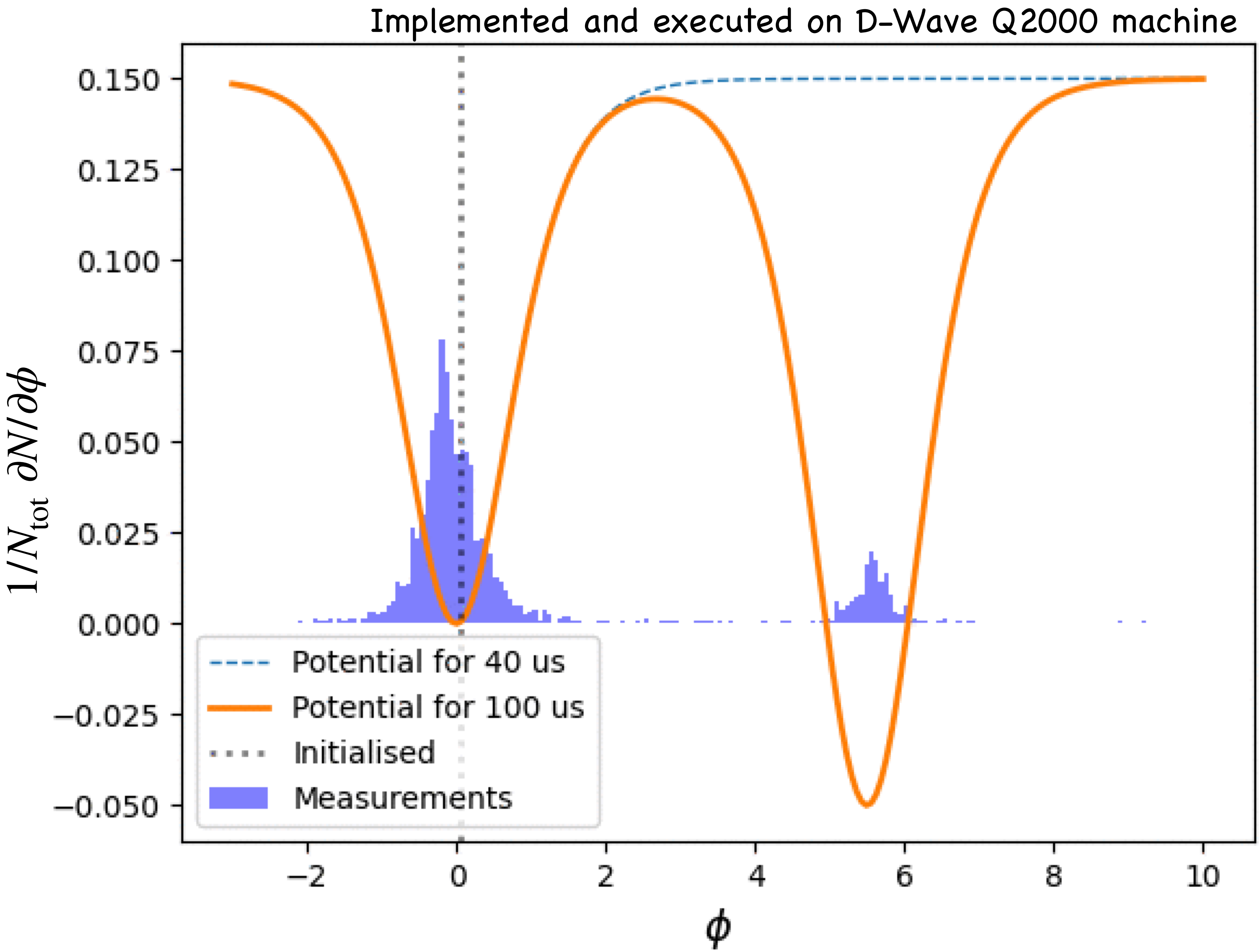


0.14
0.12
0.10
0.08
0.06
0.04
0.02
0.00
Potential
−2
0
2
4
6
8
10
field value ϕ
Potential to settle for 40 us


Implemented and executed on D-Wave Q2000 machine
0.14
0.12
0.10
0.08
0.06
0.04
0.02
0.00
$1/N_{\mathrm{tot}}\ \partial N/\partial\phi$
−2
0
2
4
6
8
10
ϕ
Potential
Initialised
Eigenfunction
Measurements

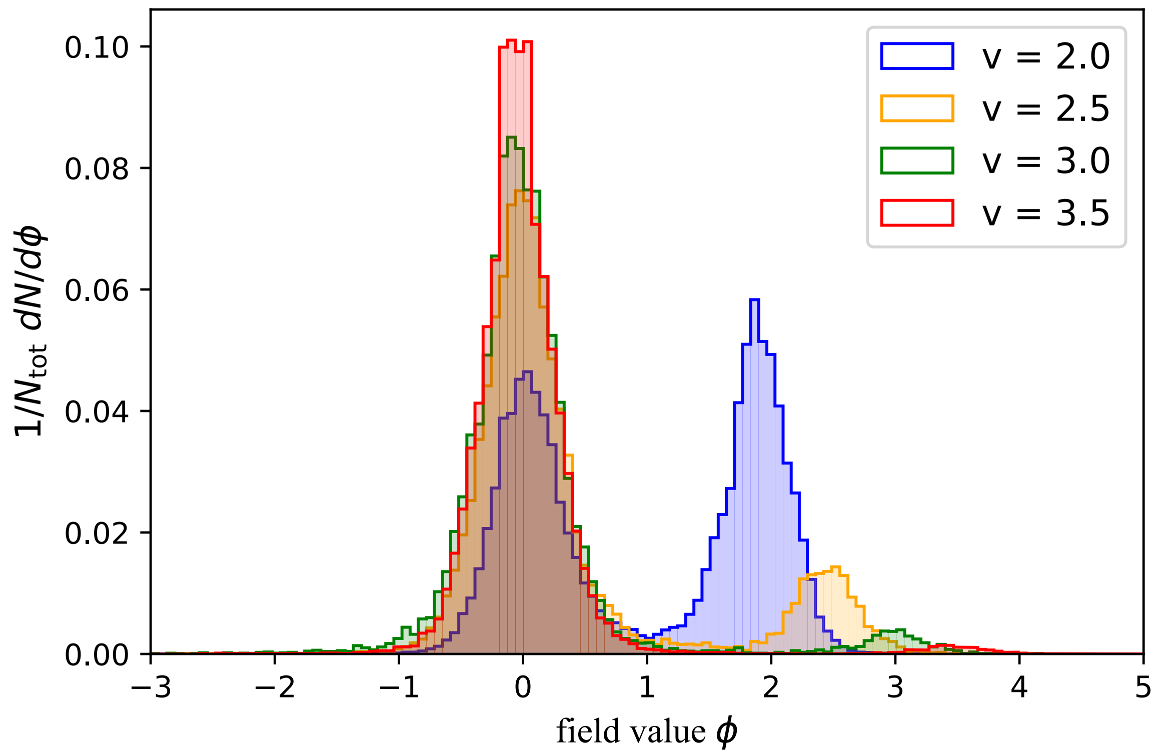


Potential
0.6
0.4
0.2
0.0
−0.2
−2
0
2
4
6
8
10
field value ϕ
Potential to settle for 40 us
Potenial after 40 us to tunnel for 100 us



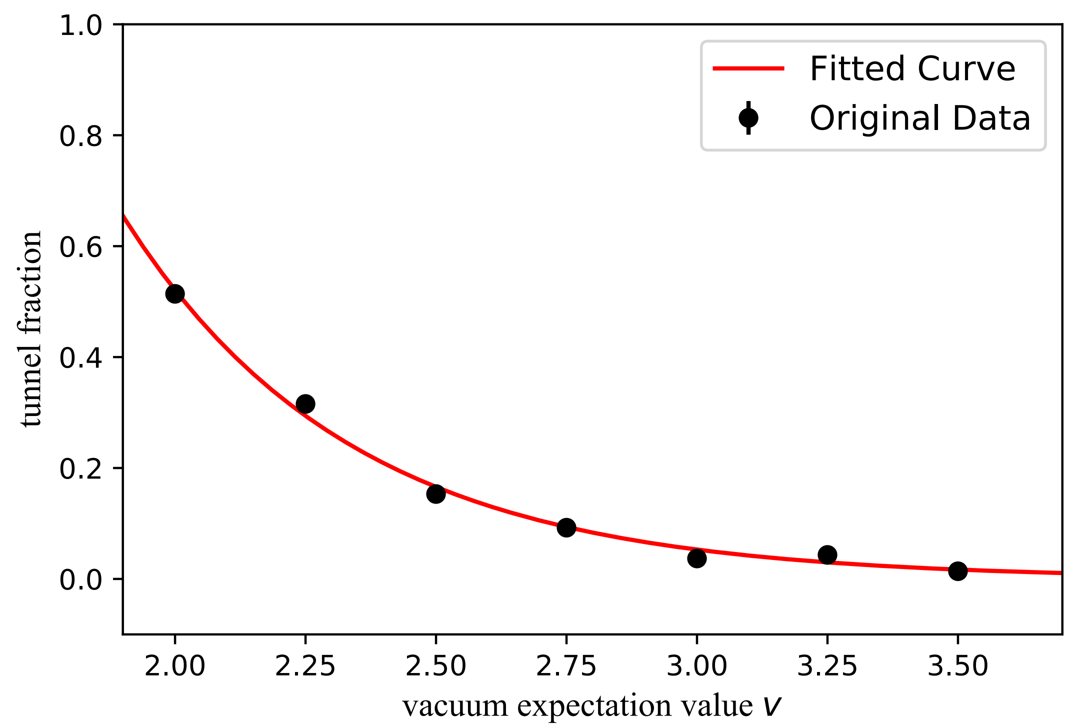
Implemented and executed on D-Wave Q2000 machine
0.150
0.125
0.100
0.075
0.050
0.025
0.000
−0.025
−0.050
$1/N_{\rm tot}\ \partial N/\partial\phi$
−2
0
2
4
6
8
10
ϕ
Potential for 40 us
Potential for 100 us
Initialised
Measurements

Results: it decays with v as expected



Perform tunnelling for

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

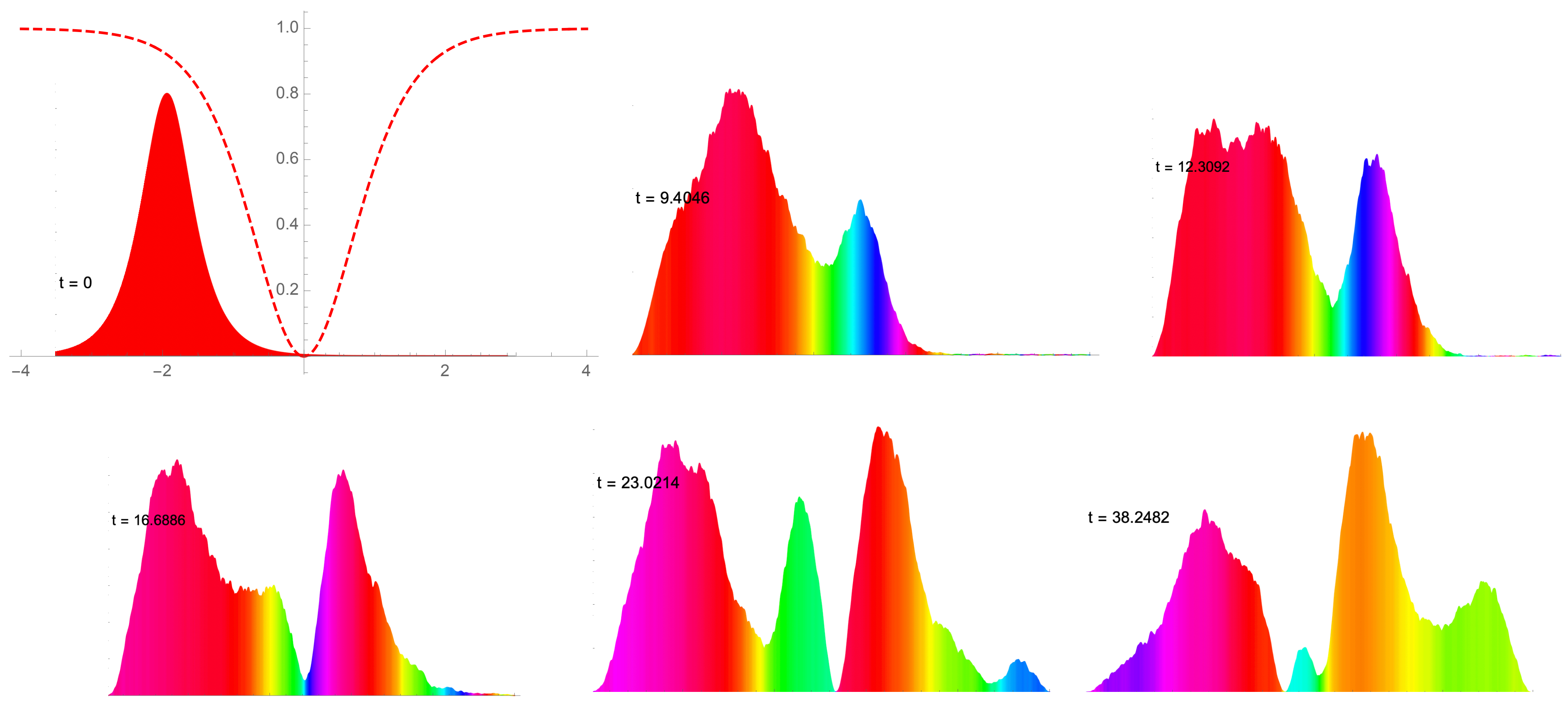


Theory: $\log\Gamma = 3.0 \times (1.66 - v)$

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

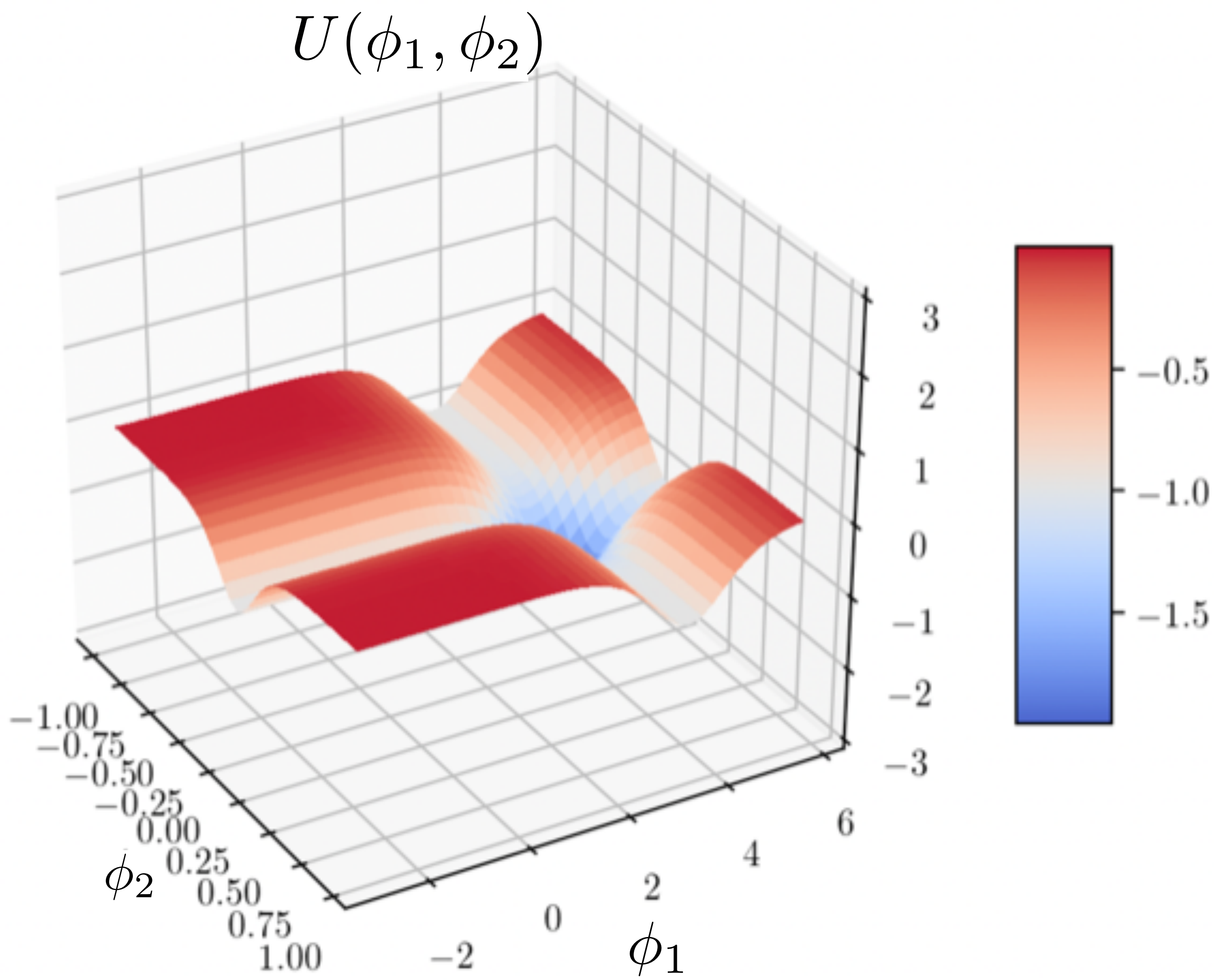
Also dynamics has characteristic behaviour. For example it still "tunnels" to the bottom of a potential even if there is no barrier: i.e. the wave function leaks across, rather than rolling as a lump —

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

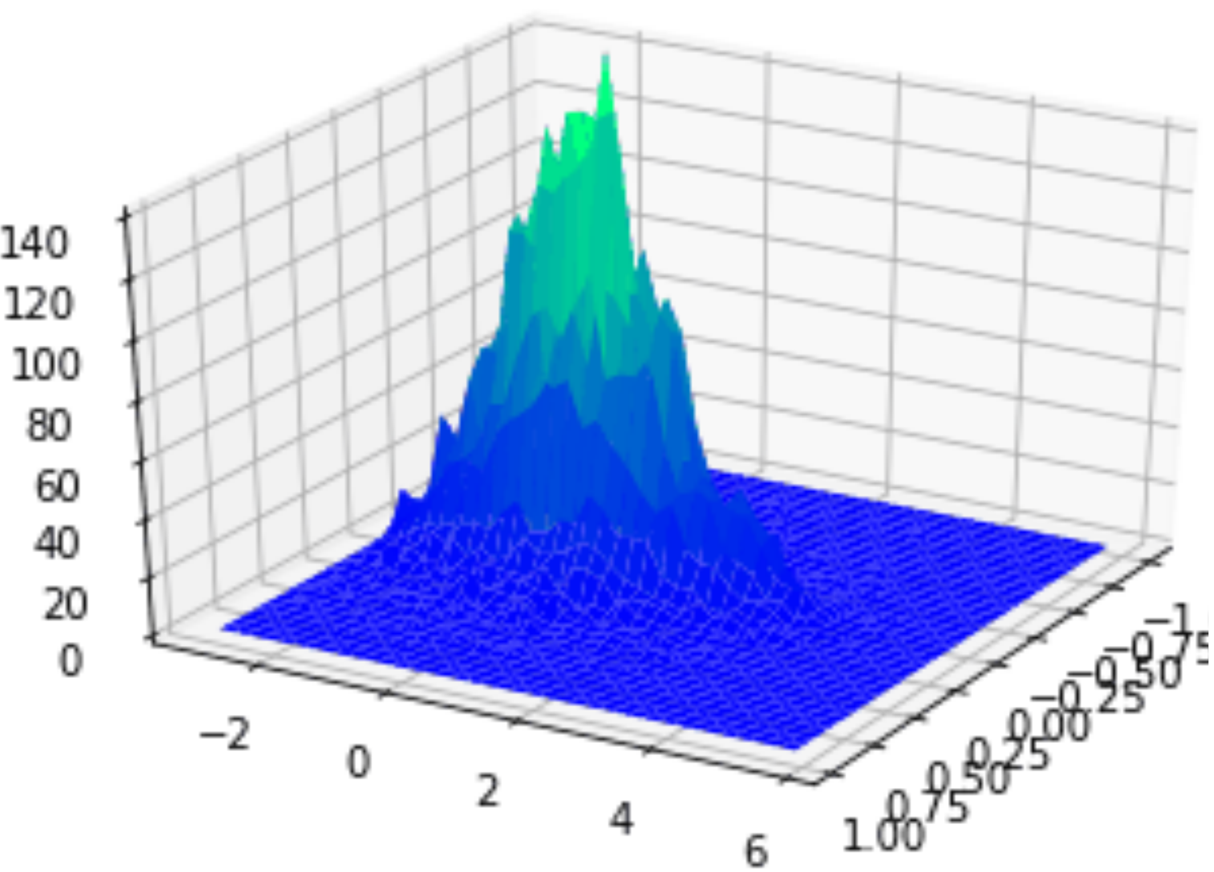


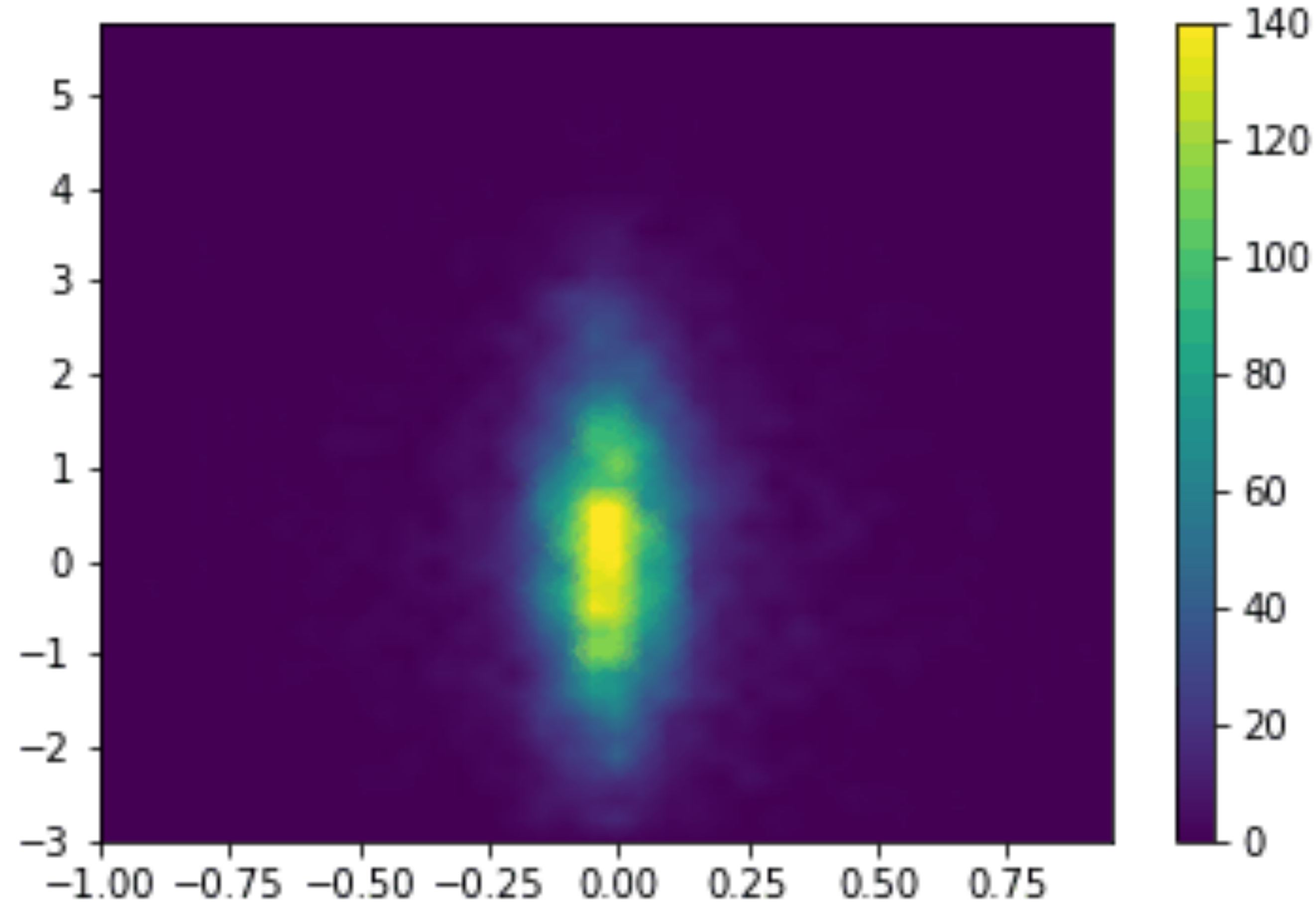
Also dynamics has characteristic behaviour. For example it still transits to the bottom of a potential even if there is no barrier i.e. the wave function leaks across, rather than rolling as a lump

2D example potential



Also dynamics has characteristic behaviour. For example it still transits 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



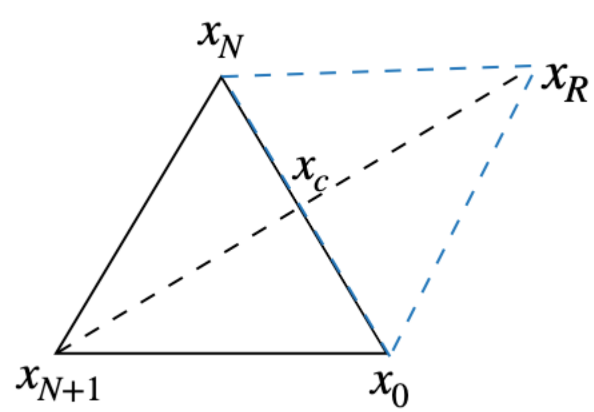


Optimisation comparison quantum vs classical

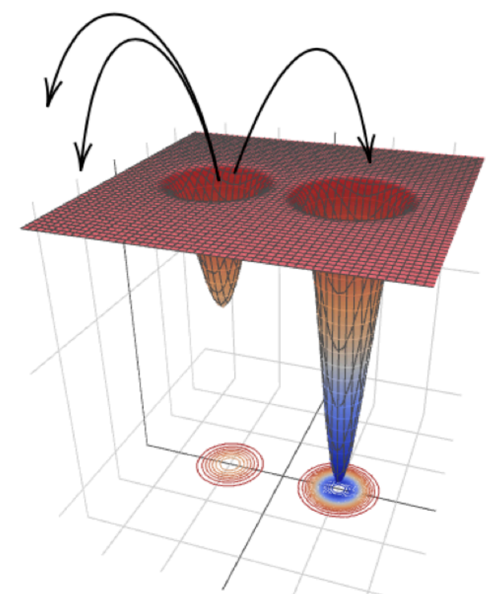
gradient descent

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

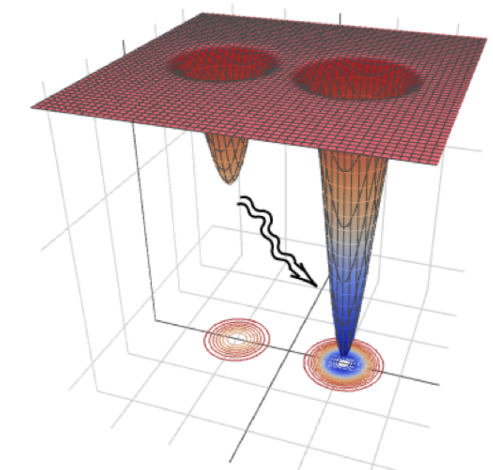
Nelder-Mead



Thermal Annealing



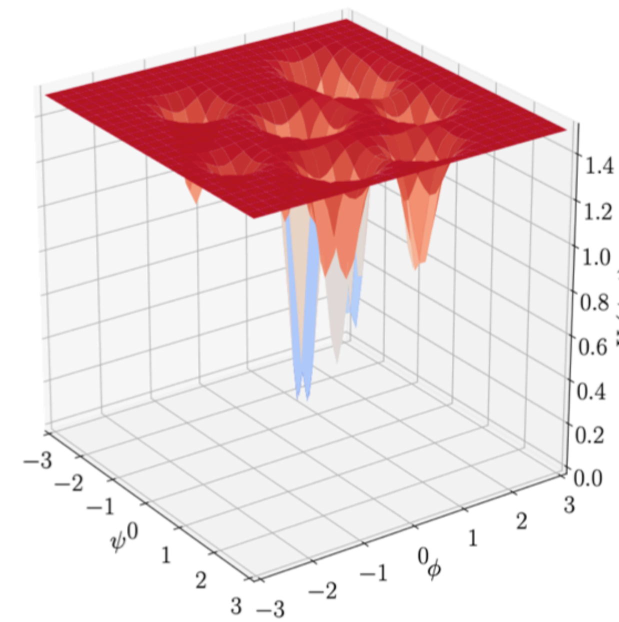
Quantum Annealing

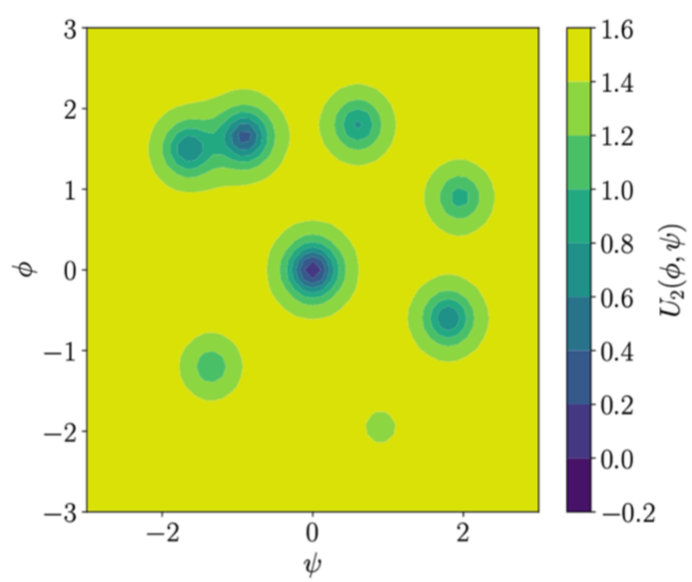


Applied to several examples in [Abel, Blance, MS '21], let's show one here:

Multi-well potential

enocded on D-Wave quantum annealer



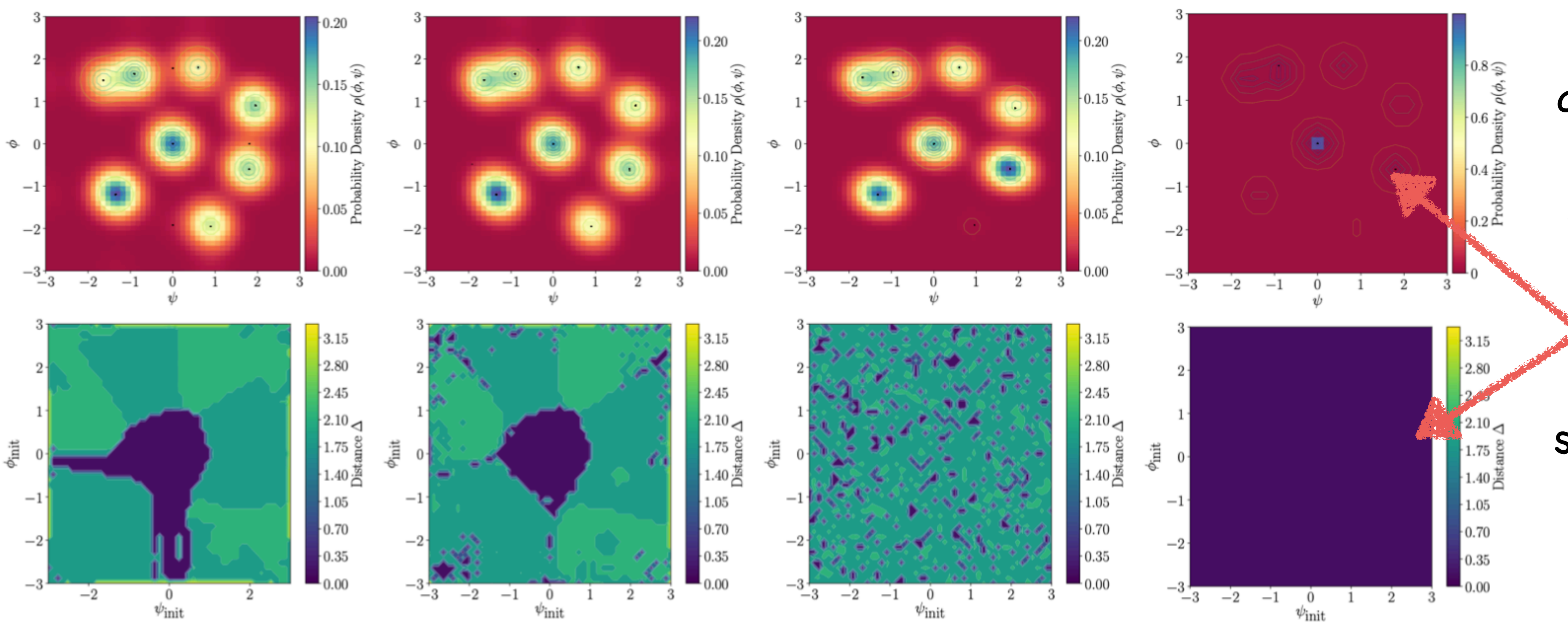


Results for Multi-well potential

- 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

[Abel, Blance, MS '21]



(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 advantage on current device

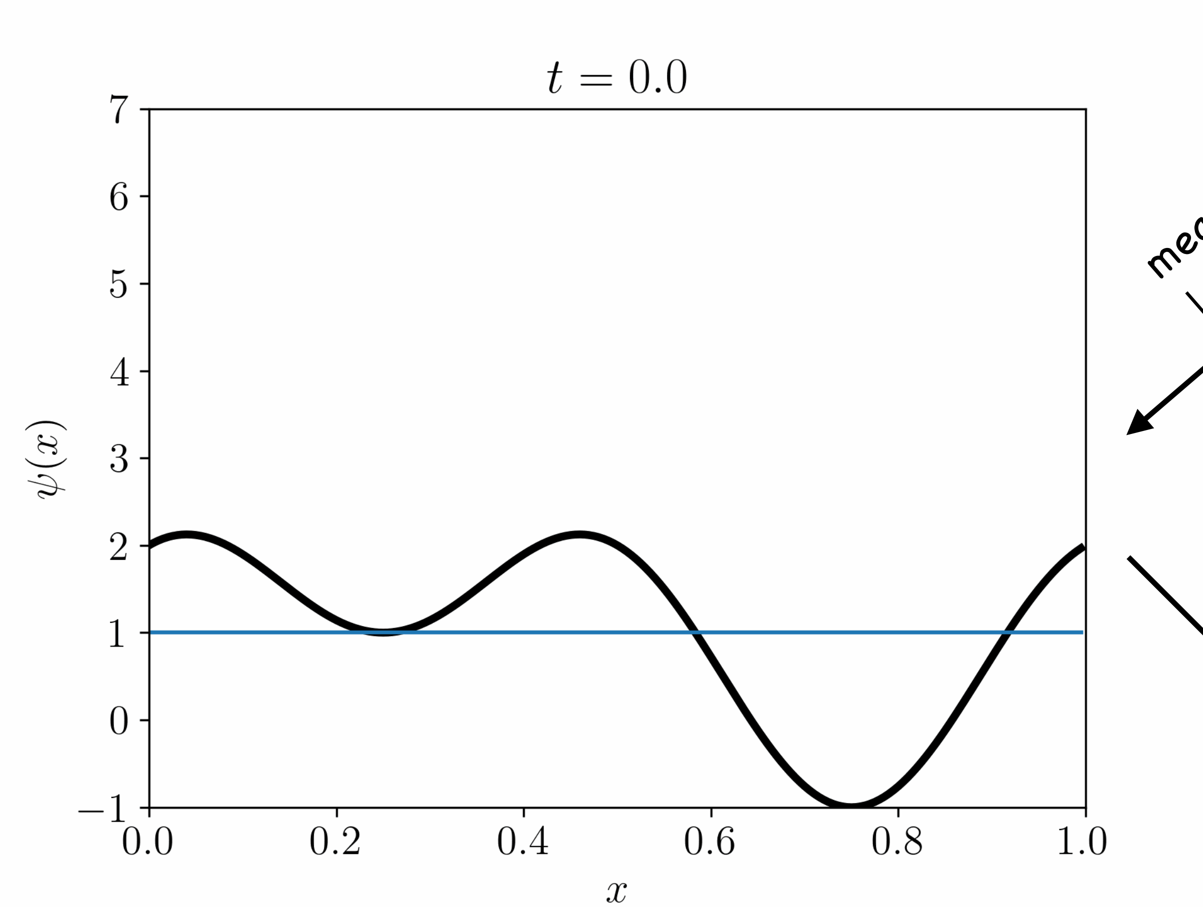
Completely Quantum Neural Networks

[Abel, Criado, MS '22]

[Abel, Criado, MS '23]

- Encode network, i.e. entire loss function, as Hamiltonian

- Gradient-free optimisation -> particularly important for discrete/binary NN

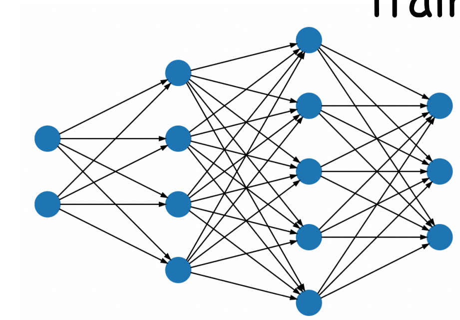


measure $|\psi(x)|^2$

$$H = (1-t)\,\frac{p^2}{2m^2} + t\;V(x)$$

$$\text{Loss} = \frac{1}{N}\sum_i^N |\phi(x,i) - \kappa_i|^2$$

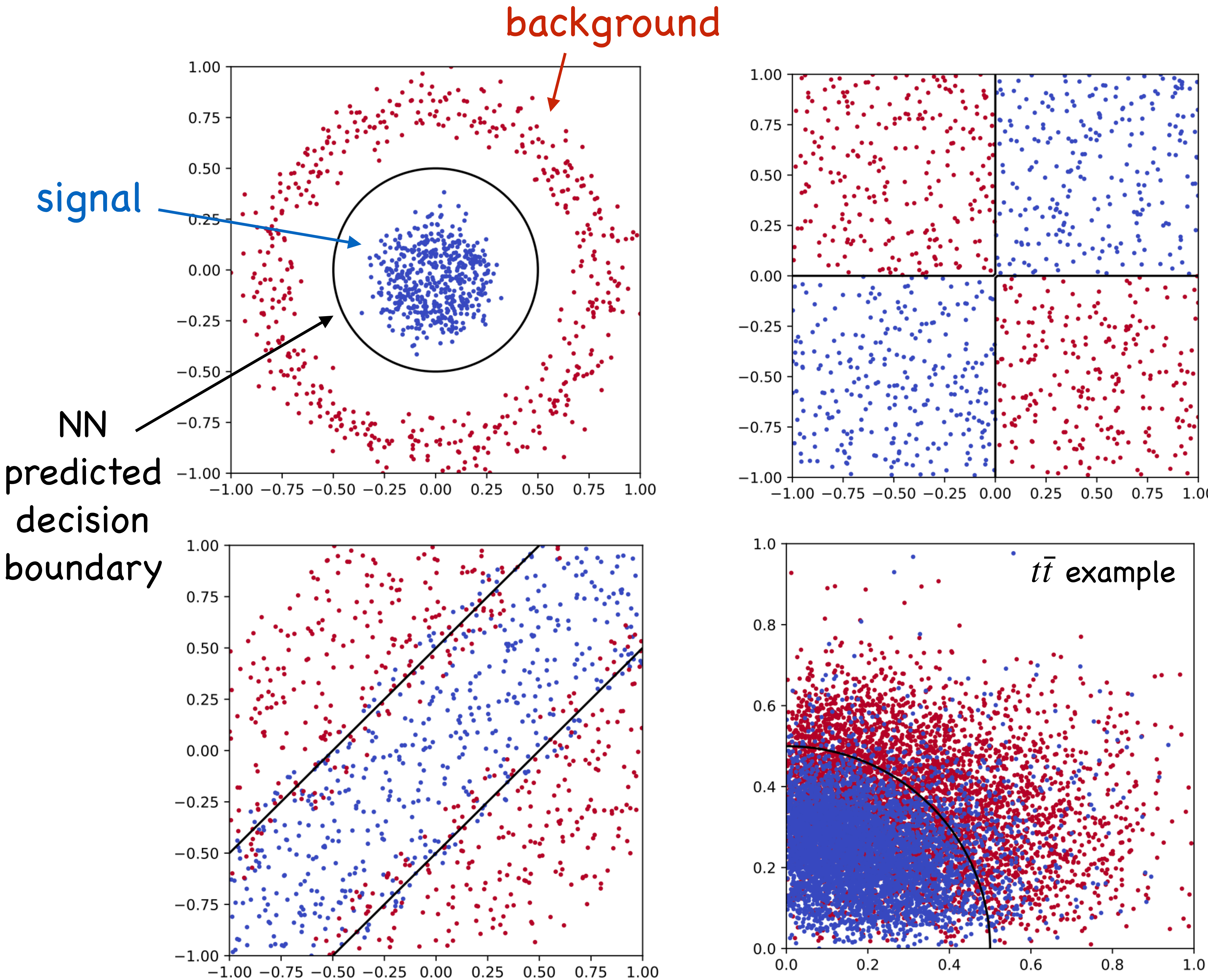
NN output with trainable weights x



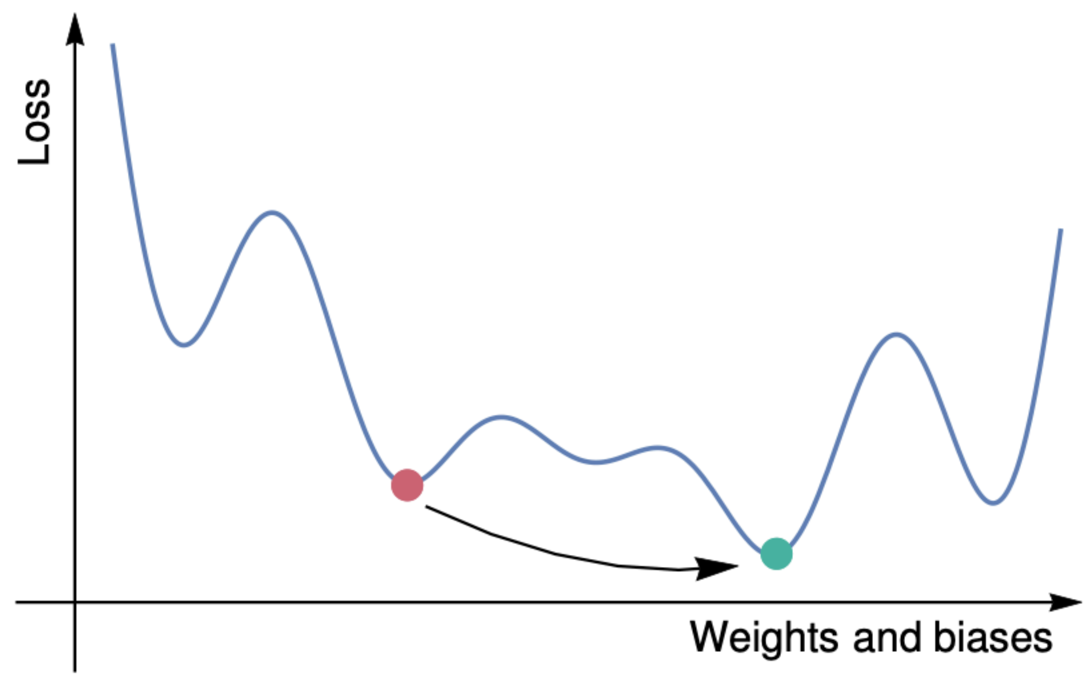
extract optimal weights x

- No hybrid training, but quantum tunnelling / AQC to train network

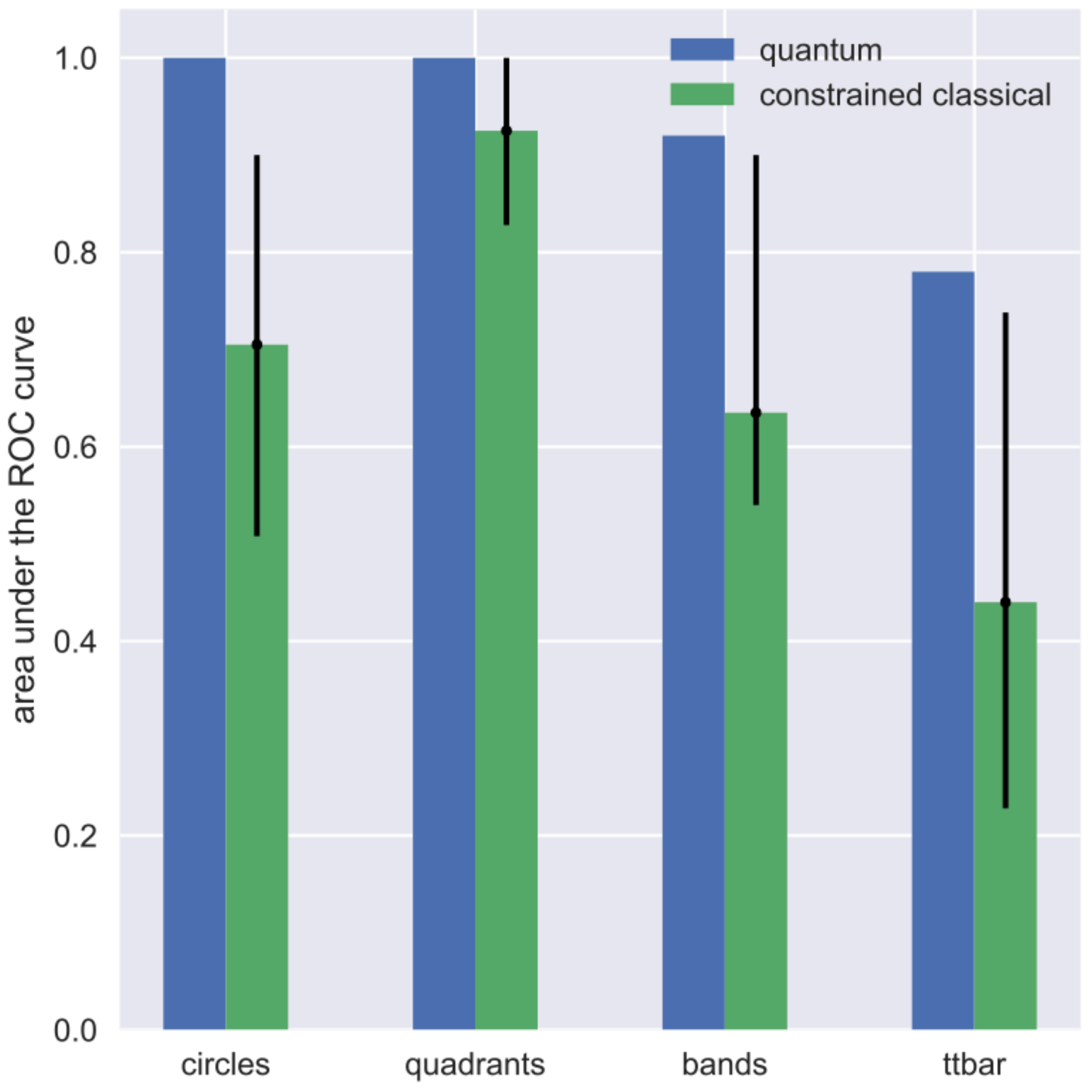
- Train fast and optimally on QC -> **read off parameters for deployment classically**


background
signal
NN
predicted
decision
boundary
$t\bar{t}$ example
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

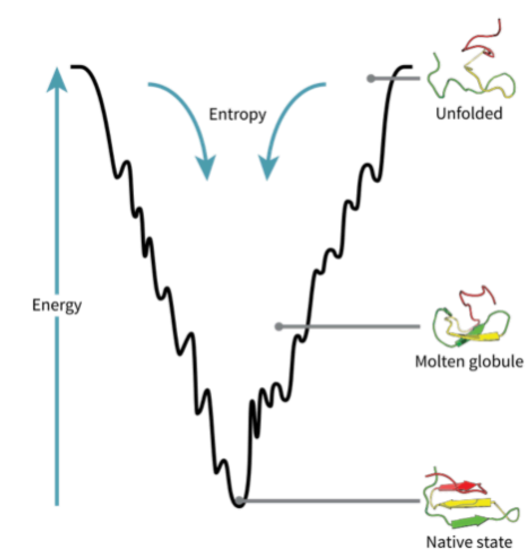
Application to differential equations and variational methods

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

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



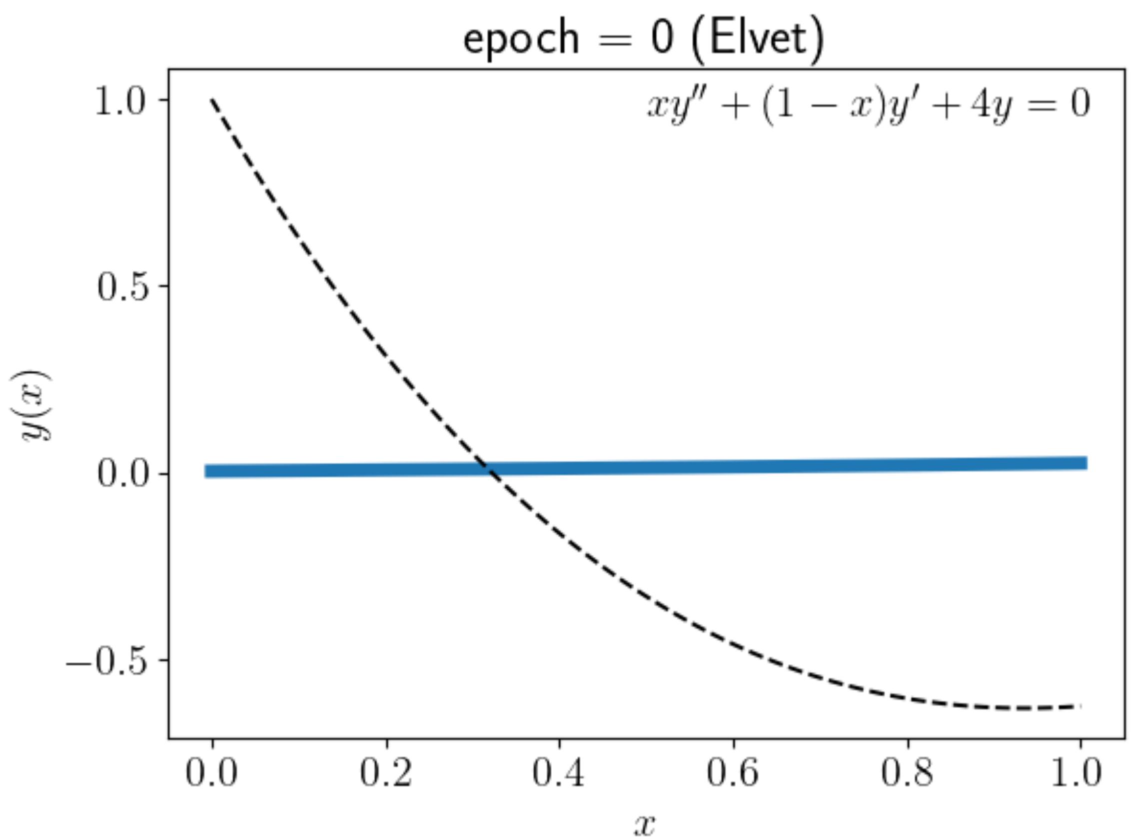
[Piscopo, MS, Waite '19]

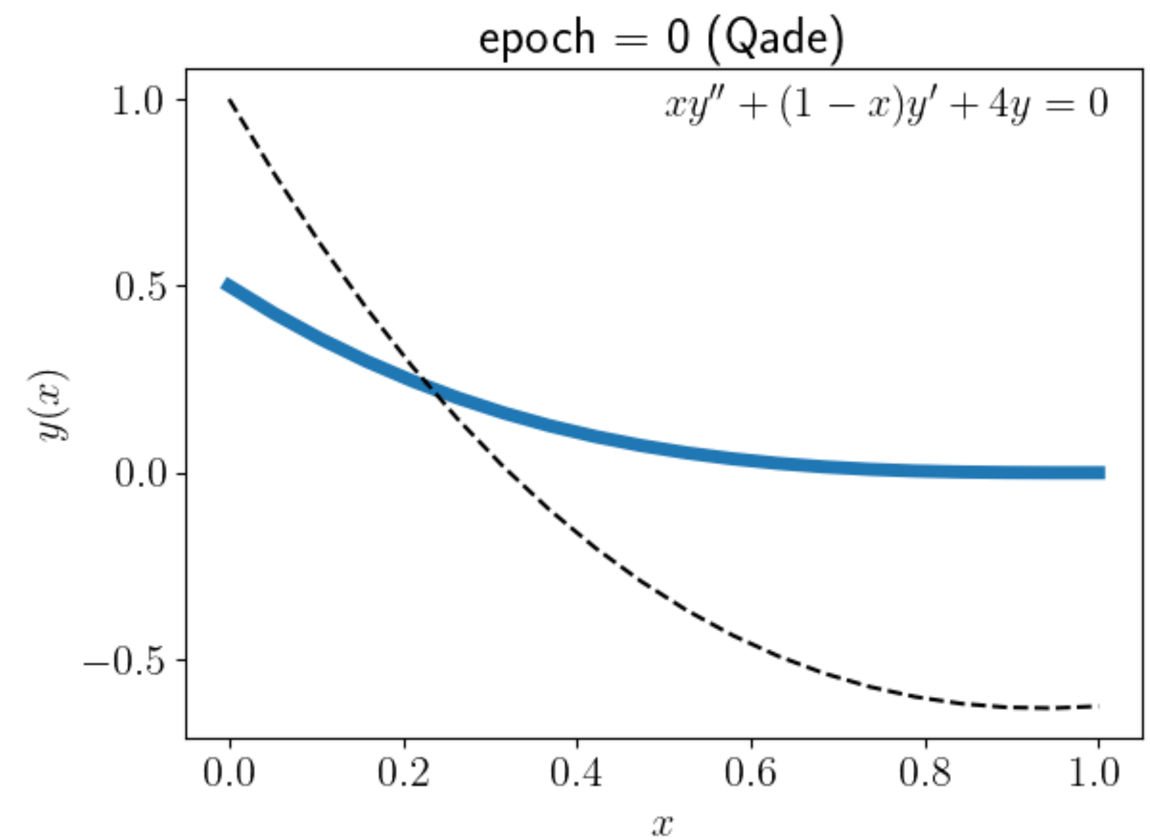
identify trial solution with network output $\hat{\phi}_m(\vec{x}) \equiv \check{N}_m(\vec{x}, \{w, \vec{b}\})$

QADE: Solving differential equations with a quantum annealer

Define Loss as: $$\mathscr{L} = \sum_i \mathrm{E}_i(f, \partial f, \ldots)^2 + \sum_j \mathrm{BC}_j(f, \ldots)^2$$

Example Laguerre differential equation $xy'' + (1-x)y' + 4y = 0$ with $y(0) = 1$ and $y(1) = L_4(1)$





Classical Neural Network approach (Elvet)

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

https://gitlab.com/elvet/elvet

Quantum algorithm (QADE)

gitlab.com/jccriado/qade [Criado, MS '22]

QFitter

Example Higgs EFT fit:

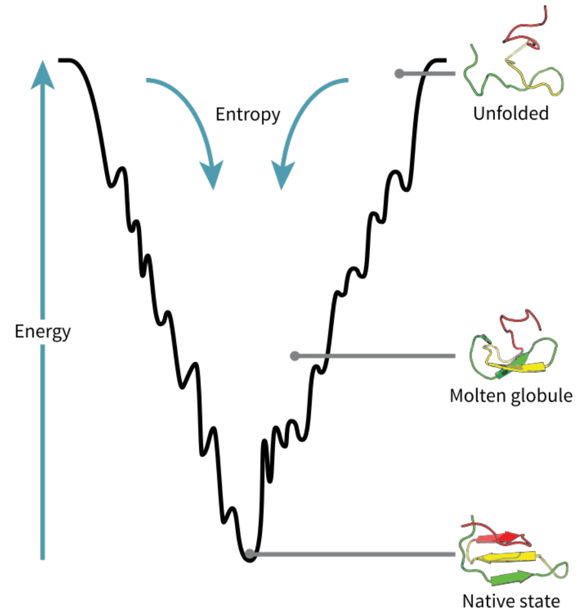
[Criado, Kogler, MS '22]

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

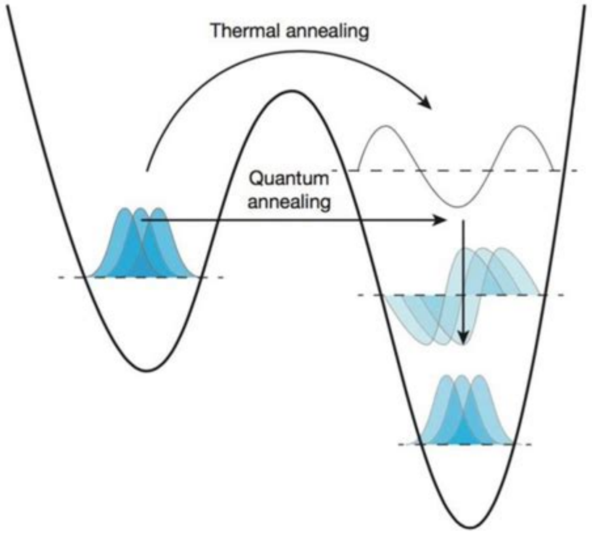
$$\chi^2 = \sum_{ij} V_a C^{-1}_{ab} V_b \qquad V_a = O_a^{(\text{exp})} - O_a^{(\text{th})}(c)$$

- 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 Computing is exciting research area that rapidly expands, supported through private and public sector. Many methods to be invented.
  ➡ Can exploit QM prop: entanglement, superposition principle and tunnelling

- Nature is inherently quantum mechanical, thus quantum simulation should be advantageous
  ➡ Analogue quantum devices provide an environment that evolves quantum mechanically. The task is to encode problems into such an environment.

- For quantum advantage in real-world applications need development of technical realisation of quantum computers (size, fault tolerance, type of operations,...)