Please, feed the quantum troll!

Alba Cervera-Lierta
University of Toronto
Q-Hack 2021

the matter lab

Menu

Appetizer

Variational Quantum Algorithms

Starter

Meta-VQE

Main

Quantum Classifier

Desert

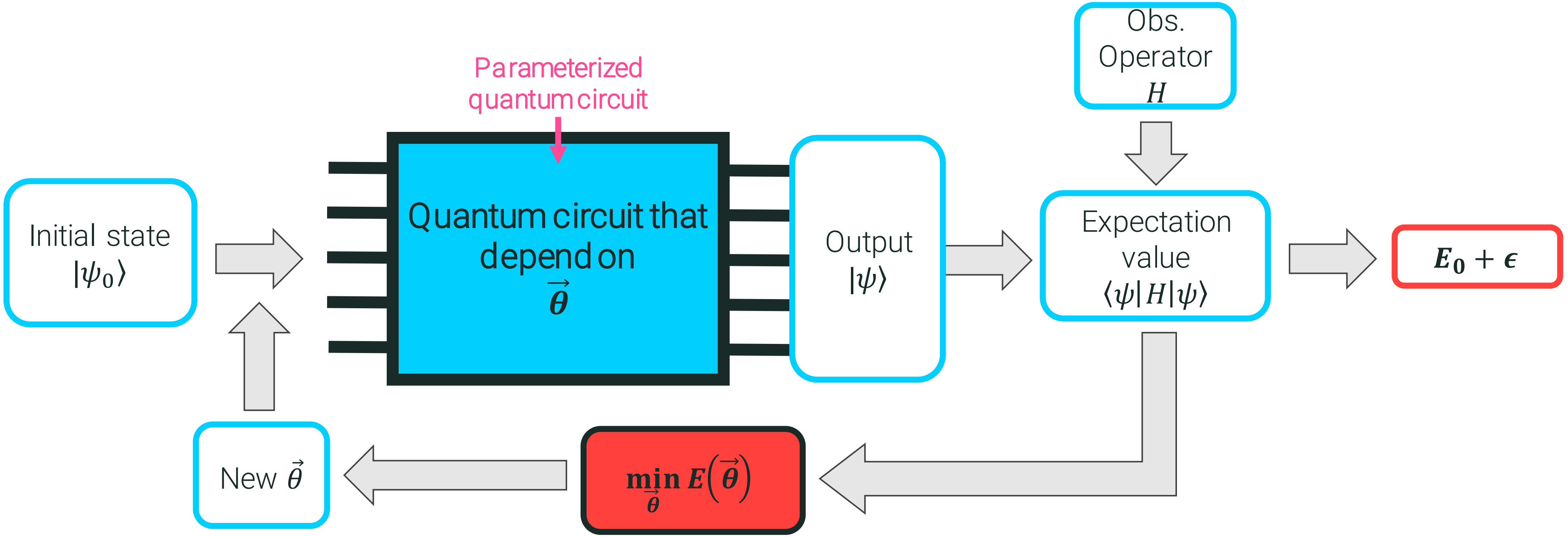
A quantum fitting algorithm

Some recipes at

https://github.com/AlbaCL/qhack21

Variational Quantum Algorithms

e.g. Variational Quantum Eigensolver, QAOA, Classifier, Autoencoder,...



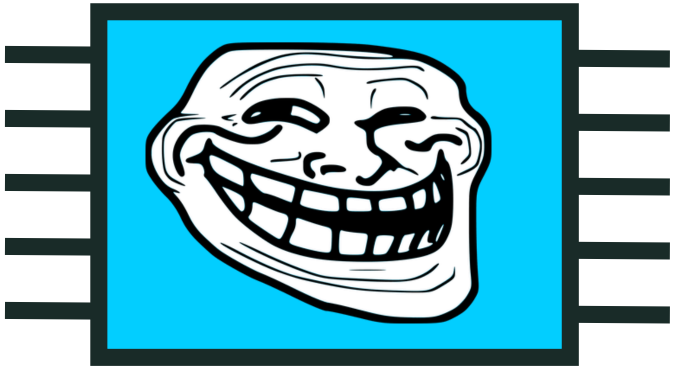
Variational principle: $\mathrm{E} = \langle\psi|H|\psi\rangle \geq E_0$

Barren Plateaus

If the parameters of the quantum circuit are initialized at random, gradients and variances of the circuit w.r.t. parameters vanish!!

The Barren Plateau problem

J. R. McClean, S. Boixo, V. N. Smelyanskiy, R. Babbush, H. Neven, Nature Commun. 9, 4812 (2018)



Some of the proposed solutions to alleviate the problem:

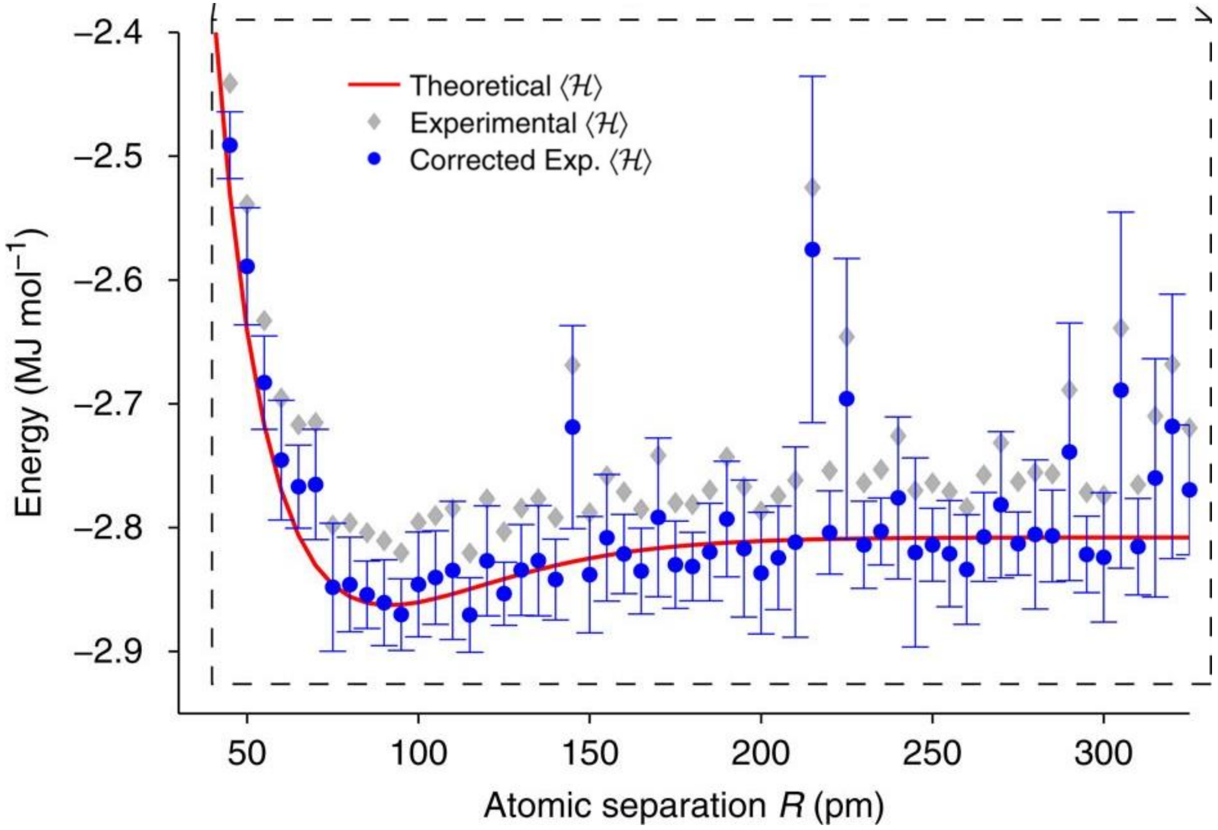
Initialize wisely:
E. Grant, L. Wossnig, M. Ostaszewski, M. Benedetti, Quantum 3, 214 (2019).

Local cost functions:
M. Cerezo, A. Sone, T. Volkoff, L. Cincio, P. J. Coles, arXiv:2001.00550 (2020).

Correlate parameters:
T. Volkoff, P. J. Coles, Quantum Sci. Technol. 6, 025008 (2021).

What's the goal of a VQE?

Bond dissociation curve of the He–H$^+$ molecule.



Hamiltonian that can be written with Pauli strings

$$\langle \mathcal{H} \rangle = \sum_{i\alpha} h^i_\alpha \langle \sigma^i_\alpha \rangle + \sum_{ij\alpha\beta} h^{ij}_{\alpha\beta} \langle \sigma^i_\alpha \sigma^j_\beta \rangle + \ \dots$$

Outputs of the quantum computer

Quantum circuit that generates the ground state of that Hamiltonian

e.g. Hartree-Fock

$$|\Psi\rangle = e^{T - T^\dagger} \, |\Phi\rangle_{\text{ref}} .$$

$$T = T_1 + T_2 + T_3 + \ldots + T_N ,$$

Unitary operation, e.g. Cluster operator

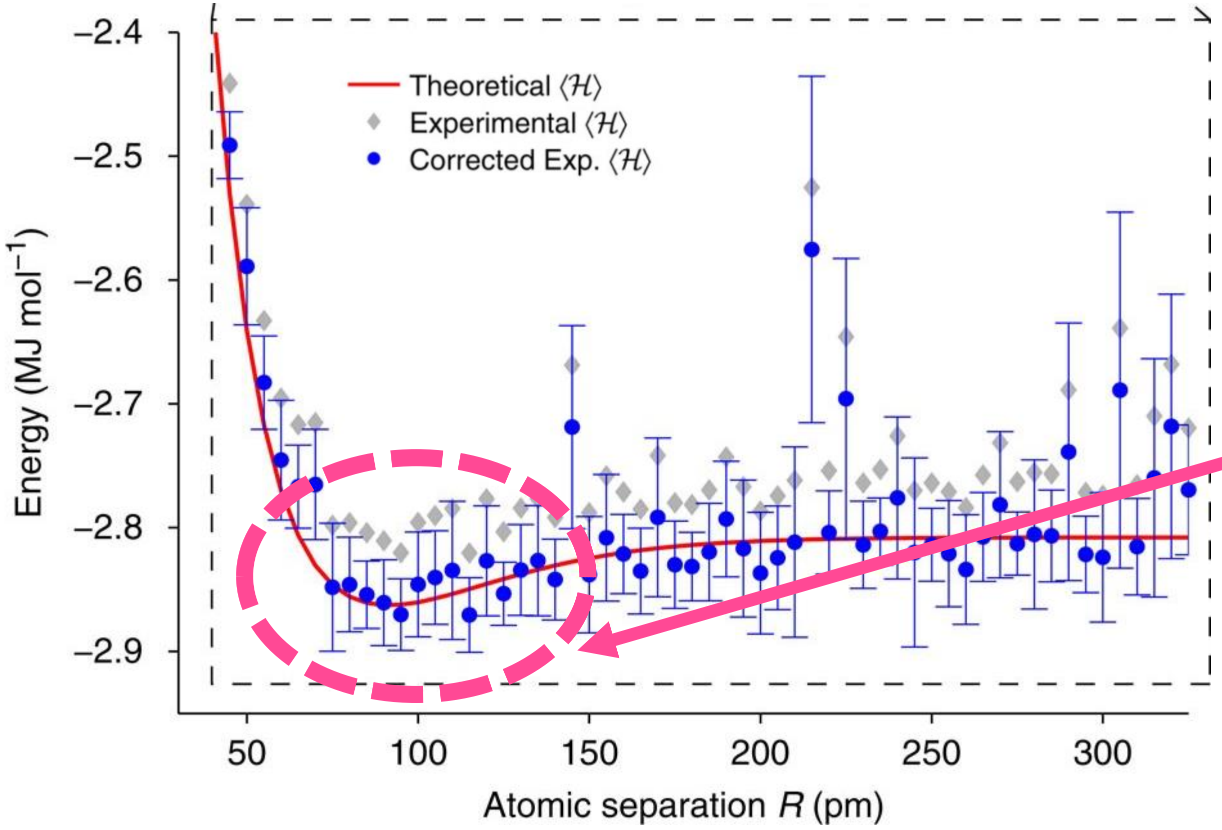
<u>GOAL</u>: find $|\psi\rangle$ that minimizes

$$\frac{\langle \psi | \mathcal{H} | \psi \rangle}{\langle \psi | \psi \rangle} .$$

A. Peruzzo, J. McClean, P. Shadbolt, M.-H. Yung, X.-Q. Zhou, P. J. Love, A. Aspuru-Guzik, J. L. O'Brien, Nature Comm. 5, 4213 (2014)

What's the *true* goal of VQE?

Bond dissociation curve of the He–H⁺ molecule.



Find the atomic separation that minimizes the energy
$\min\langle H(R)\rangle$

To obtain this you need to scan from 0 to 300.

Each blue point is a VQE, that is, you have to prepare, run and optimize the quantum circuit.

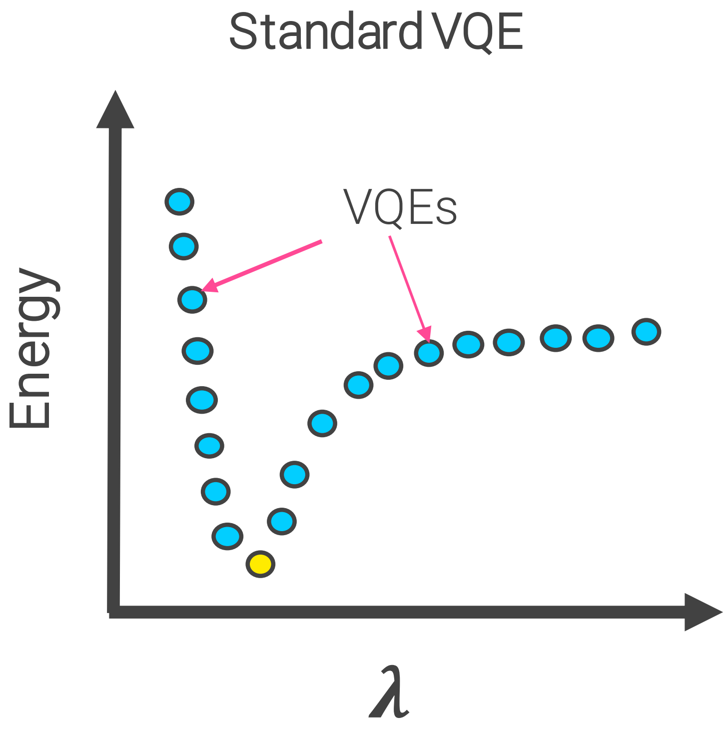
Can we avoid to compute these uninteresting points?

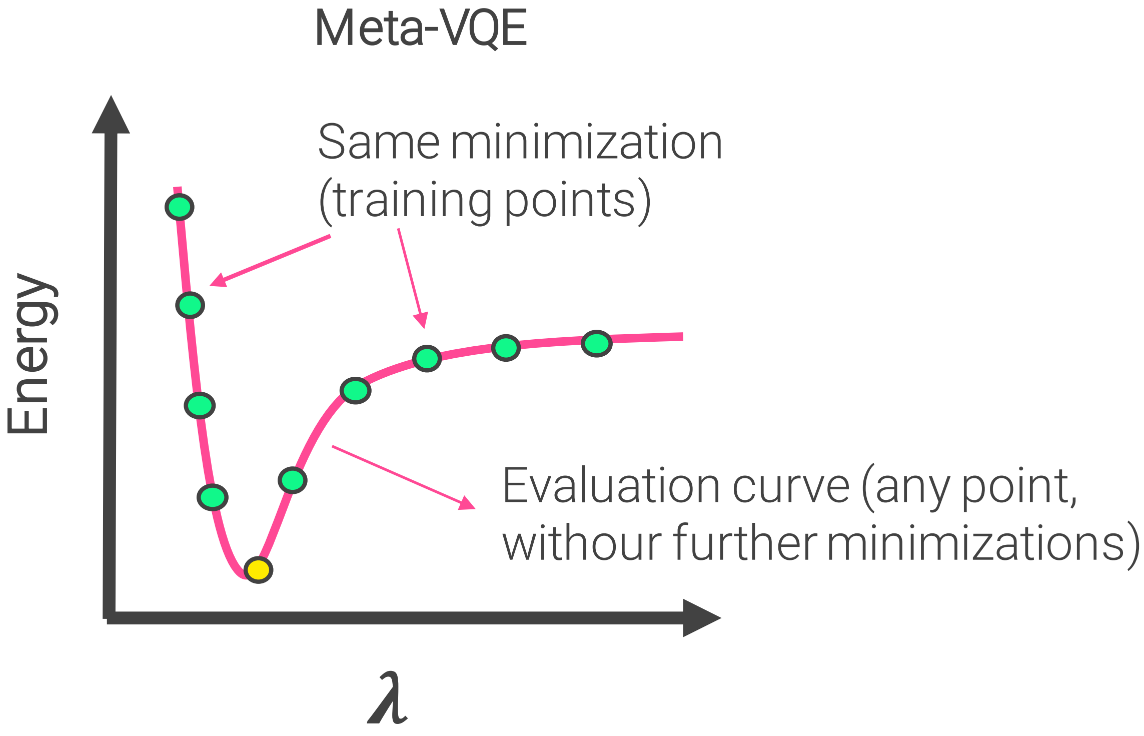
A. Peruzzo, J. McClean, P. Shadbolt, M.-H. Yung, X.-Q. Zhou, P. J. Love, A. Aspuru-Guzik, J. L. O'Brien, Nature Comm. 5, 4213 (2014)

The Meta-VQE

Parameterized Hamiltonian $H(\vec{\lambda})$

<u>Goal:</u> to find the quantum circut that encodes the ground state of the Hamiltonian for any value of $\vec{\lambda}$





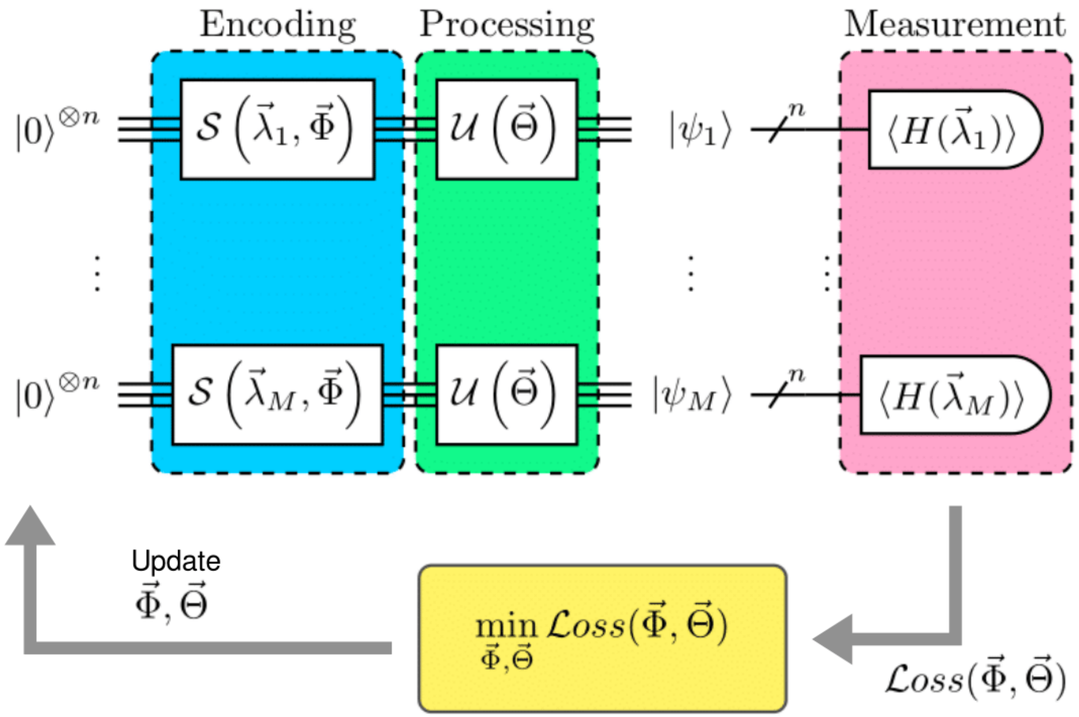
The Meta-VQE

Feed the troll with the $\vec{\lambda}$!

Parameterized Hamiltonian $H\left(\vec{\lambda}\right)$

Training points: $\vec{\lambda}_i$ for $i = 1, \dots, M$

Loss function with all $\left\langle H(\vec{\lambda}_i) \right\rangle$



<u>Goal:</u> to find the quantum circut that encodes the ground state of the Hamiltonian for any value of $\vec{\lambda}$

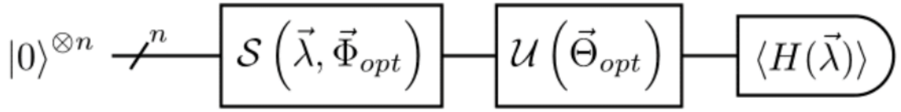
Output: $\vec{\Phi}_{opt}$ and $\vec{\Theta}_{opt}$

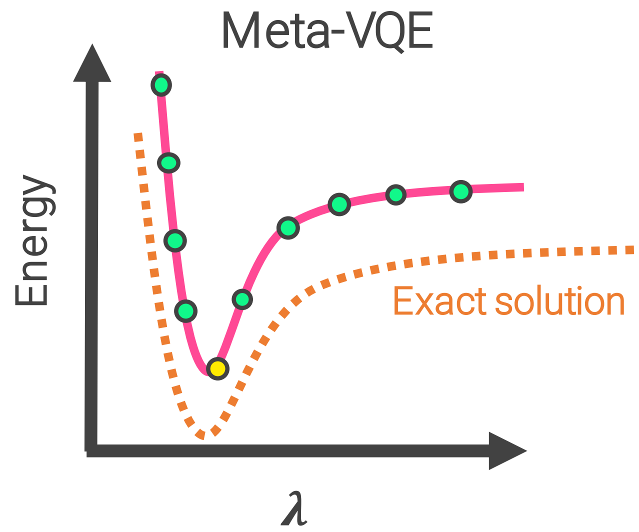
See also: K. Mitarai, T. Yan, K. Fujii, Phys. Rev. Applied **11**, 044087 (2019)

The Meta-VQE output

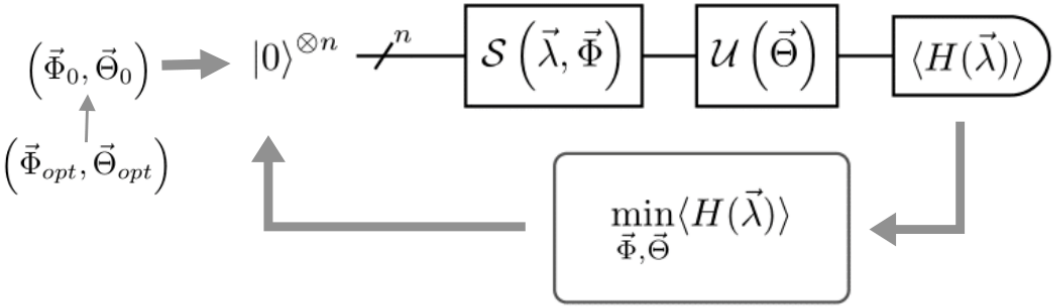
Output: $\vec{\Phi}_{opt}$ and $\vec{\Theta}_{opt}$

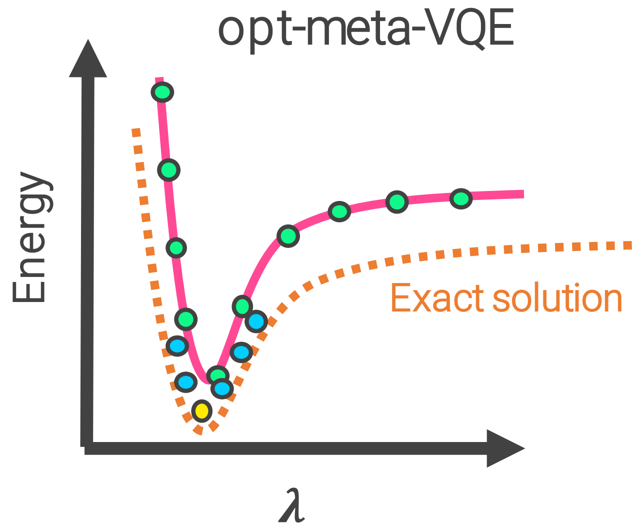
Option 1: run the circuit with test $\vec{\lambda}$ and obtain the g.s. energy profile.





Option 2: use $\vec{\Phi}_{opt}$ and $\vec{\Theta}_{opt}$ as starting point of a standard VQE optimization (*opt-meta-VQE*)





Example: H_4 molecule

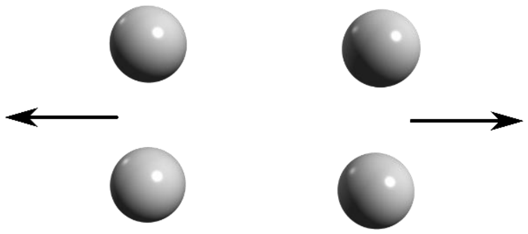
H_4 molecule in 8 spin-orbitals (STO-3G basis set)

Ansatz: k-UpCCGSD (k=2 for these results)

Linear encoding: $\theta = \alpha + d\beta$

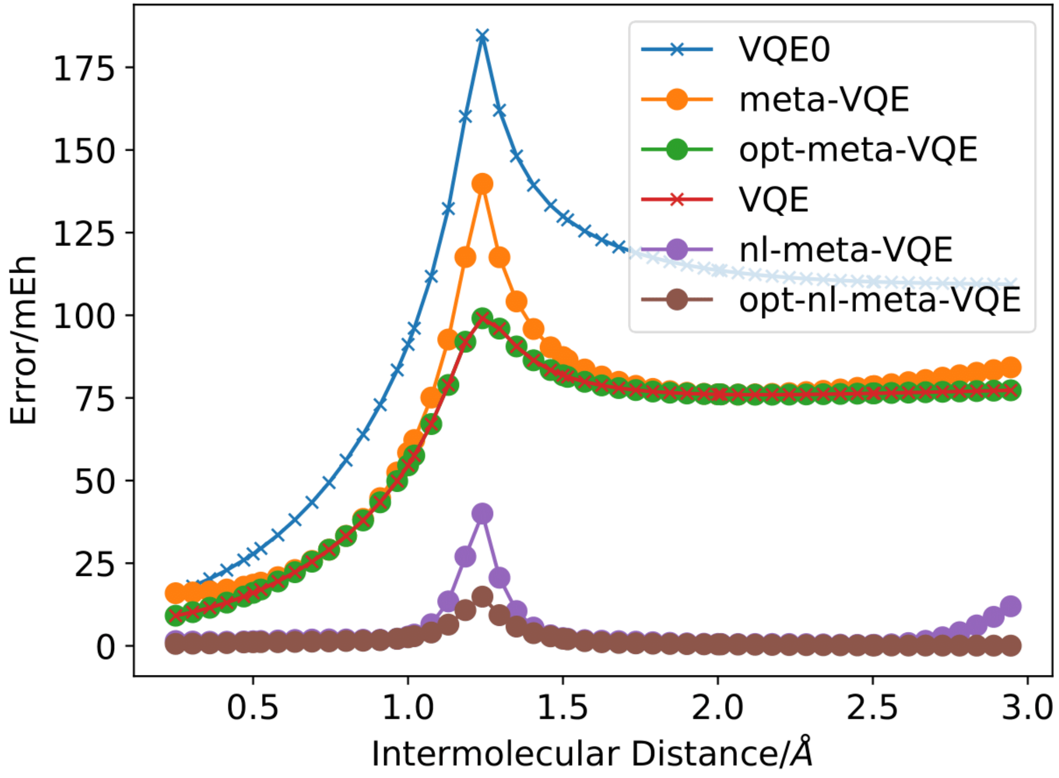
Hamiltonian Parameter
(intermolecular distance)

Non-linear encoding: $\theta = \alpha e^{\beta(\gamma - d)} + \delta$ (floating Gaussians)



Check the code tutorial for details!

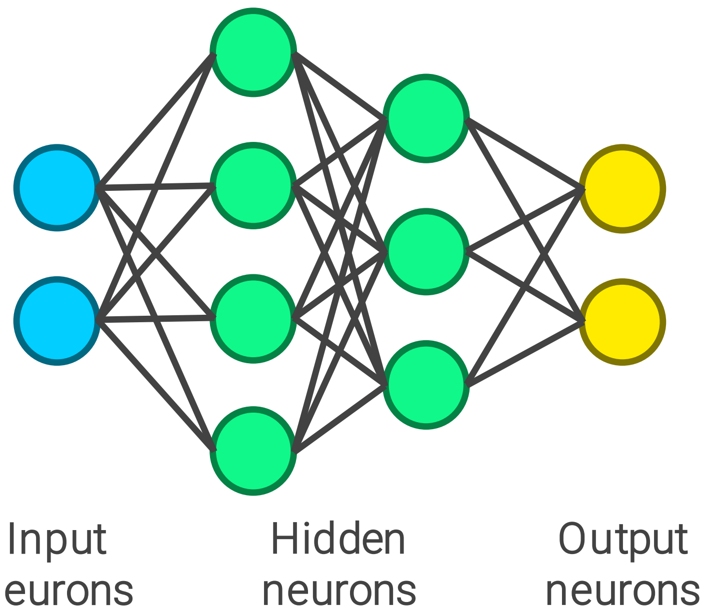
https://github.com/AlbaCL/qhack21



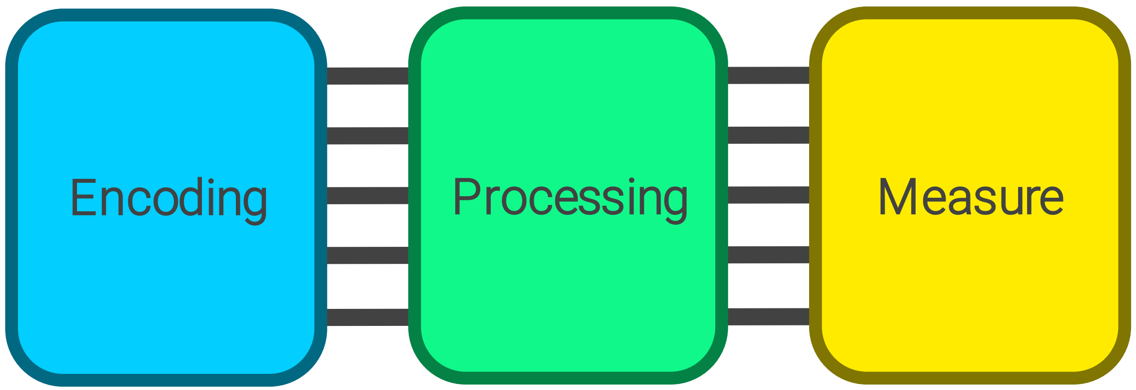
ACL, J. S. Kottmann and A. Aspuru-Guzik, arXiv:2009.13545 [quant-ph] (2020)

From classical to quantum NN

Classical



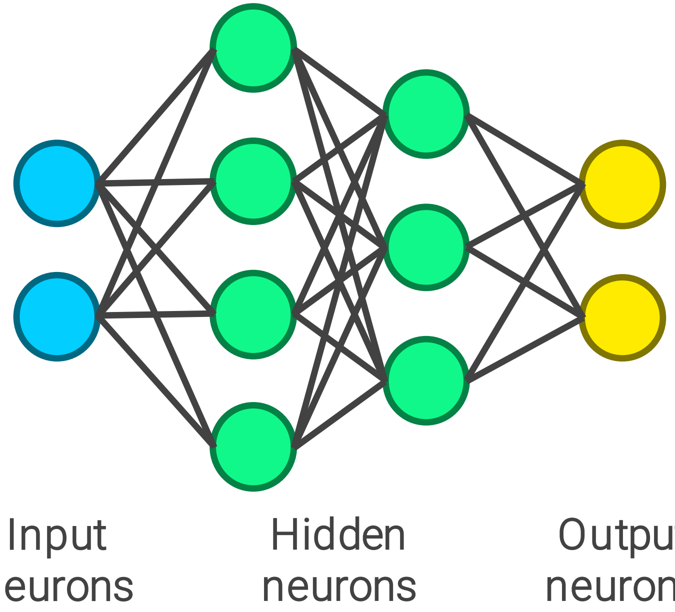
Quantum
(circuit centric)



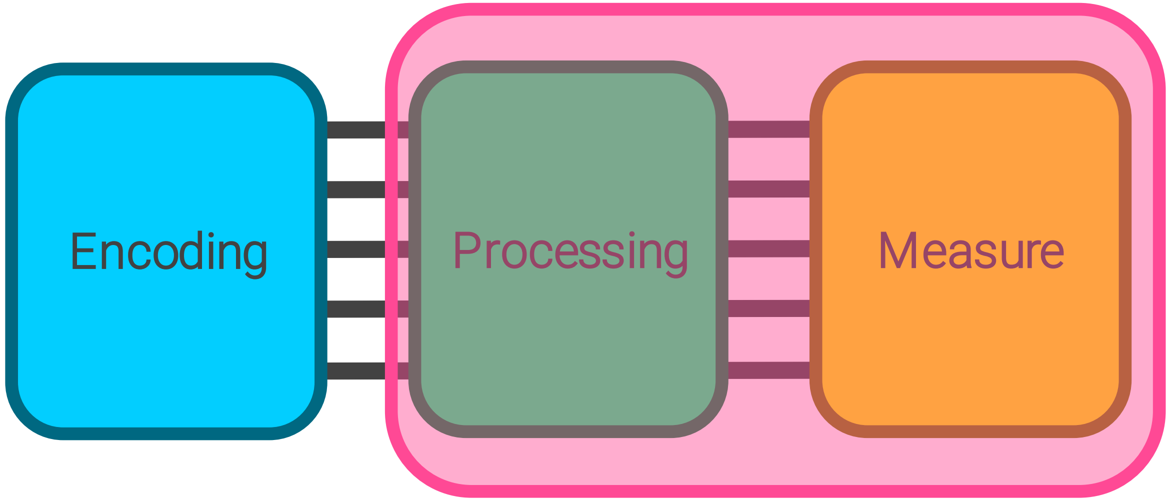
K Mitarai, M Negoro, M Kitagawa, K Fujii Phys. Revs A 98 (3), 032309 (2018)
E. Farhi and H.Neven, arXiv:1802.06002 (2018)
M. Schuld and N. Killoran, Phys. Rev. Lett. 122, 040504 (2019)
M. Schuld, A. Bocharov, K. M. Svore, and N. Wiebe, Phys. Rev. A 101, 032308 (2020)

From classical to quantum NN

Classical



Quantum
(circuit centric)



Find the correct measurement basis

M. Schuld, arXiv:2101.11020 [quant-ph] (2021)

K Mitarai, M Negoro, M Kitagawa, K Fujii Phys. Revs A 98 (3), 032309 (2018)

E. Farhi and H.Neven, arXiv:1802.06002 (2018)

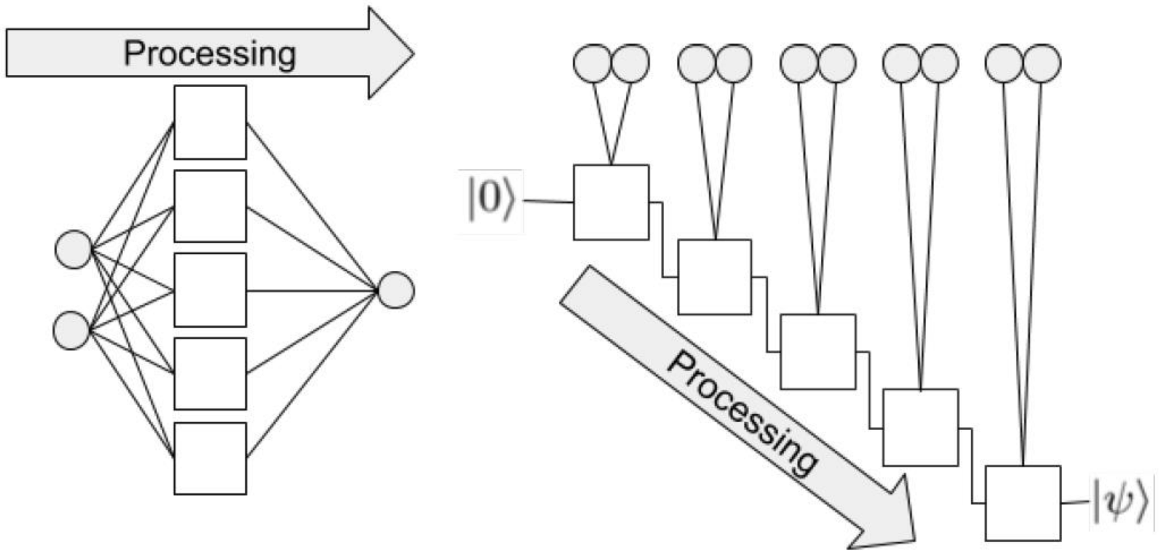
M. Schuld and N. Killoran, Phys. Rev. Lett. 122, 040504 (2019)

M. Schuld, A. Bocharov, K. M. Svore, and N. Wiebe, Phys. Rev. A 101, 032308 (2020)

The minimal QNN

What is the most simple (but universal) NN? → Single hidden layer NN

What is the most simple (but universal) QNN? → Single-qubit QNN



(a) Neural network (b) Quantum classifier

Encoding the data

A product of free single-qubit unitaries can be written with another single-qubit unitary

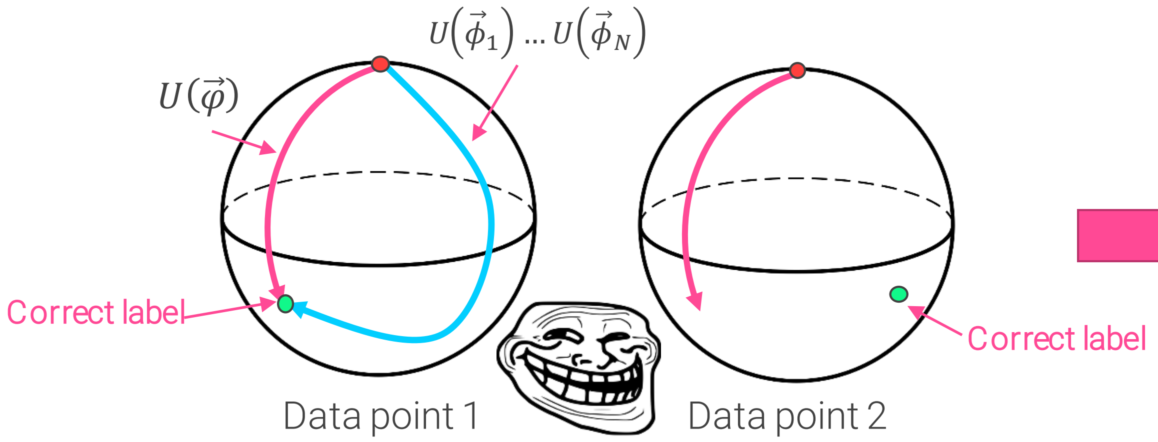
Insufficient to carry out any non-trivial task

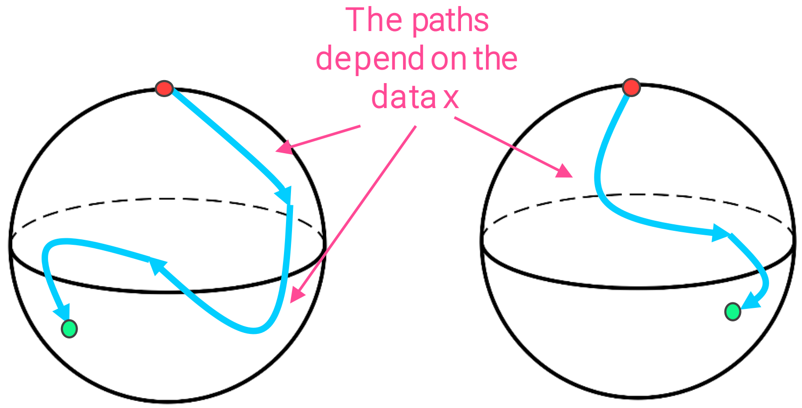
$$U\left(\vec{\phi}_1\right) \dots U\left(\vec{\phi}_N\right) \equiv U(\vec{\varphi})$$

If we add some fixed parameter dependency (the data), the operation becomes flexible and data-depedent.

Data re-uploading

$$\mathcal{U}(\vec{\phi}, \vec{x}) \equiv U(\vec{\phi}_N)U(\vec{x}) \dots U(\vec{\phi}_1)U(\vec{x})$$





A. Pérez-Salinas, ACL,, E. Gil-Fuster and J. I. Latorre, Quantum 4, 226 (2020), arXiv:1907.02085

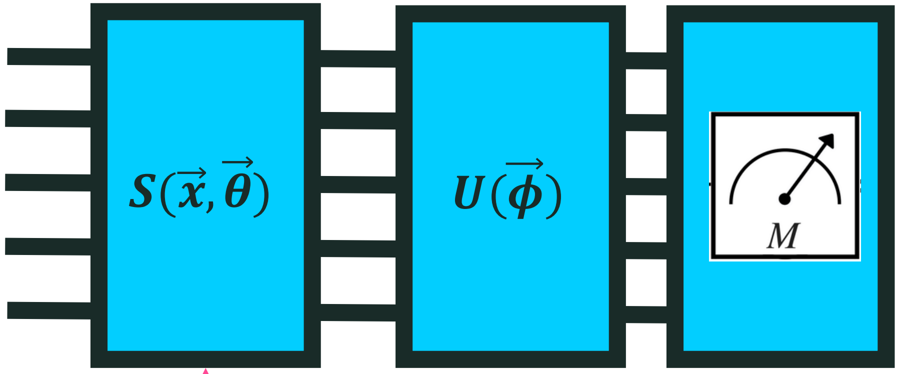
Quantum Feature Maps

$$|\psi_0\rangle \rightarrow |\psi(\vec{x},\vec{\theta})\rangle \rightarrow |\psi(\vec{x},\vec{\theta},\vec{\phi})\rangle \rightarrow |gs\,(\vec{x})\rangle$$

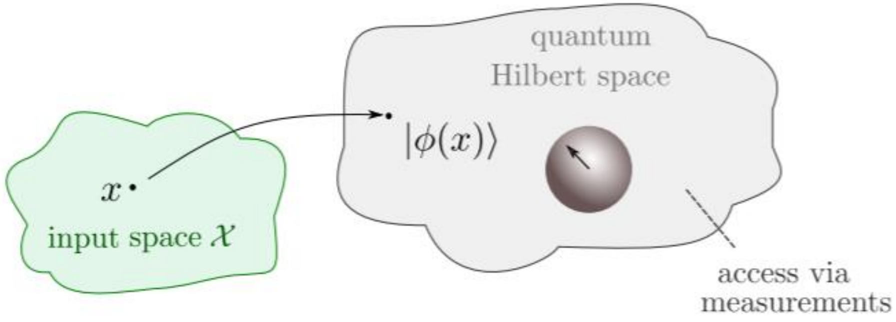
Encode the data

Rotate to the correct measurement basis

Find the parameters that minimze the energy (measured in the computational basis)



Feed the troll with the $\vec{x}$!



Data re-uploading layers

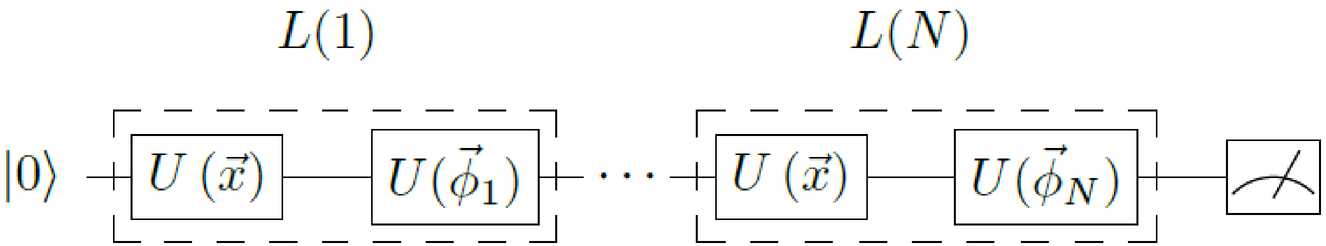
The total unitary is divided into layers.
Each layer encodes the data.

$$\mathcal{U}(\vec{\phi}, \vec{x}) = L(N) \dots L(1)$$

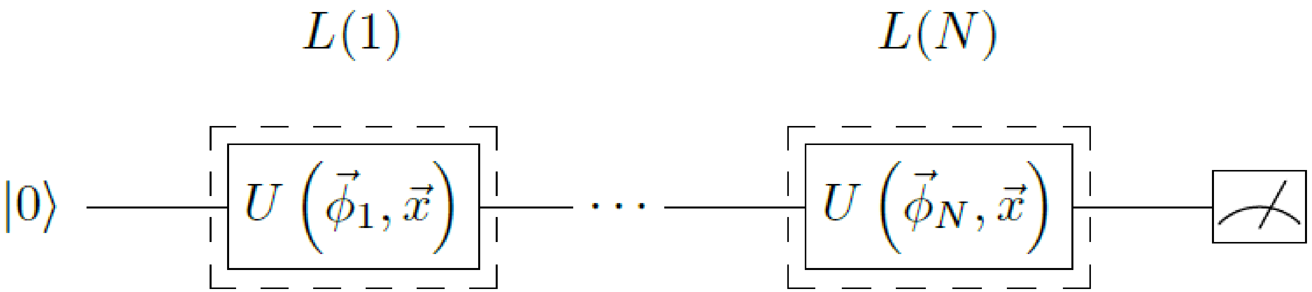
$$L(i) \equiv U(\vec{\phi}_i) U(\vec{x})$$

Single operation

$$L(i) = U\left(\vec{\theta}_i + \vec{w}_i \circ \vec{x}\right)$$

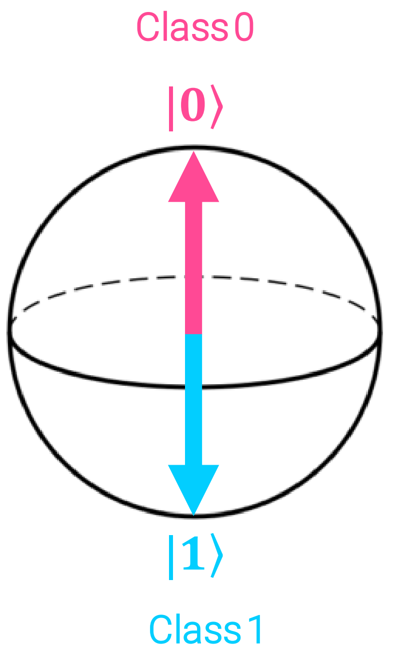


(a) Original scheme



(b) Compressed scheme

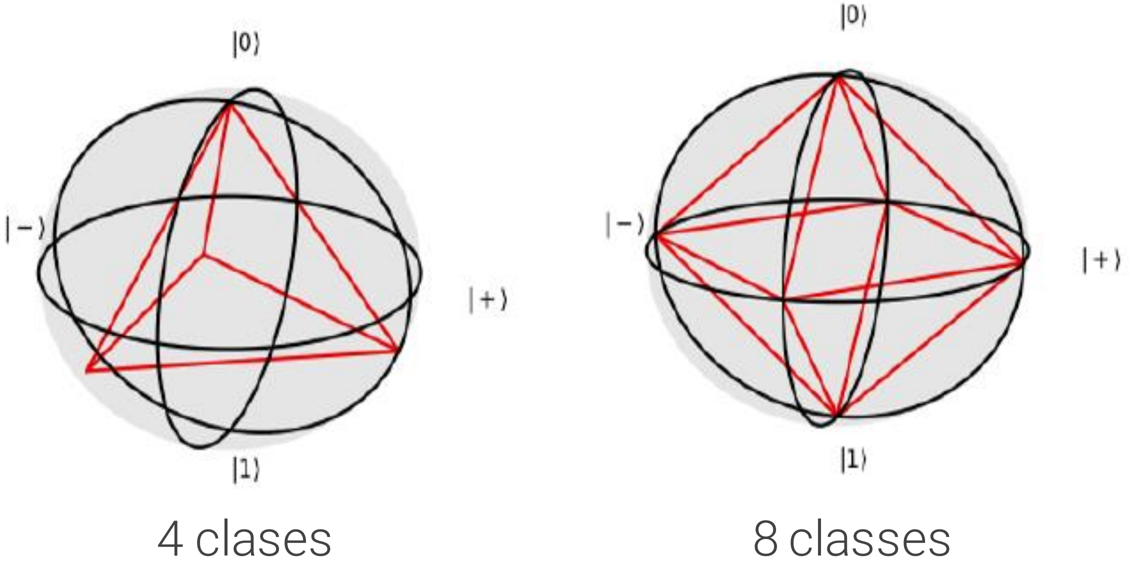
Target states



Convenient: choose the most ortogonal states to define each target state.

Single-qubit → Divide the Bloch sphere into Nclass sections

C. W. Helstrom, *Quantum detection and estimation theory*, Academic Press New York (1976).



Extension for multi-qubits:
S. Lloyd, M. Schuld, A. Ijaz, J. Izaac, N. Killoran, arXiv:2001.03622

Measurement and cost function

Target state: one for each label/class.

Compute the fidelity (overlap) between the quantum circuit state and the target state

Training points

Circuit state wavefunction

$$\chi_f^2(\vec{\theta}, \vec{w}) = \sum_{\mu=1}^{M} \left(1 - |\langle \tilde{\psi}_s | \psi(\vec{\theta}, \vec{w}, \vec{x}_\mu) \rangle|^2 \right)$$

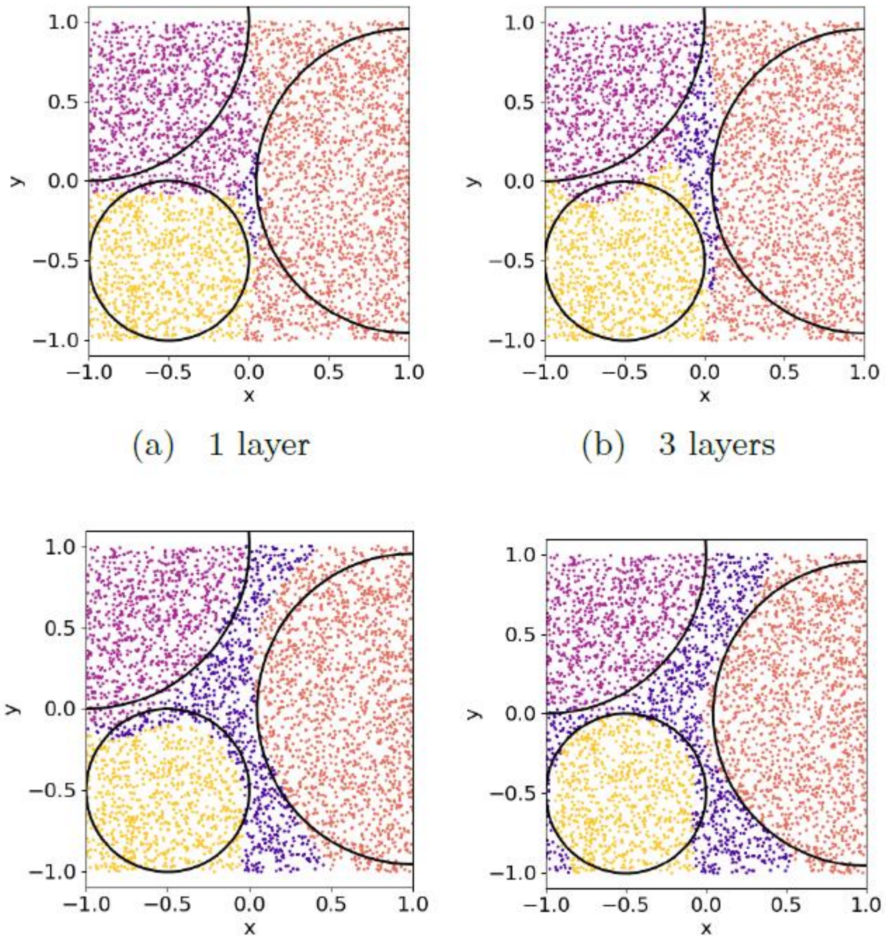
Target state wavefunction

Weighted fidelity (for multiclassification):
compute the overlap w.r.t. target state – distance w.r.t. other class target state

classes

$$\chi_{wf}^2(\vec{\alpha}, \vec{\theta}, \vec{w}) = \frac{1}{2} \sum_{\mu=1}^{M} \left(\sum_{c=1}^{\mathcal{C}} \left(\alpha_c F_c(\vec{\theta}, \vec{w}, \vec{x}_\mu) - Y_c(\vec{x}_\mu) \right)^2 \right)$$

Example 2D data: 3 circles



(a) 1 layer

(b) 3 layers

(c) 4 layers

(d) 10 layers

Qubits	χ^2_f			χ^2_{wf}				
	1	2		1	2		4	
Layers		No Ent.	Ent.		No Ent.	Ent.	No Ent.	Ent.
1	0.73	0.56	–	0.75	0.81	–	0.88	–
2	0.79	0.77	0.78	0.76	0.90	0.83	0.90	0.89
3	0.79	0.76	0.75	0.78	0.88	0.89	0.90	0.89
4	0.84	0.80	0.80	0.86	0.84	0.91	0.90	0.90
5	0.87	0.84	0.81	0.88	0.87	0.89	0.88	0.92
6	0.90	0.88	0.86	0.85	0.88	0.89	0.89	0.90
8	0.89	0.85	0.89	0.89	0.91	0.90	0.88	0.91
10	0.91	0.86	0.90	0.92	0.90	0.91	0.87	0.91

Training/test points = 200/4000
Random accuracy = 25%

A. Pérez-Salinas, ACL,, E. Gil-Fuster and J. I. Latorre, Quantum 4, 226 (2020), arXiv:1907.02085

Fitting functions with a quantum circuit

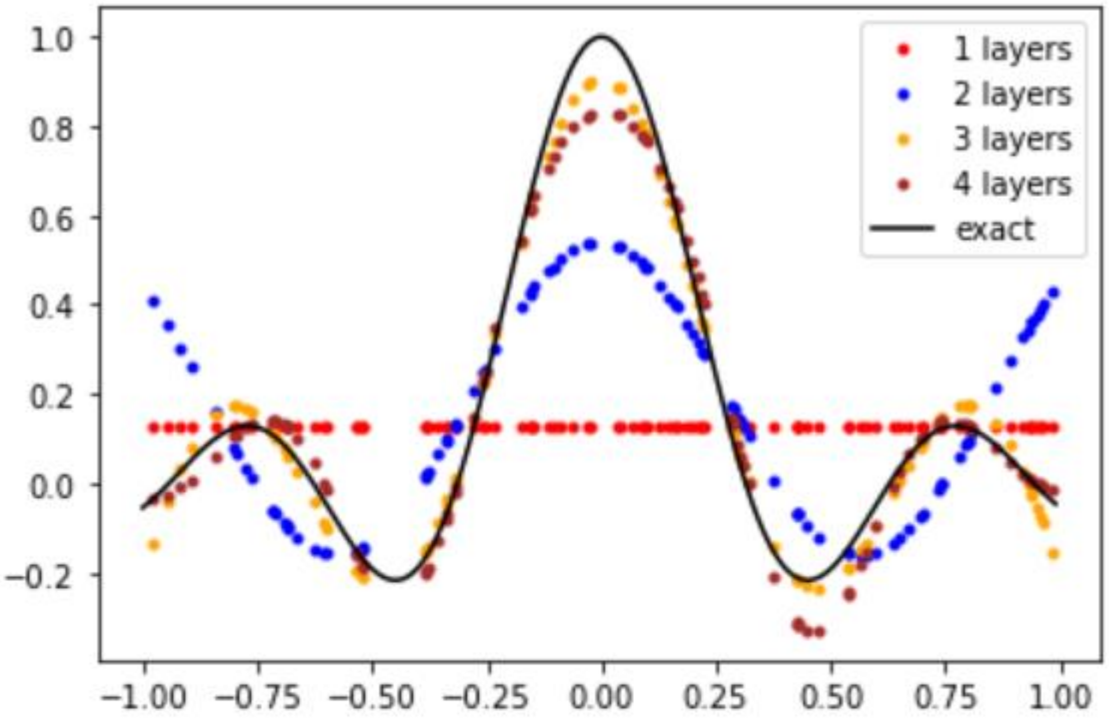
Quantum circuits can be theoretically written as partial Fourier series and, therefore, they can be universal function approximators. The more data re-uploading, the more precision can be achieved.

Example: one qubit with linear encoding layers and cost function:

$$\chi^2 = \frac{1}{M}\sum_{j=1}^{M}\left(\langle Z(x_j)\rangle - f(x_j)\right)^2$$

Check the code tutorial for details!

https://github.com/AlbaCL/qhack21



M. Schuld, R. Sweke, J. J. Meyer, arXiv:2008.08605 [quant-ph]

Overview

By using a variational strategy, we have to be careful to avoid the Barren Plateau problem, so we can optimize our loss function without getting stuck in local minima.

Parameter correlation, as the one introduced by a data-reuploading strategy, may alliviate the vanishing gradients problem.

Designing a good feature map (a.k.a. data encoding) is crucial for QML algorithms.

One qubit can approximate any continous function and be used as a multi-class classifier.

By using the encoding strategies from QML, we can improve variational quantum algorithms such as VQE to reduce the number of optimizations by learning the ground state energy profile.

Some questions

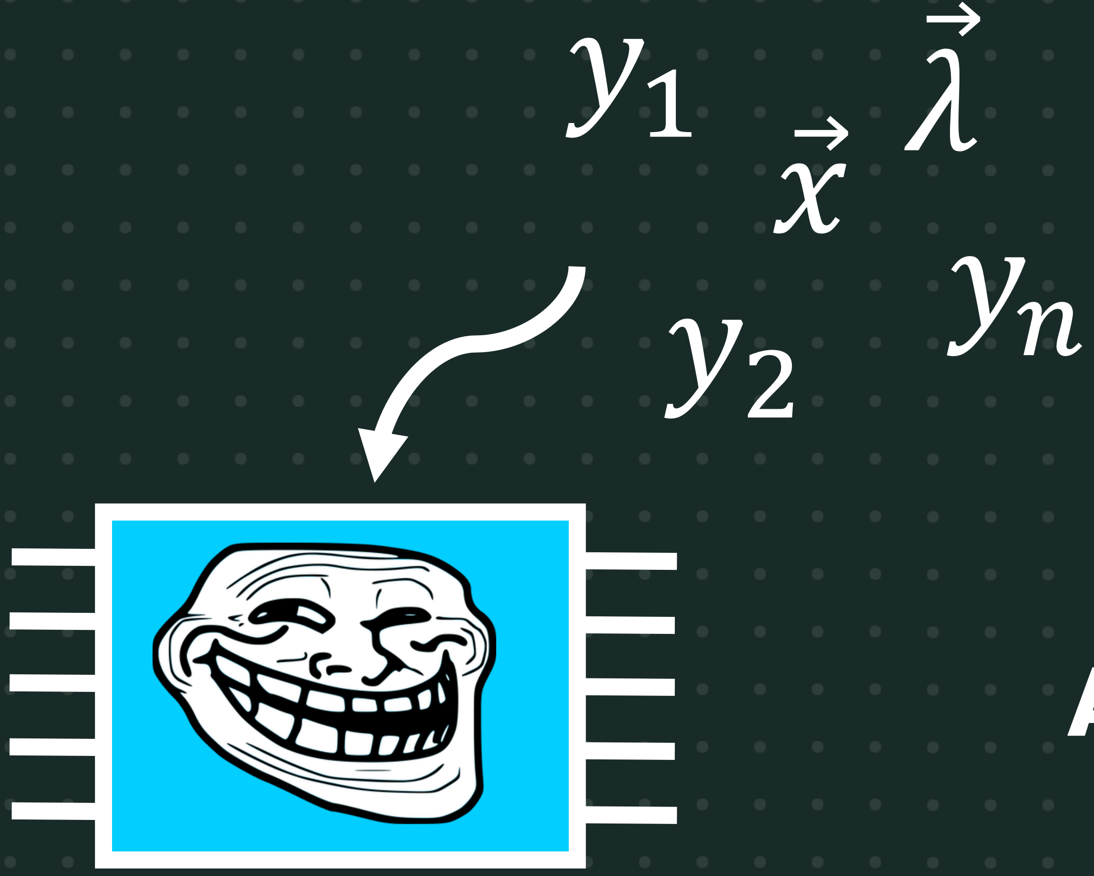
- What about quantum data? Can we adapt these algorithms to be feed with quantum instead of classical data?

- What happens exactly when we use multiple qubits? Can we use less parameters to obtain the same accuracy in the results?

- What is the best circuit ansatz (entangling gates, etc)?

Check the references and code tutorials for details!

https://github.com/AlbaCL/qhack21



Thanks!

And don't forget to feed the quantum troll!