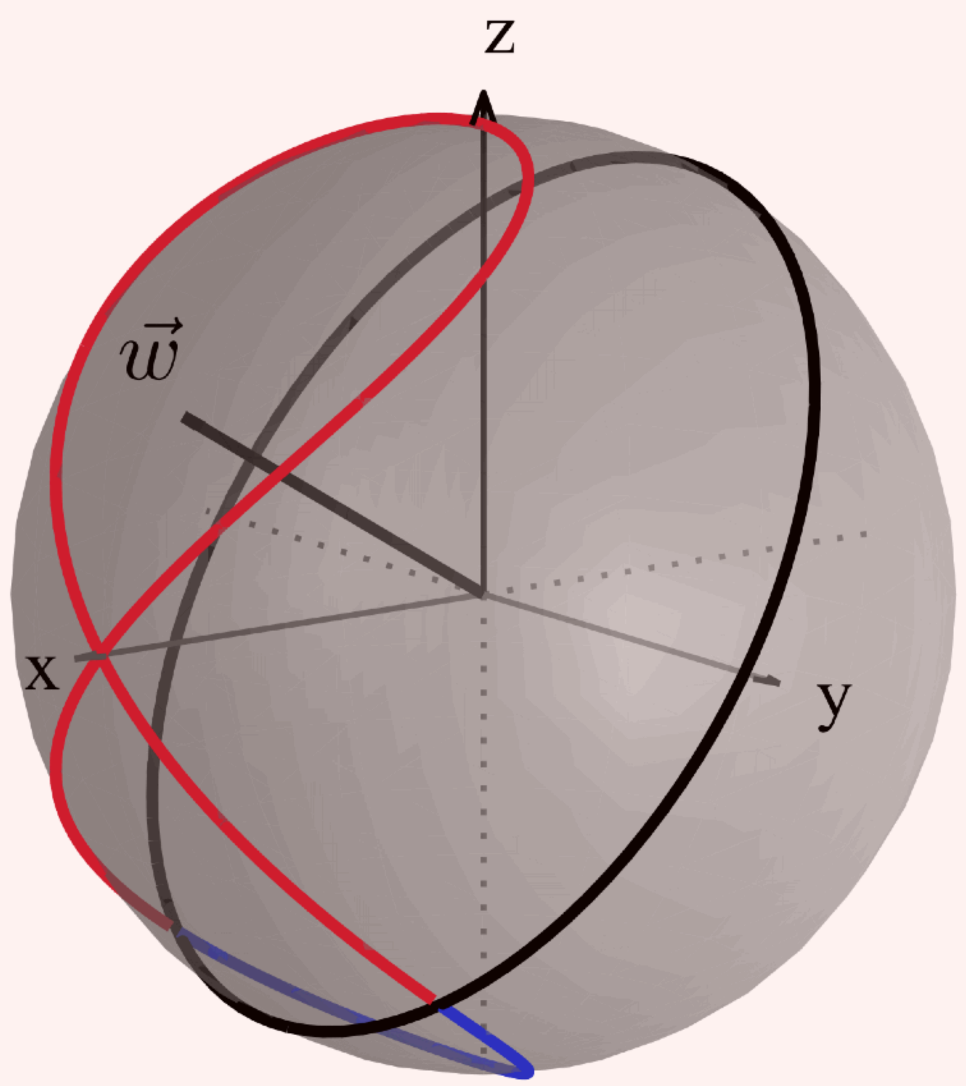


Basic programming for quantum machine learning with Qiskit and PennyLane

NITheP Mini-school

UNIVERSITY OF KWAZULU-NATAL ™
INYUVESI YAKWAZULU-NATALI

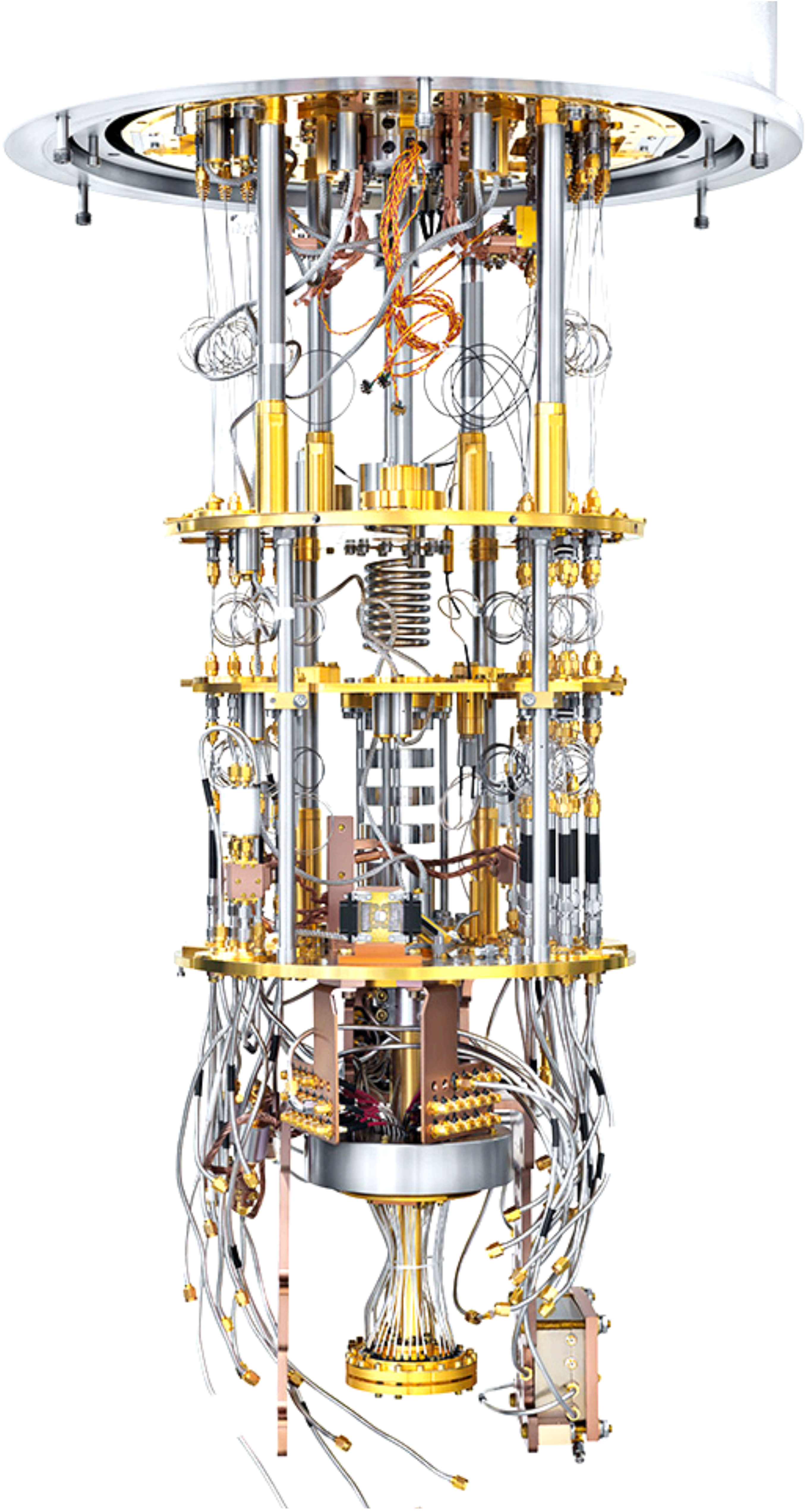
Amira Abbas
29/09/2020

NITheP
National Institute for Theoretical Physics

Thank you!

Previously

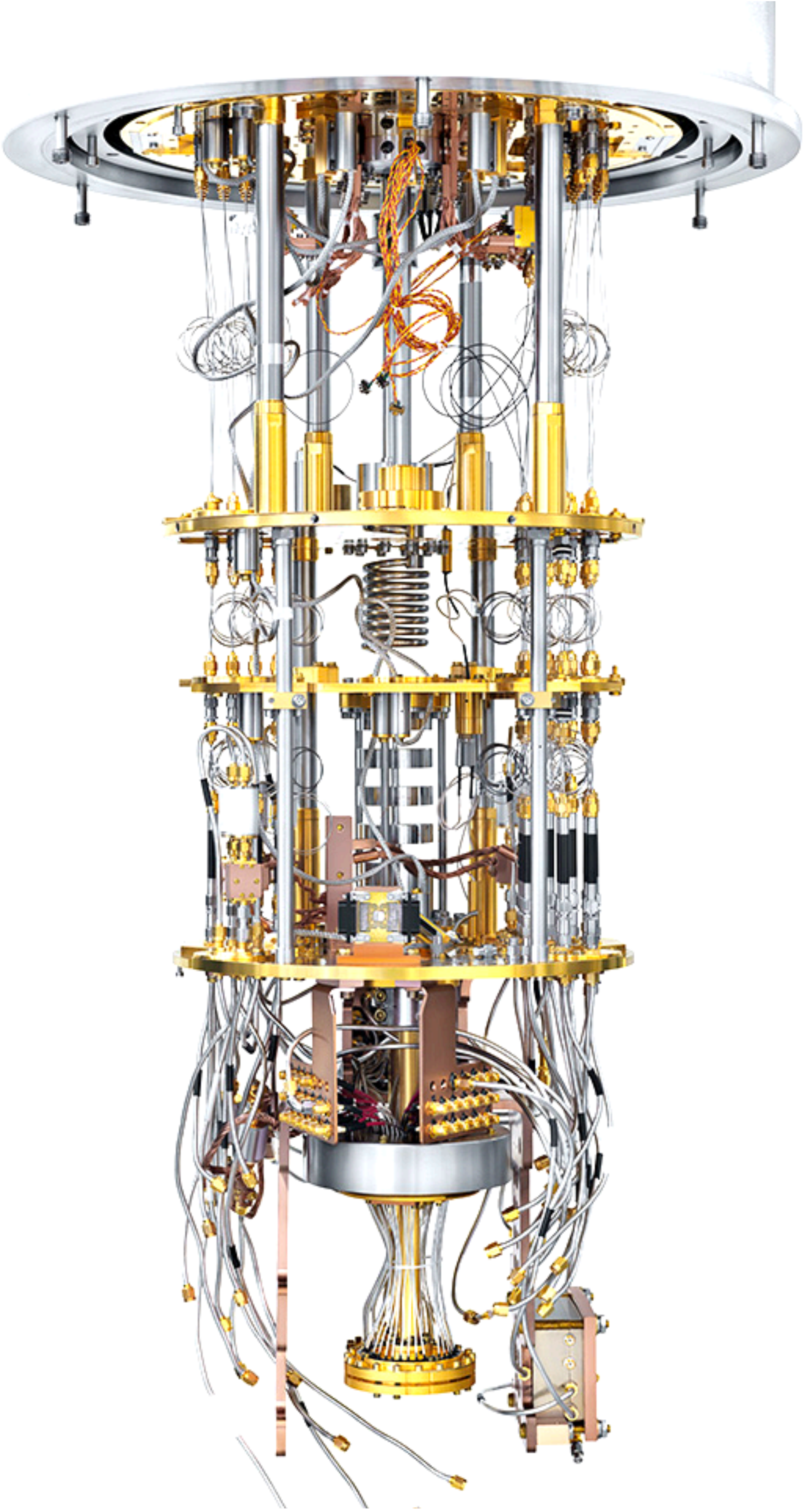
Quantum models



Previously

Quantum models

Coded a variational quantum classifier

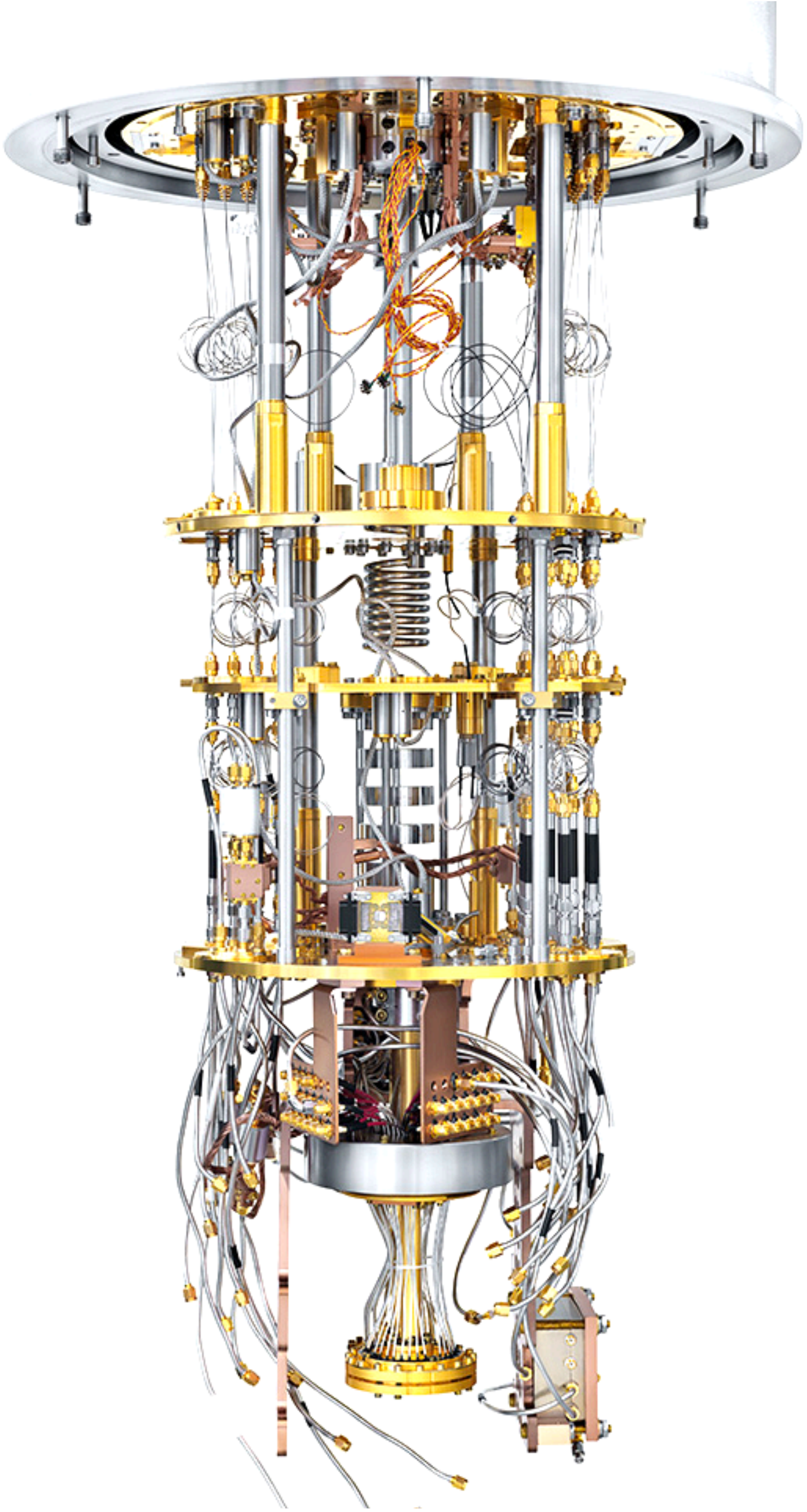


Agenda

Quantum models

Coded a variational quantum classifier

Barren plateaus



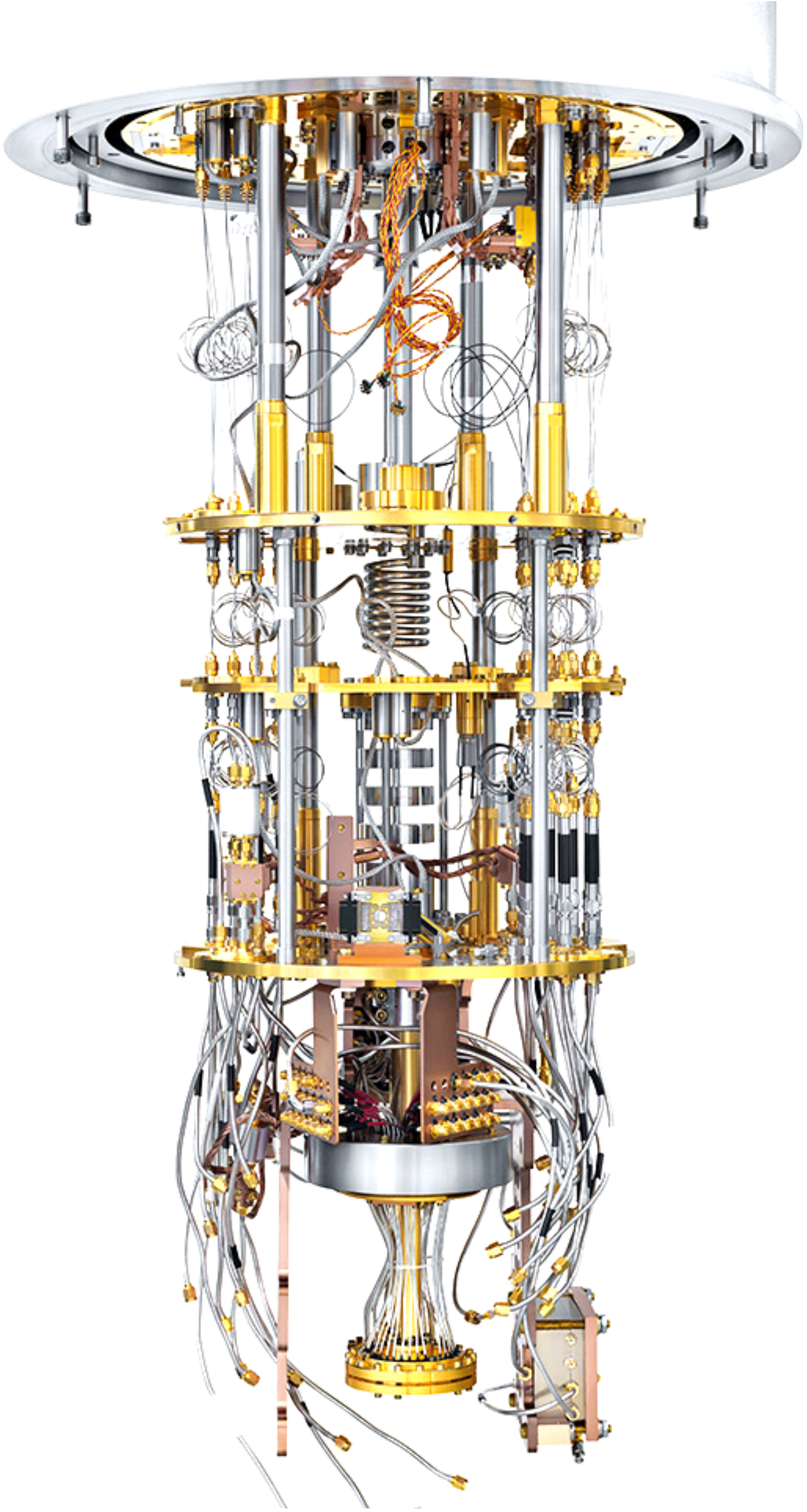
Agenda

Quantum models

Coded a variational quantum classifier

Barren plateaus

Quantum neural networks



Agenda

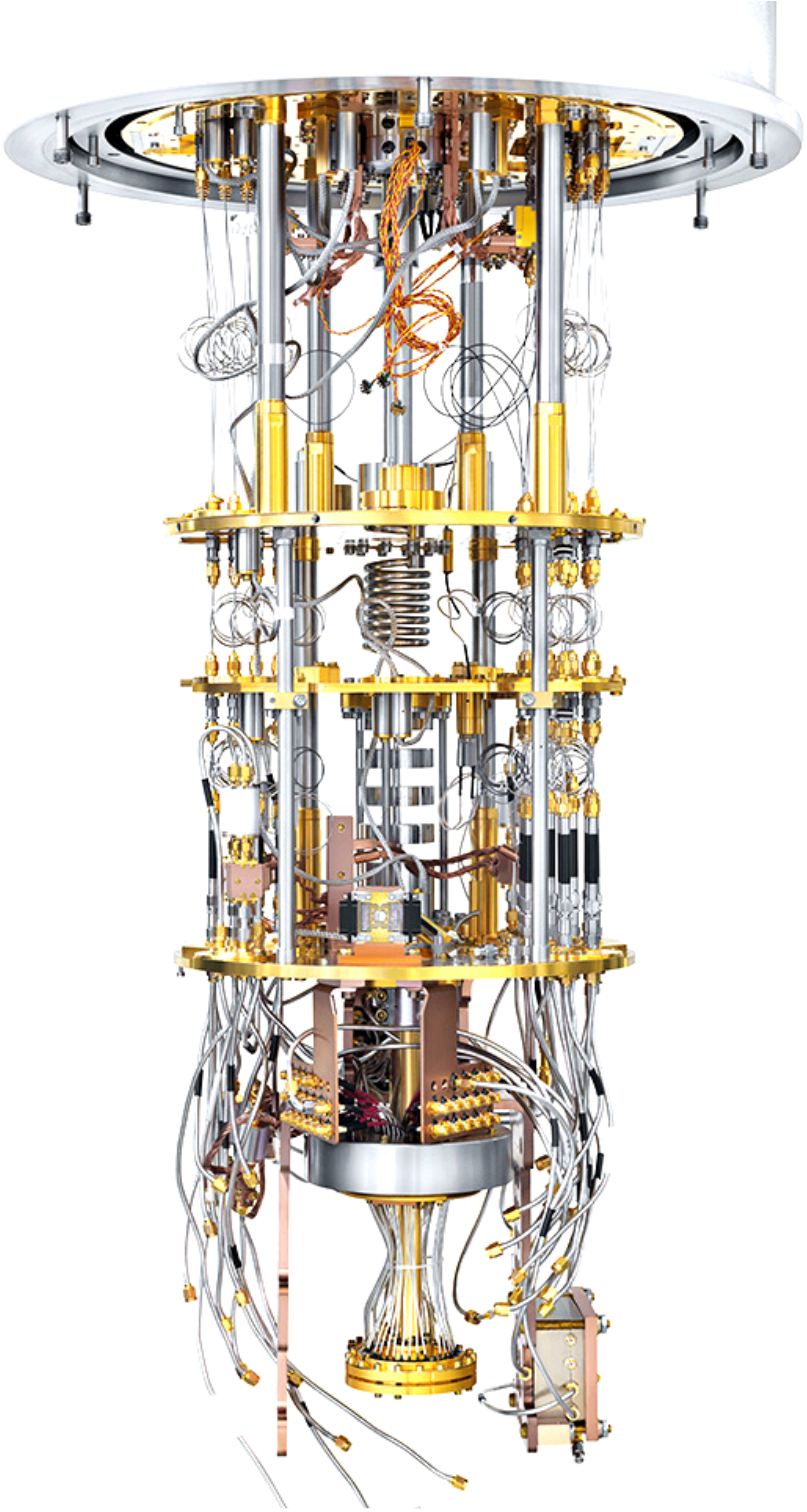
Quantum models

Coded a variational quantum classifier

Barren plateaus

Quantum neural networks

Loss landscapes



Agenda

Quantum models

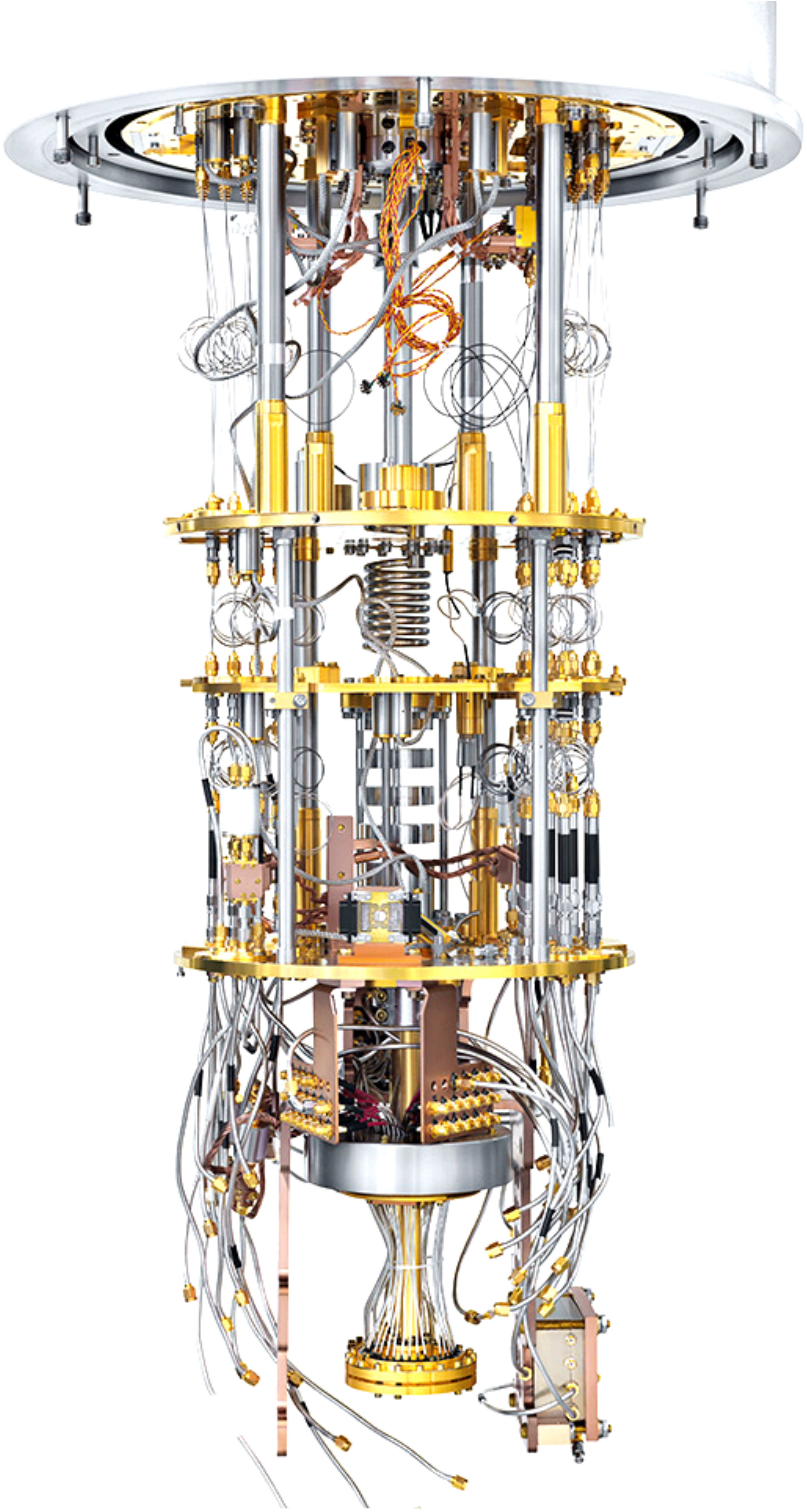
Coded a variational quantum classifier

Barren plateaus

Quantum neural networks

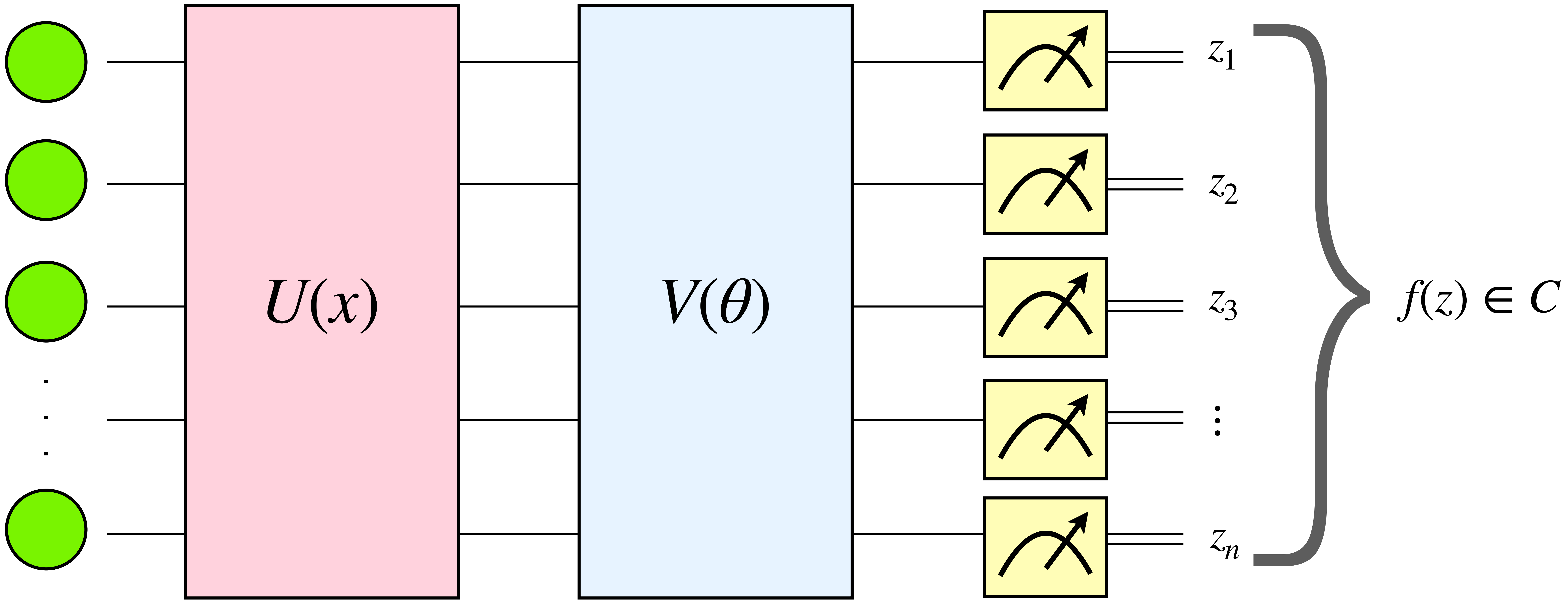
Loss landscapes

Demonstrate barren plateaus using PennyLane

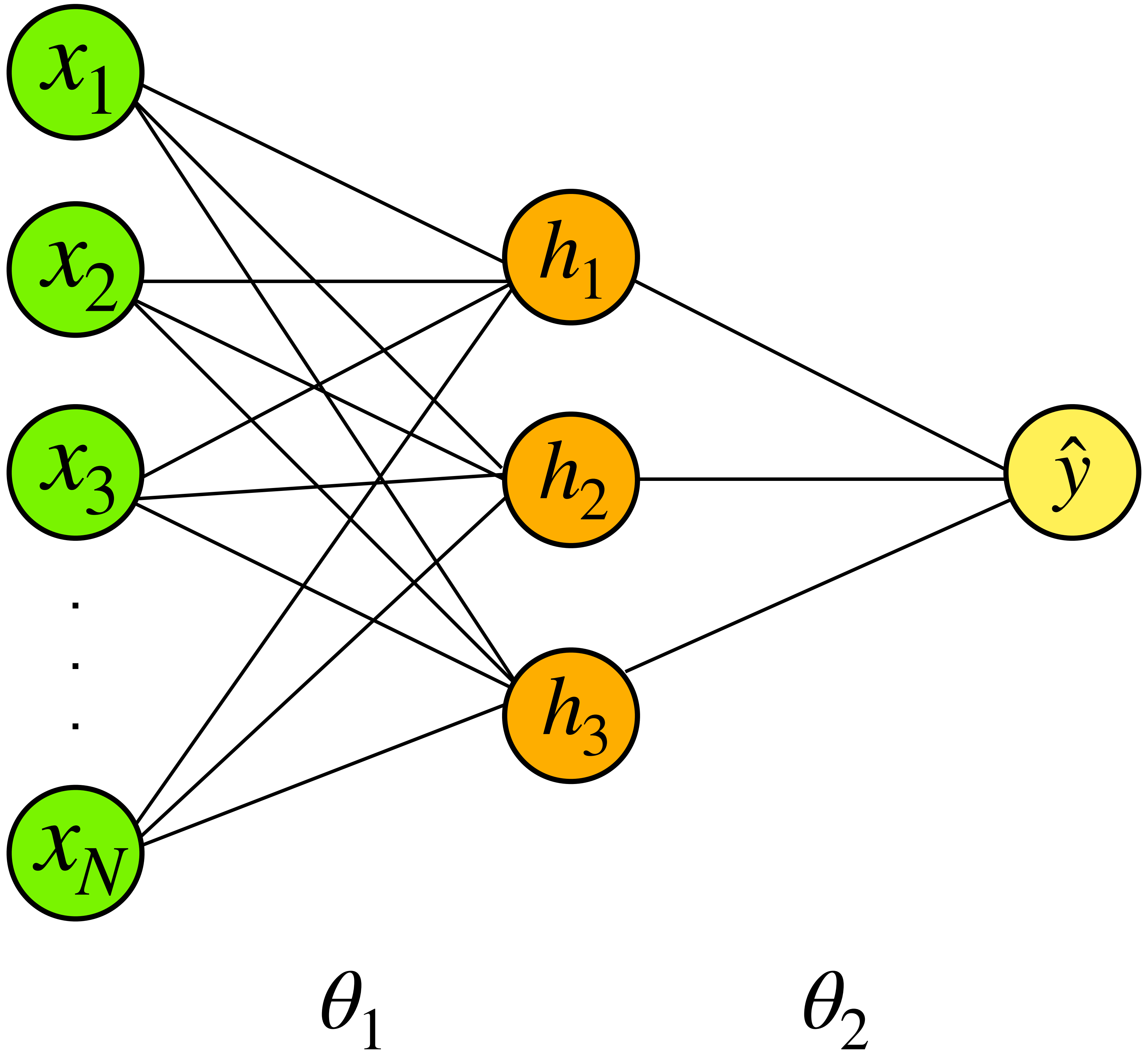


Quantum neural network

Quantum variational classifier

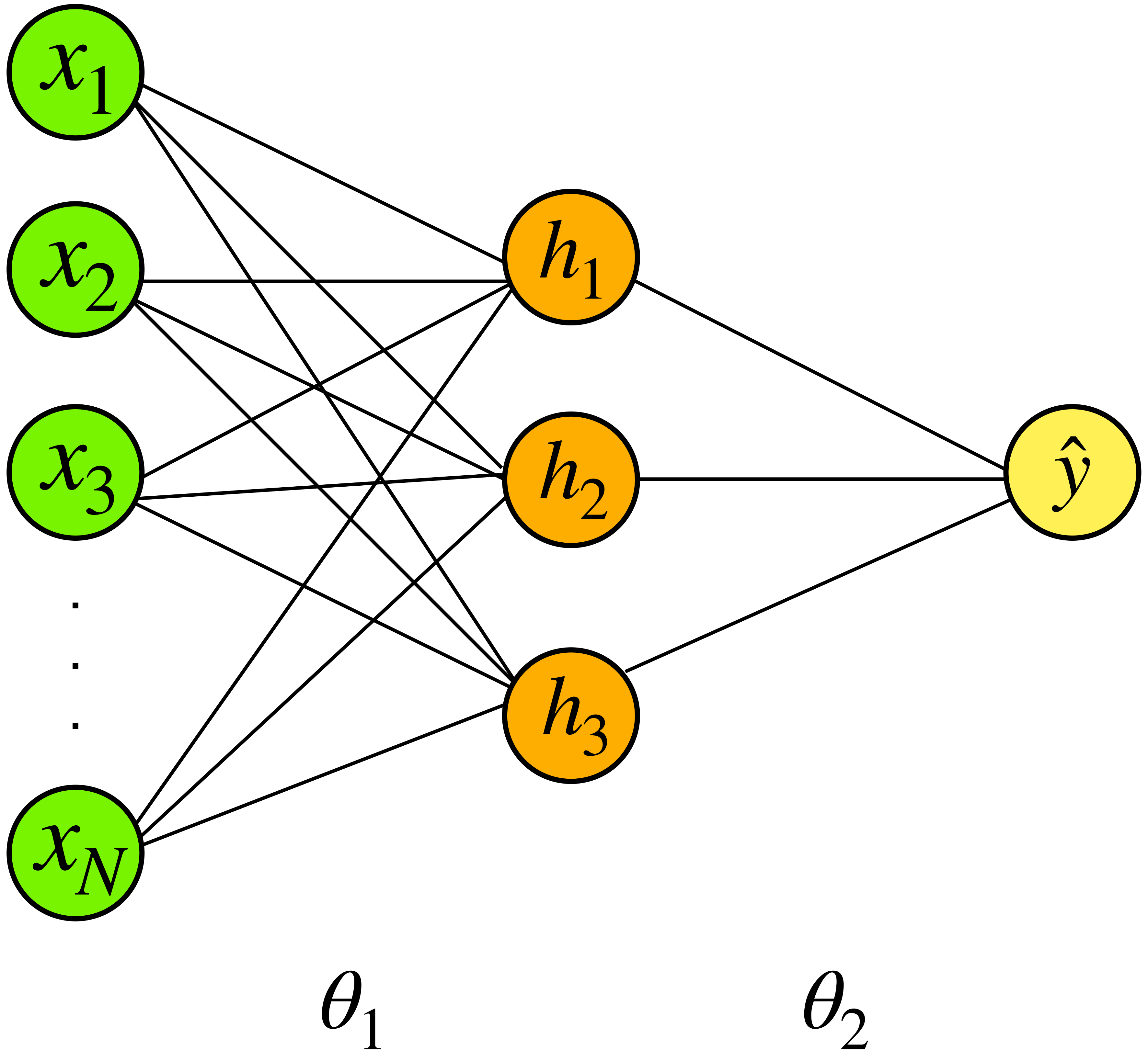


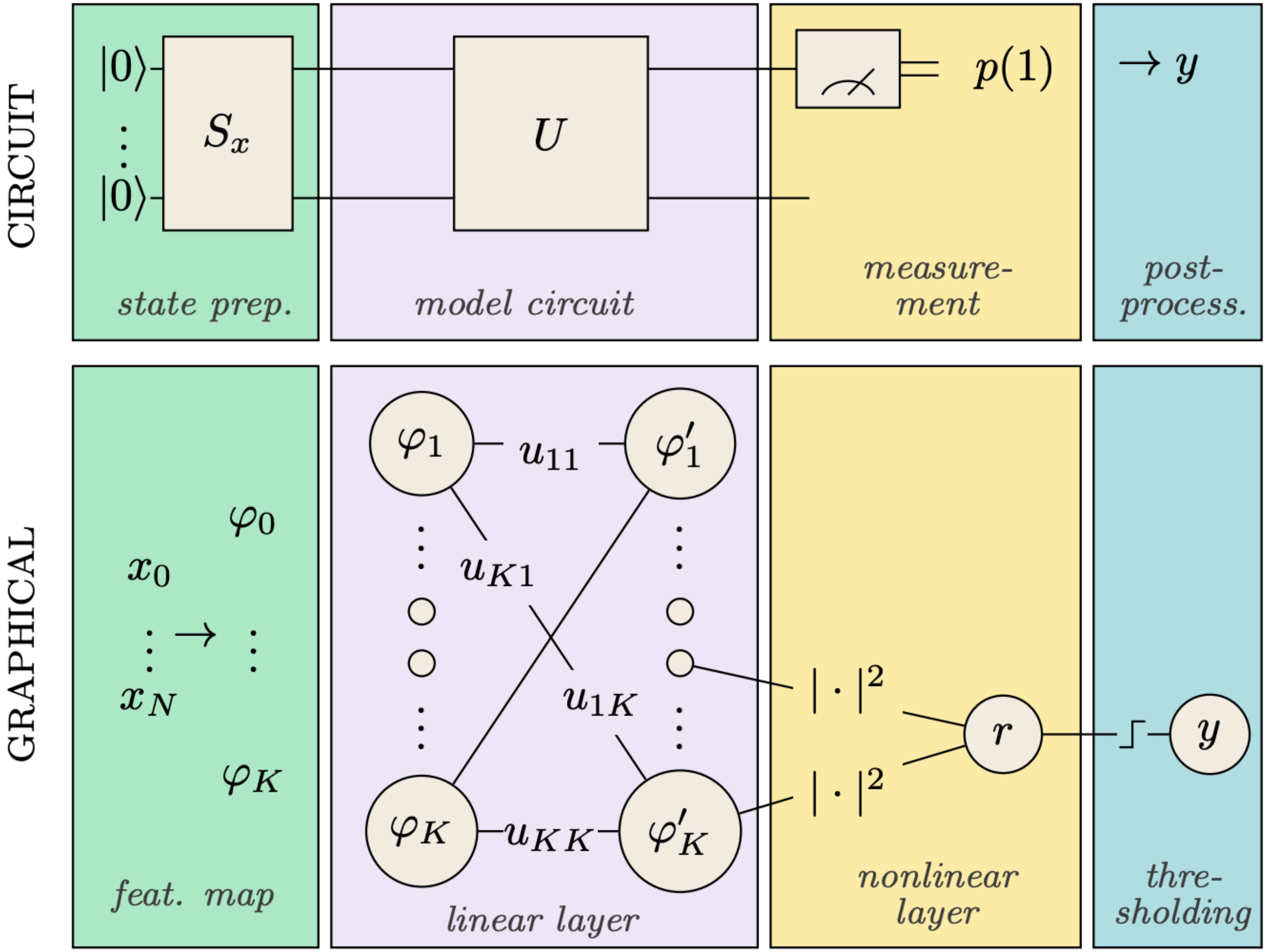
Linear neural network



$$\hat{y} = \theta_2(\theta_1 x)$$

Quantum neural networks

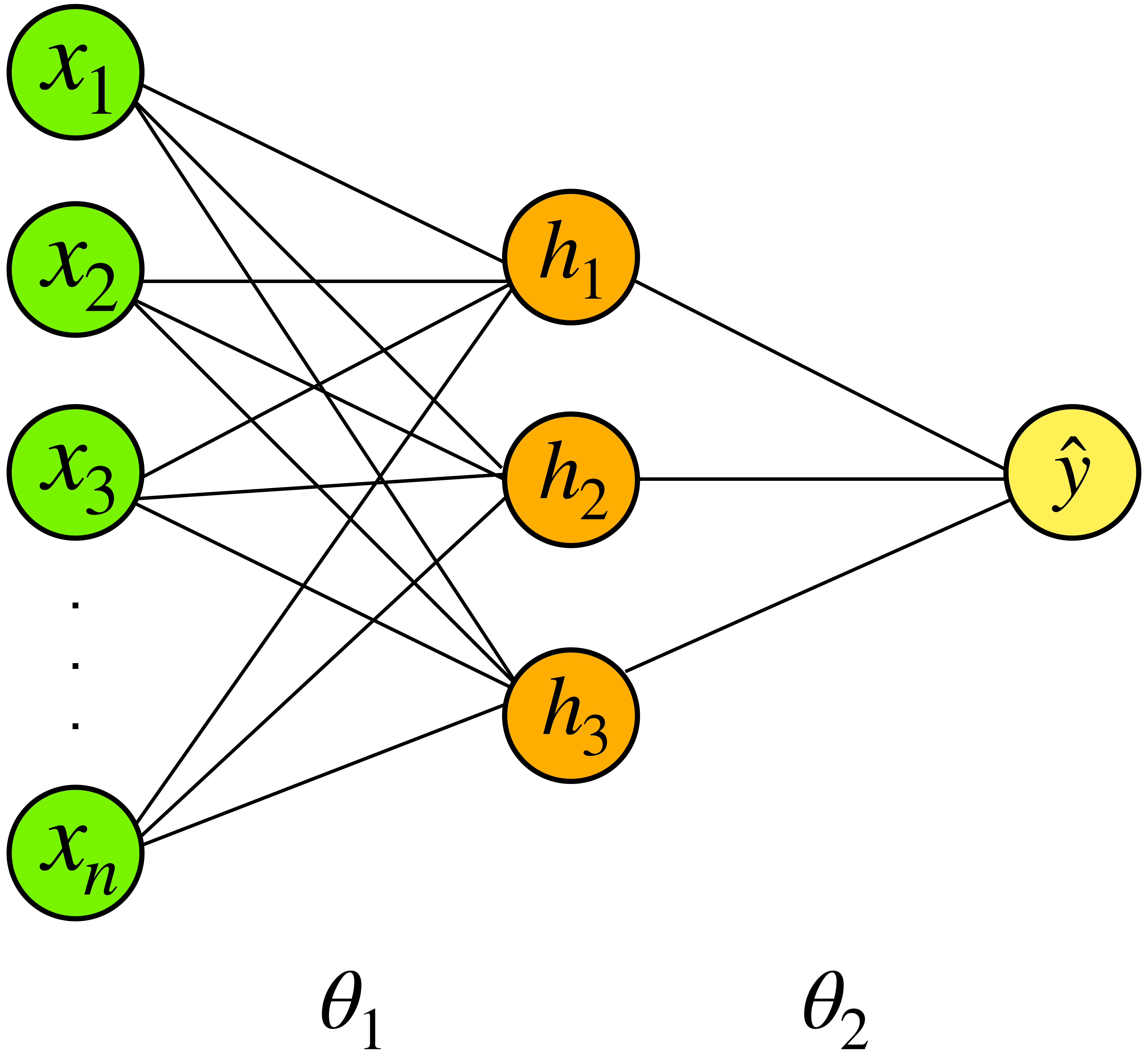




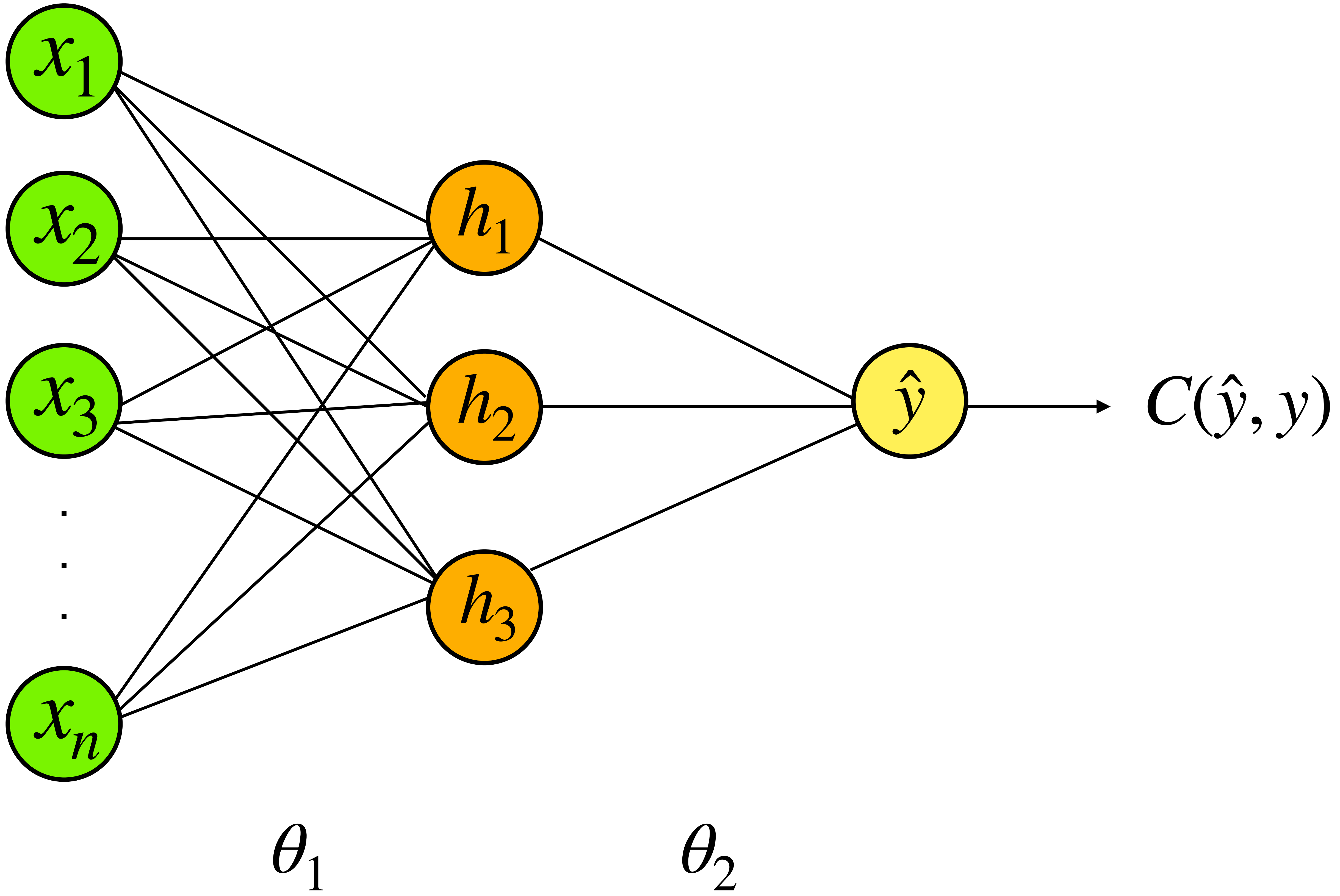
Source:Schuld, Maria, et al. "Circuit-centric quantum classifiers." Physical Review A 101.3 (2020): 032308.

Classical optimisation

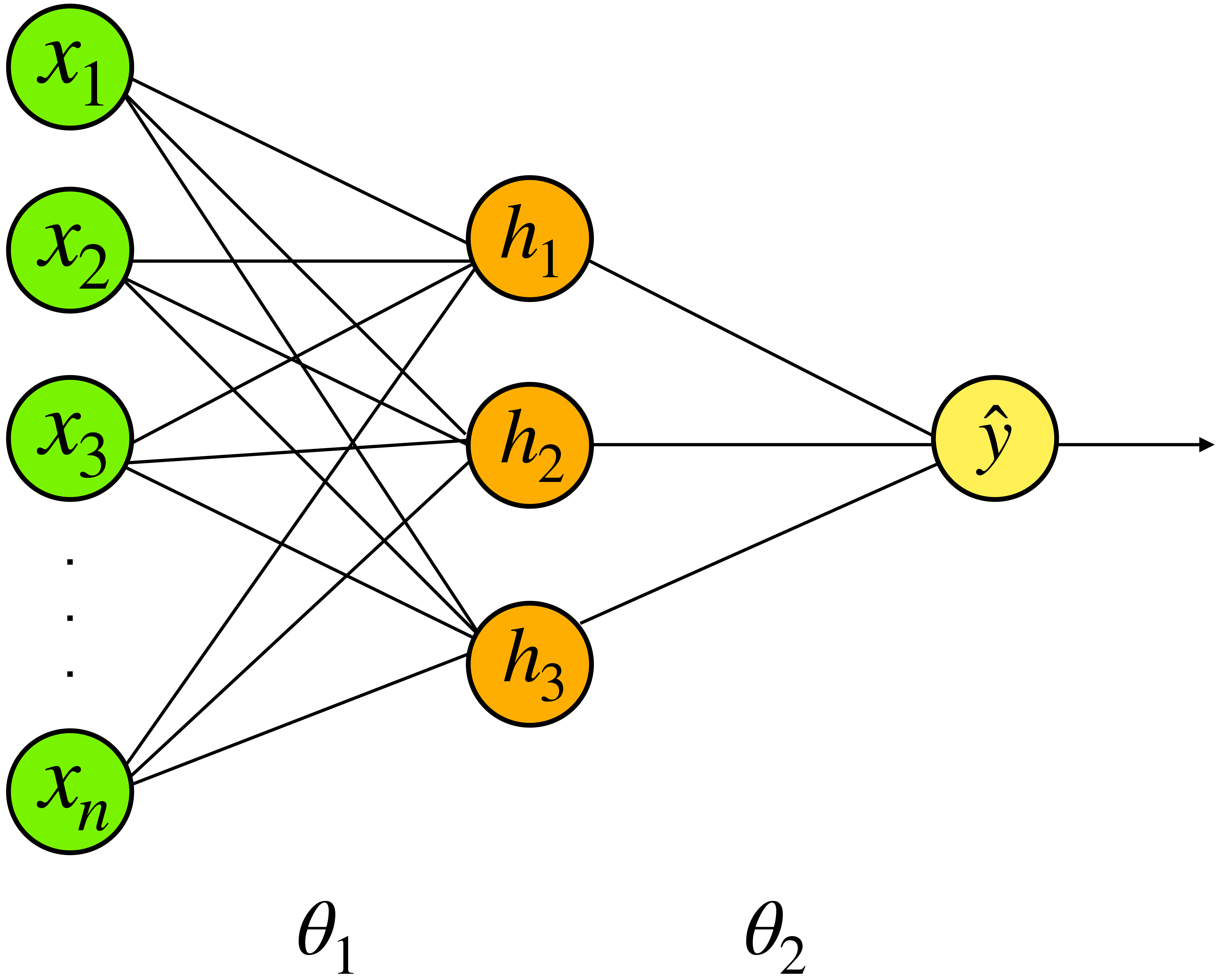
Classical neural networks



Classical neural networks

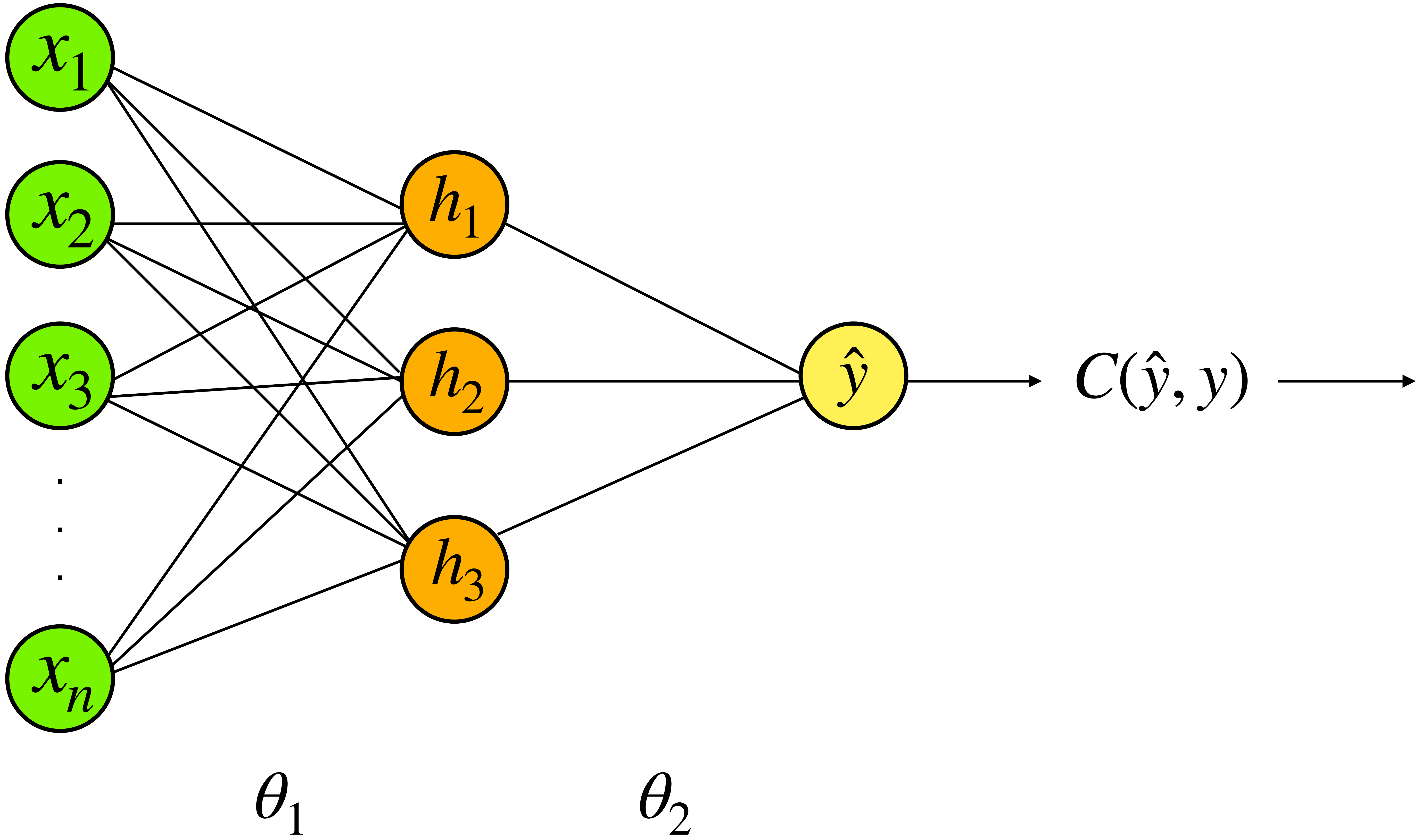


Classical neural networks

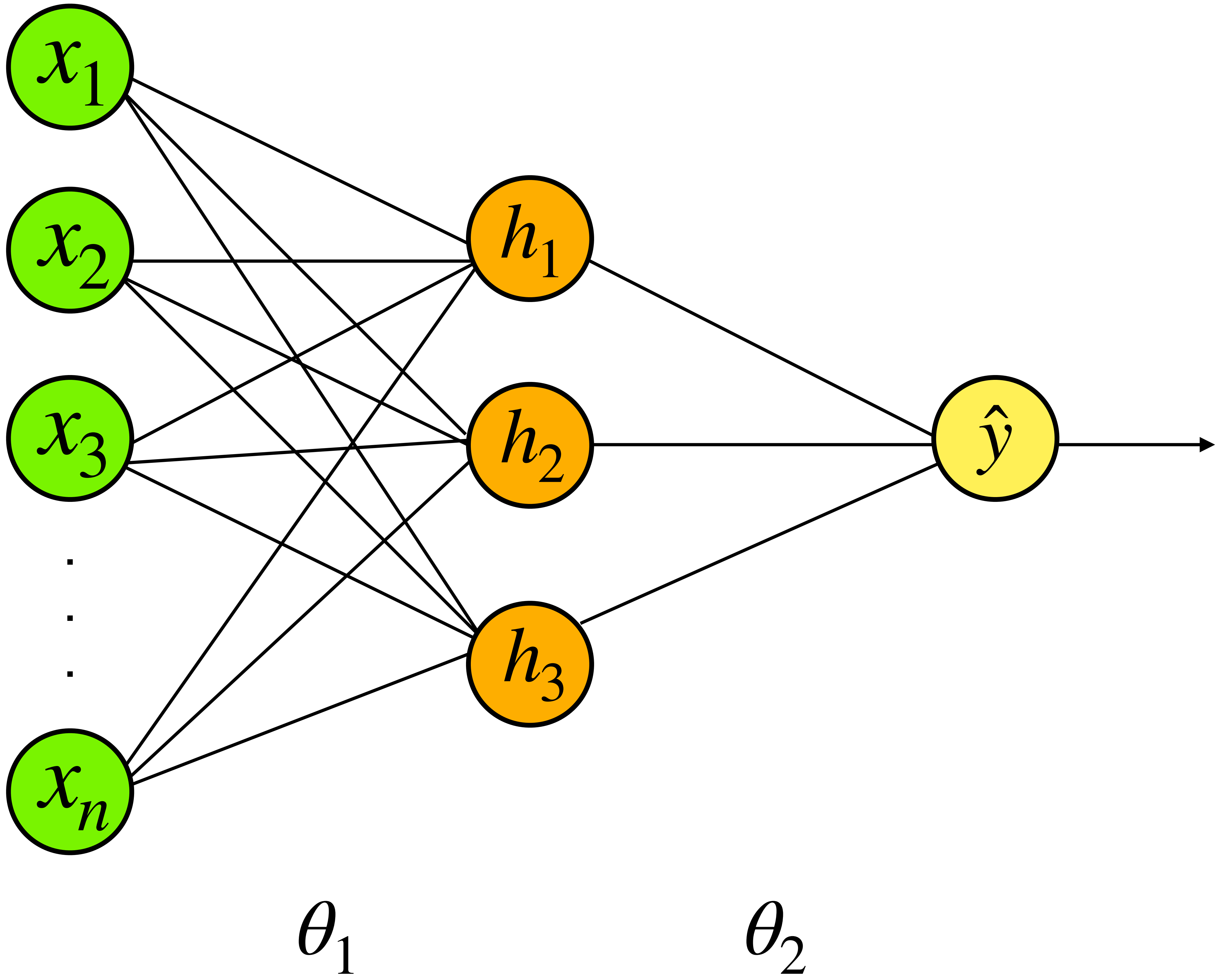


$$C(\hat{y}, y) = (\hat{y} - y)^2$$

Classical neural networks



Classical neural networks



$C(\hat{y}, y) \longrightarrow \dfrac{\partial C(\hat{y}, y)}{\partial \theta}$

Resource: https://www.jeremyjordan.me/neural-networks-training/

Gradient descent

$$\theta_{\text{new}} = \theta - \eta \frac{\partial C(\hat{y}, y)}{\partial \theta}$$

Resource: https://www.jeremyjordan.me/neural-networks-training/

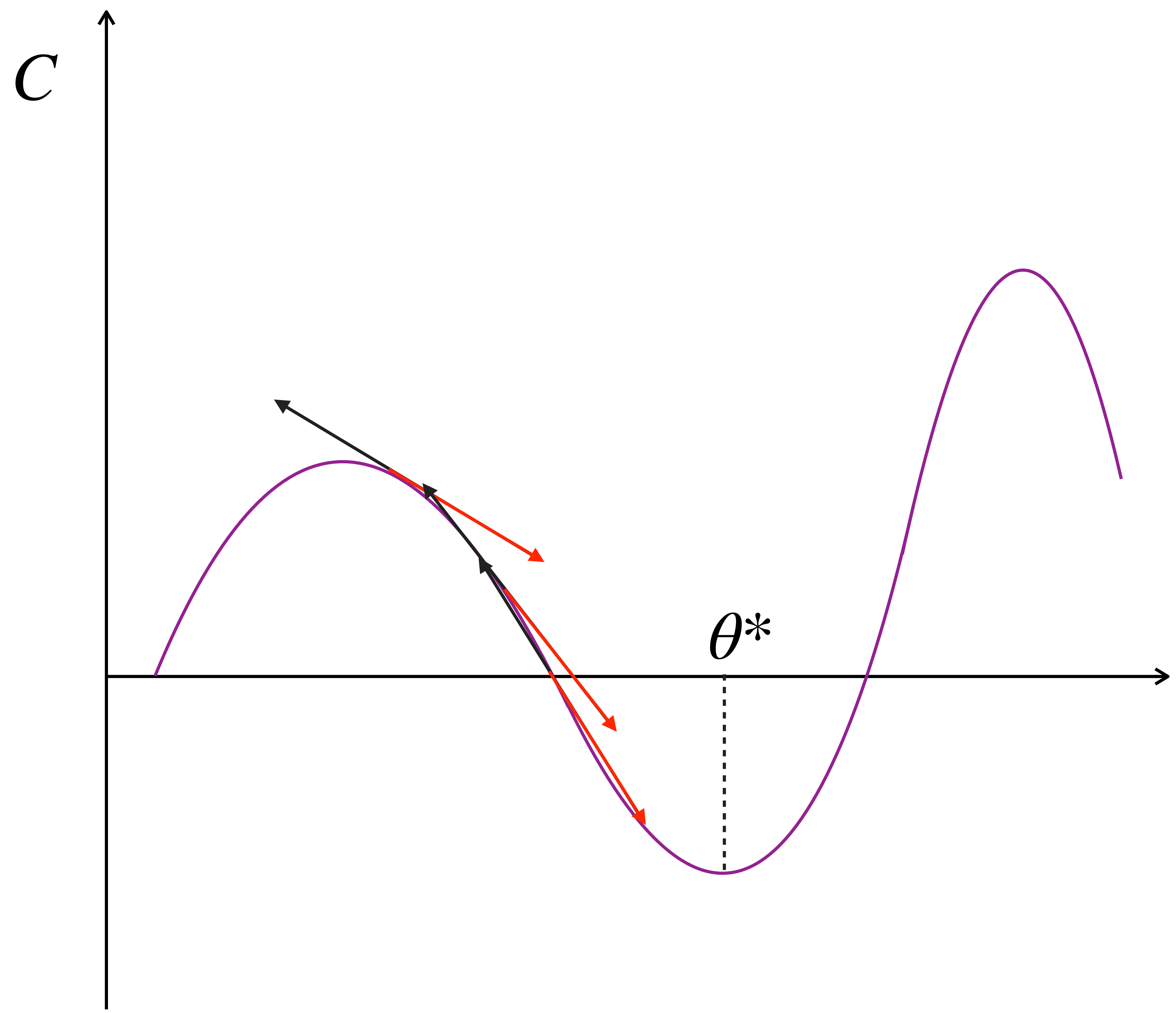
Gradient descent

$$\theta^* = \theta - \eta \frac{\partial C(\hat{y}, y)}{\partial \theta}$$

Resource: https://www.jeremyjordan.me/neural-networks-training/

Gradient descent

$$\theta^* = \theta - \eta \frac{\partial C(\hat{y}, y)}{\partial \theta}$$



Resource: https://www.jeremyjordan.me/neural-networks-training/

Vanishing and exploding gradients

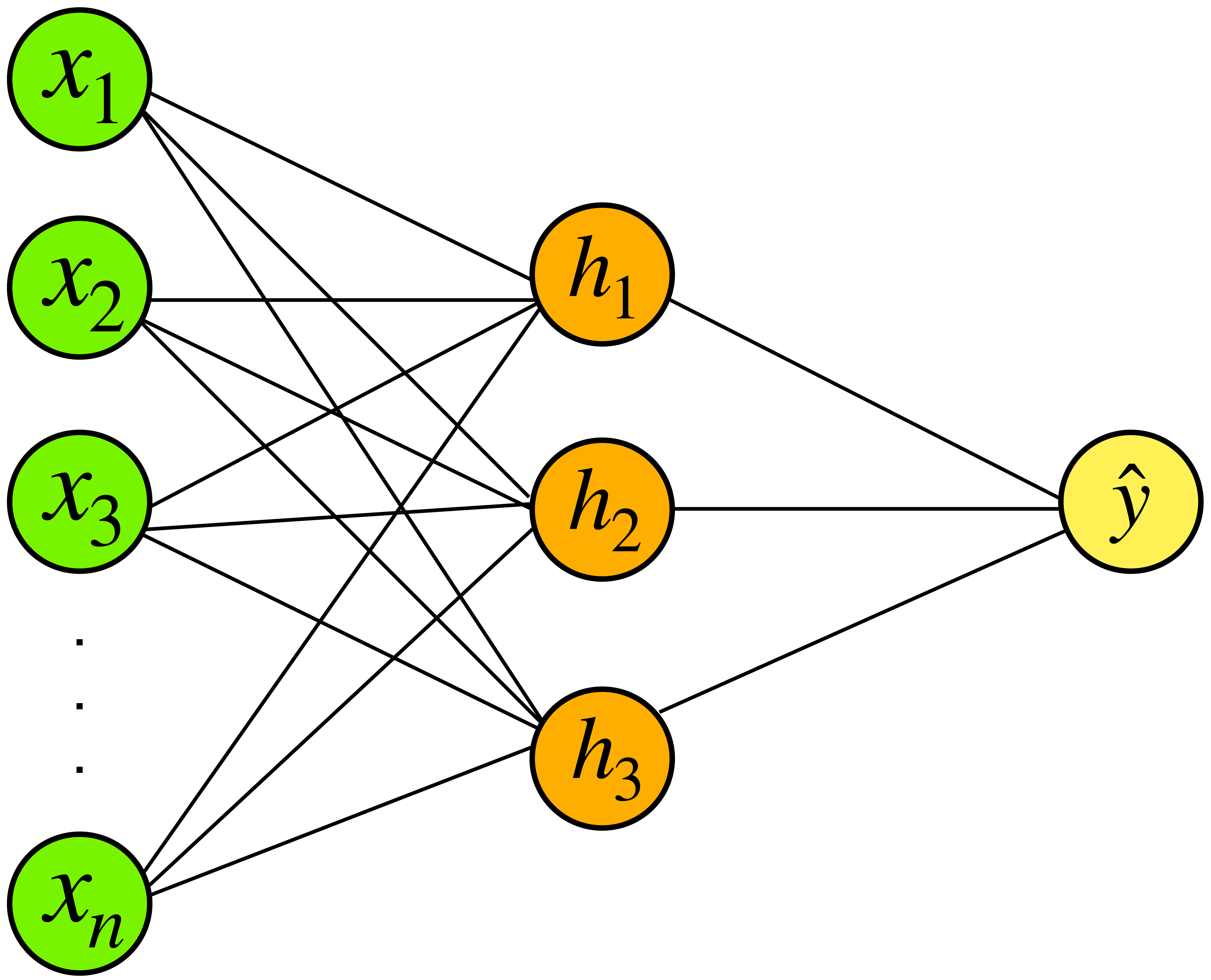
Classical neural networks

$$\theta^* = \theta - \eta \boxed{\frac{\partial C(\hat{y}, y)}{\partial \theta}}$$

Resource: https://www.jeremyjordan.me/neural-networks-training/

Classical neural networks

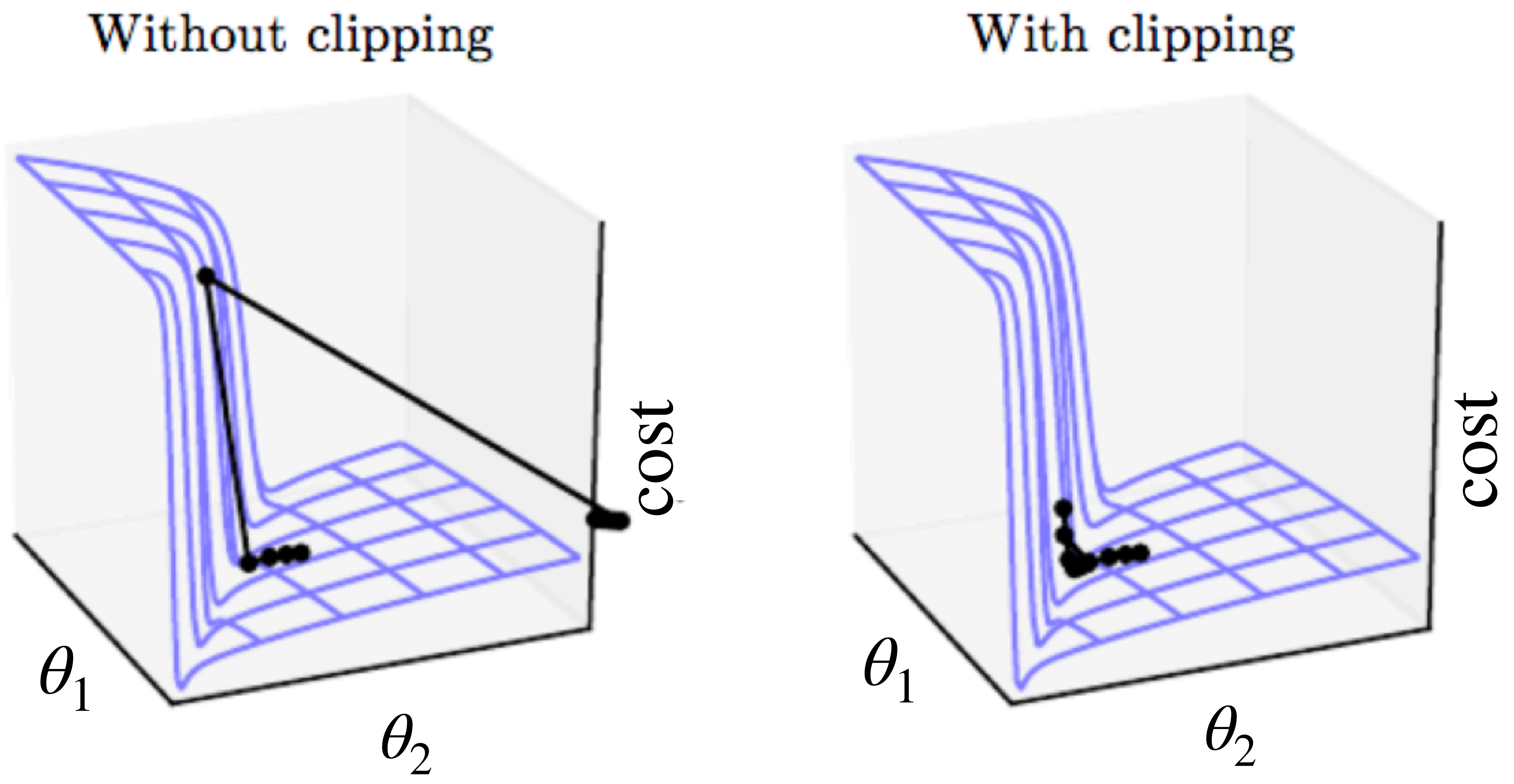
$$\theta^* = \theta - \eta \boxed{\frac{\partial C(\hat{y}, y)}{\partial \theta}}$$



Using the chain rule can cause gradients to vanish/explode

Classical neural networks

- Gradient clipping



Classical neural networks

- Gradient clipping
- Weight regularisation

L1 regularization on least squares:

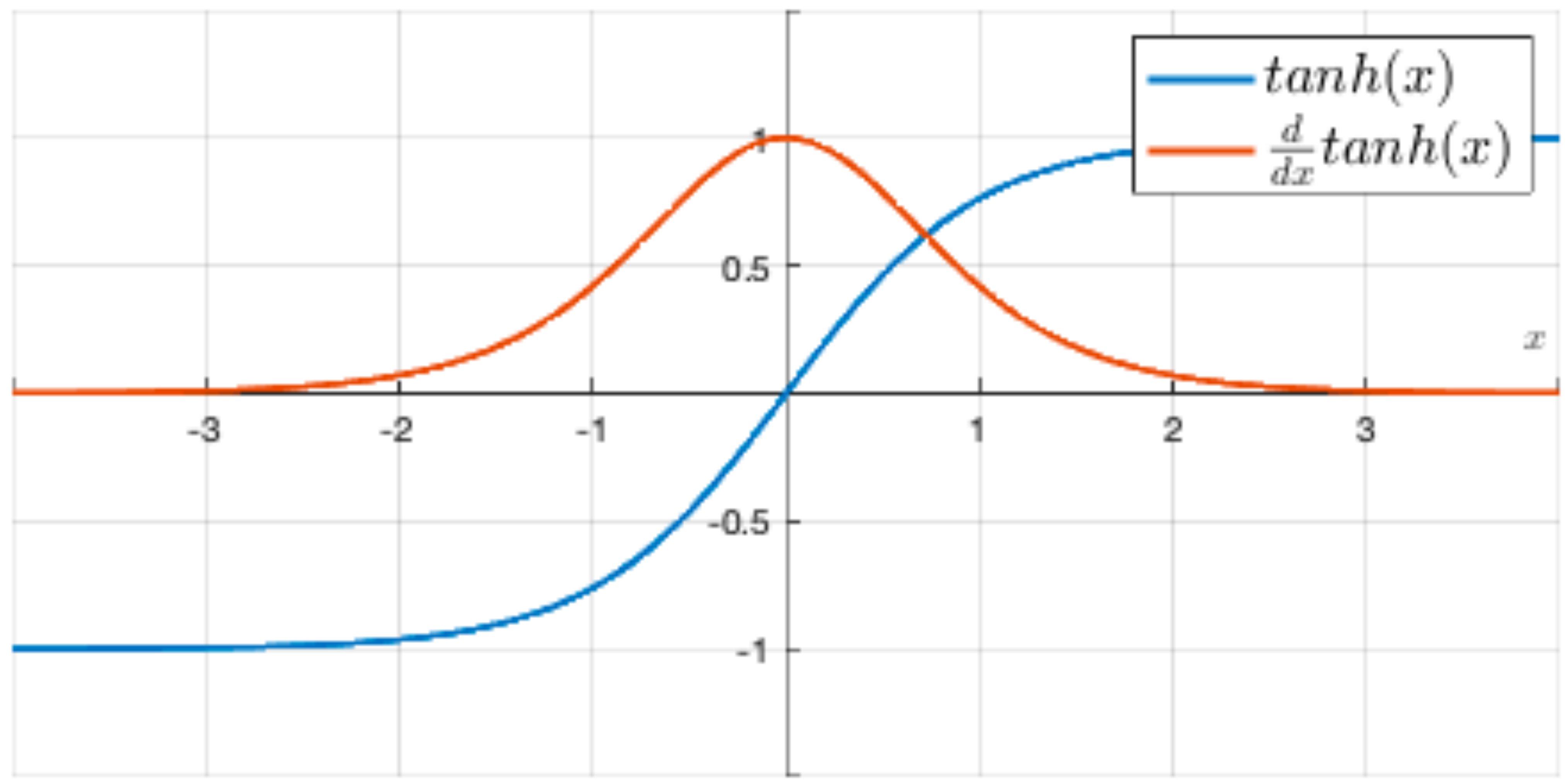
$$\mathbf{w}^* = \arg\min_{\mathbf{w}} \sum_j \left(t(\mathbf{x}_j) - \sum_i w_i h_i(\mathbf{x}_j) \right)^2 + \boxed{\lambda \sum_{i=1}^{k} |w_i|}$$

L2 regularization on least squares:

$$\mathbf{w}^* = \arg\min_{\mathbf{w}} \sum_j \left(t(\mathbf{x}_j) - \sum_i w_i h_i(\mathbf{x}_j) \right)^2 + \boxed{\lambda \sum_{i=1}^{k} w_i^2}$$

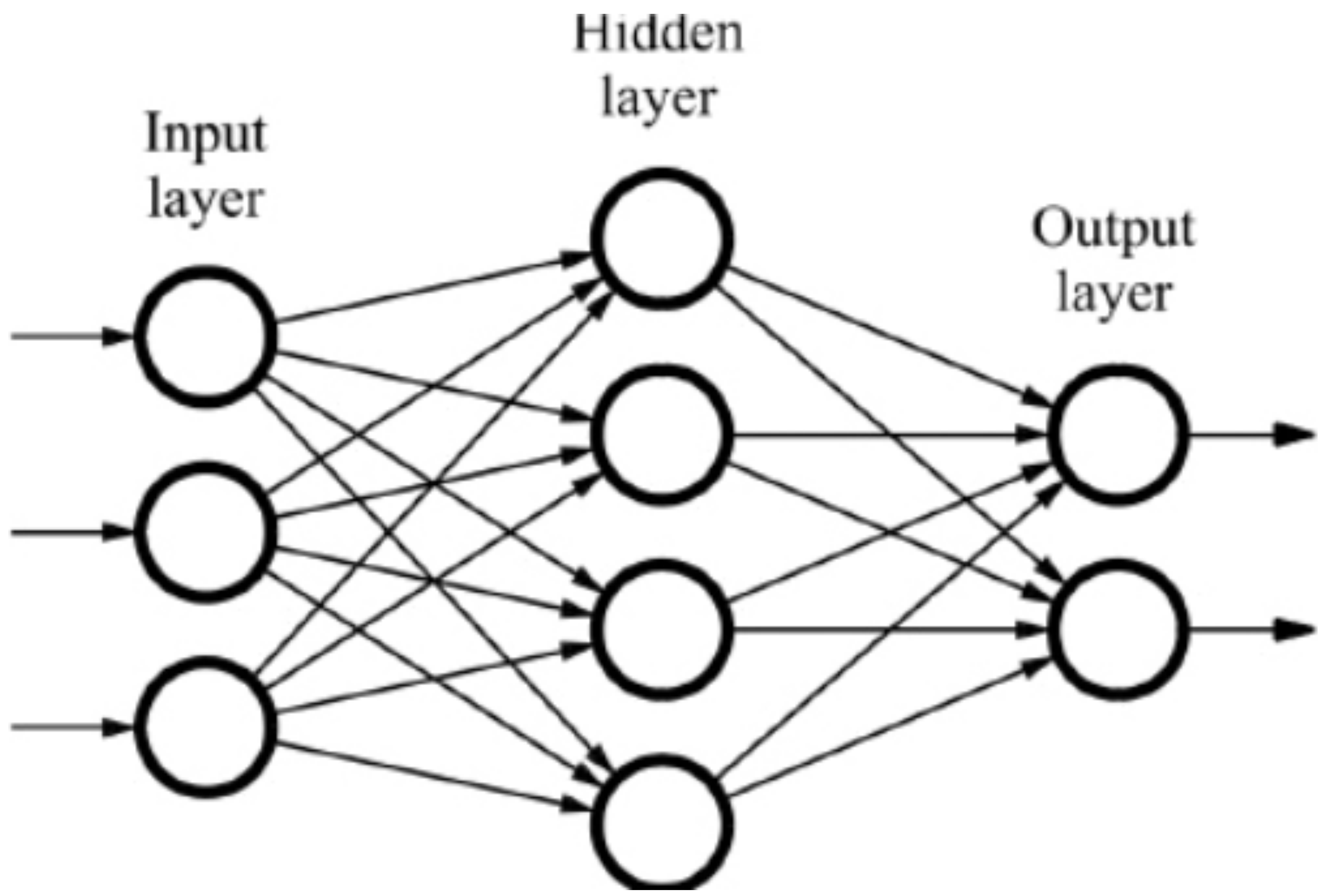
Classical neural networks

- Gradient clipping
- Weight regularisation
- Avoiding certain activation functions



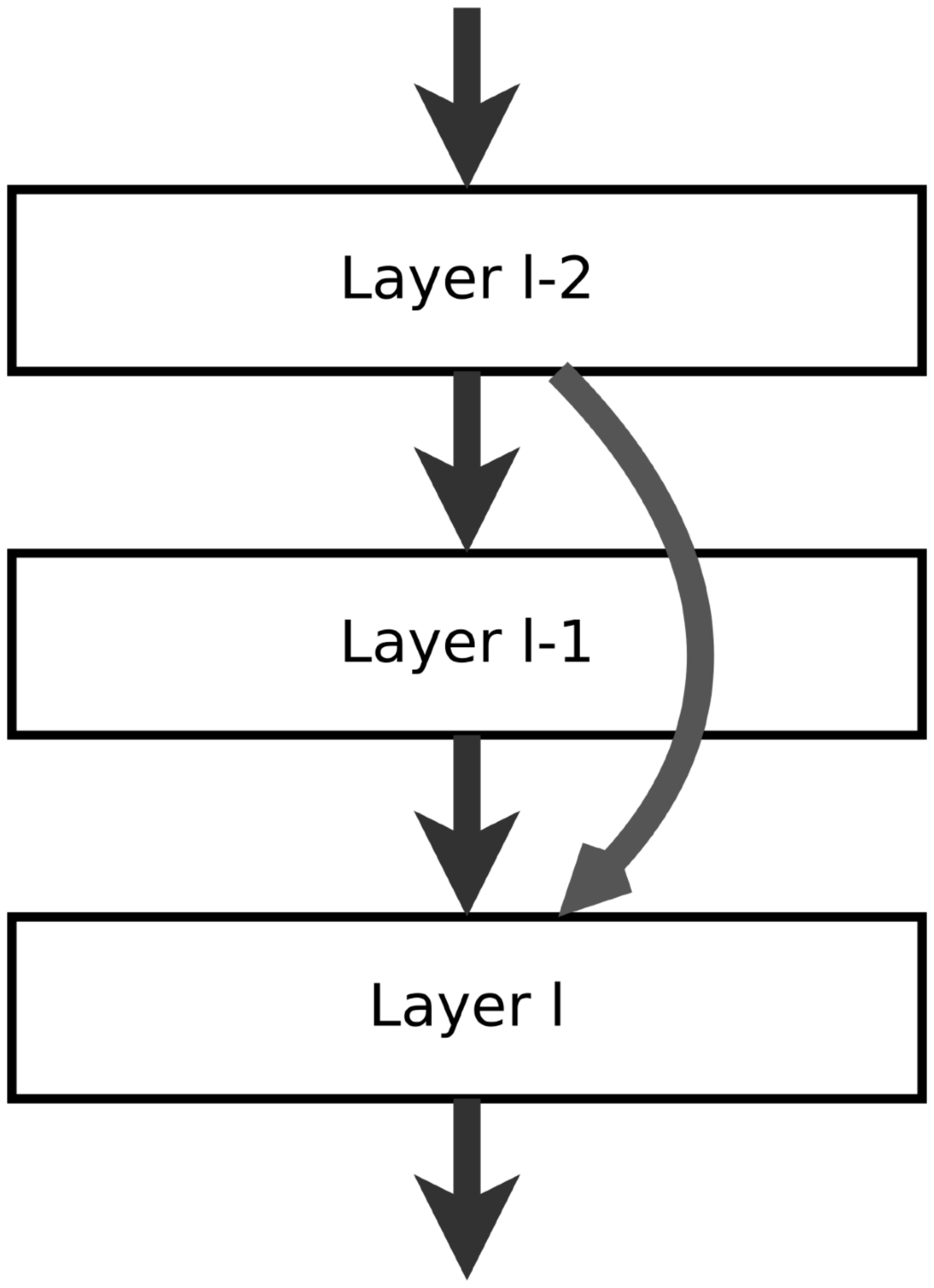
Classical neural networks

- Gradient clipping
- Weight regularisation
- Avoiding certain activation functions
- Redesigning the network

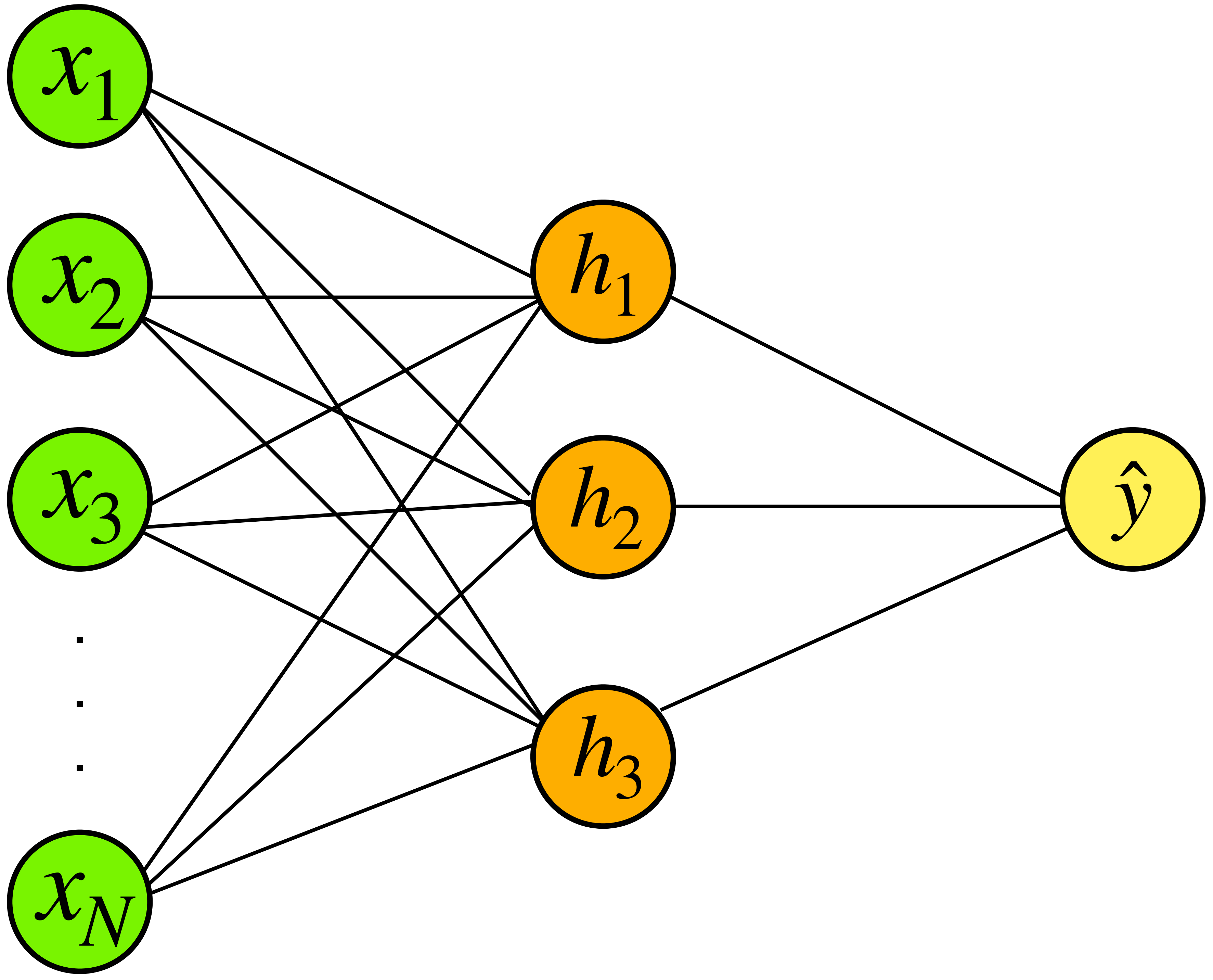


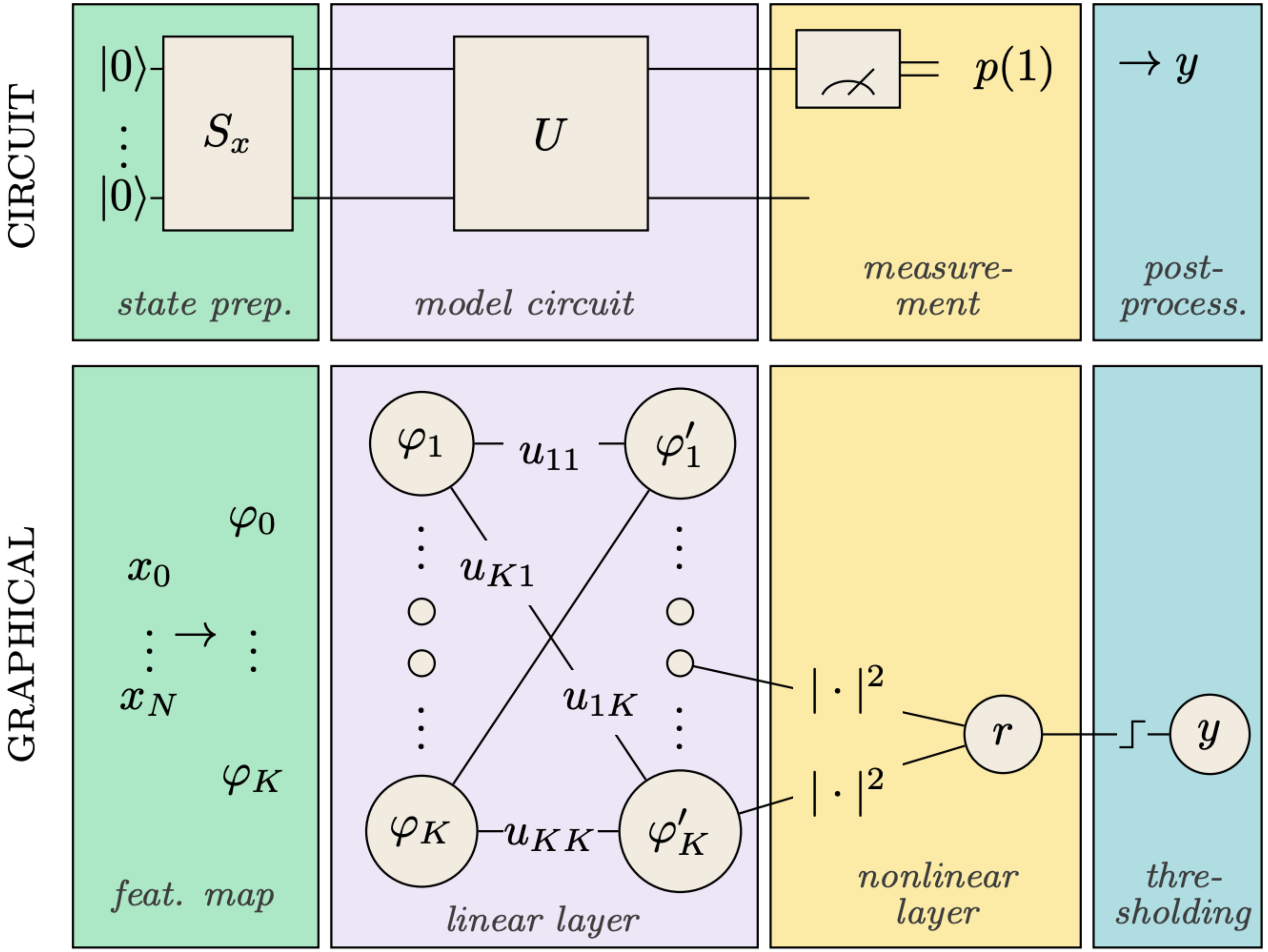
Classical neural networks

- Gradient clipping
- Weight regularisation
- Avoiding certain activation functions
- Redesigning the network
- Residual networks



Quantum neural networks

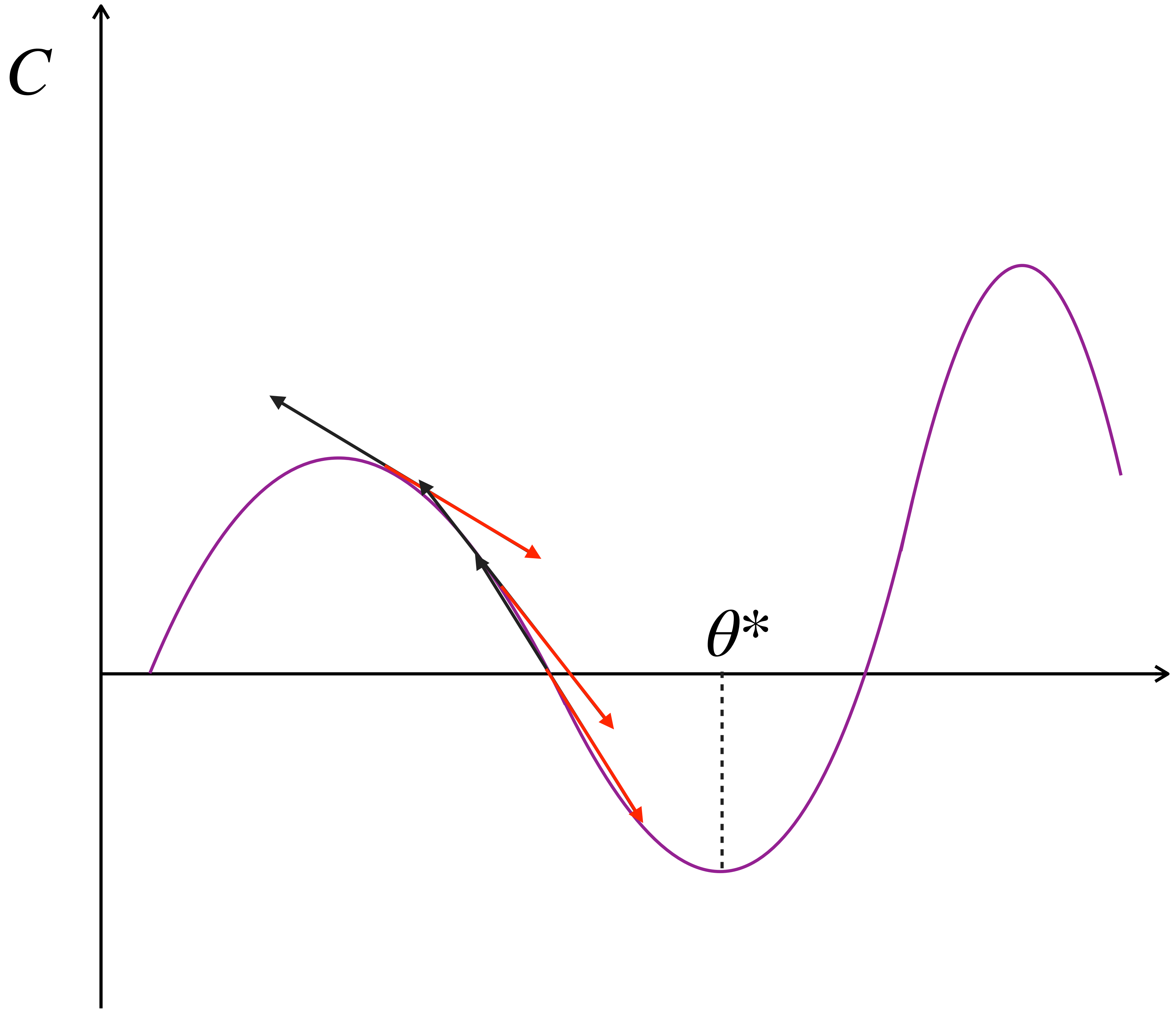




Source:Schuld, Maria, et al. "Circuit-centric quantum classifiers." Physical Review A 101.3 (2020): 032308.

Gradient descent

$$\theta^* = \theta - \eta \frac{\partial C(\hat{y}, y)}{\partial \theta}$$



Resource: https://www.jeremyjordan.me/neural-networks-training/

Barren plateaus

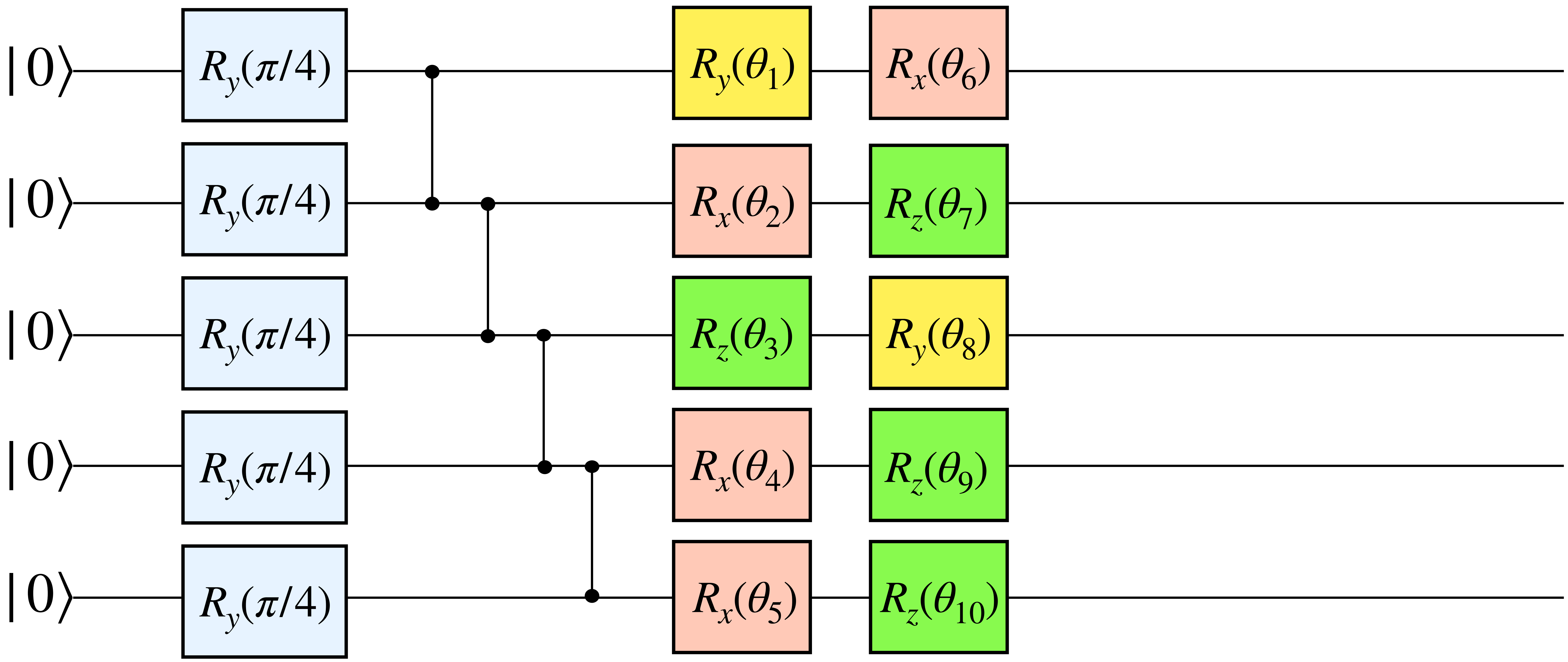
Random initialisation

If θ is initialised randomly, gates are drawn randomly

Source: McClean, Jarrod R., et al. "Barren plateaus in quantum neural network training landscapes." Nature communications 9.1 (2018): 1-6.

Random initialisation

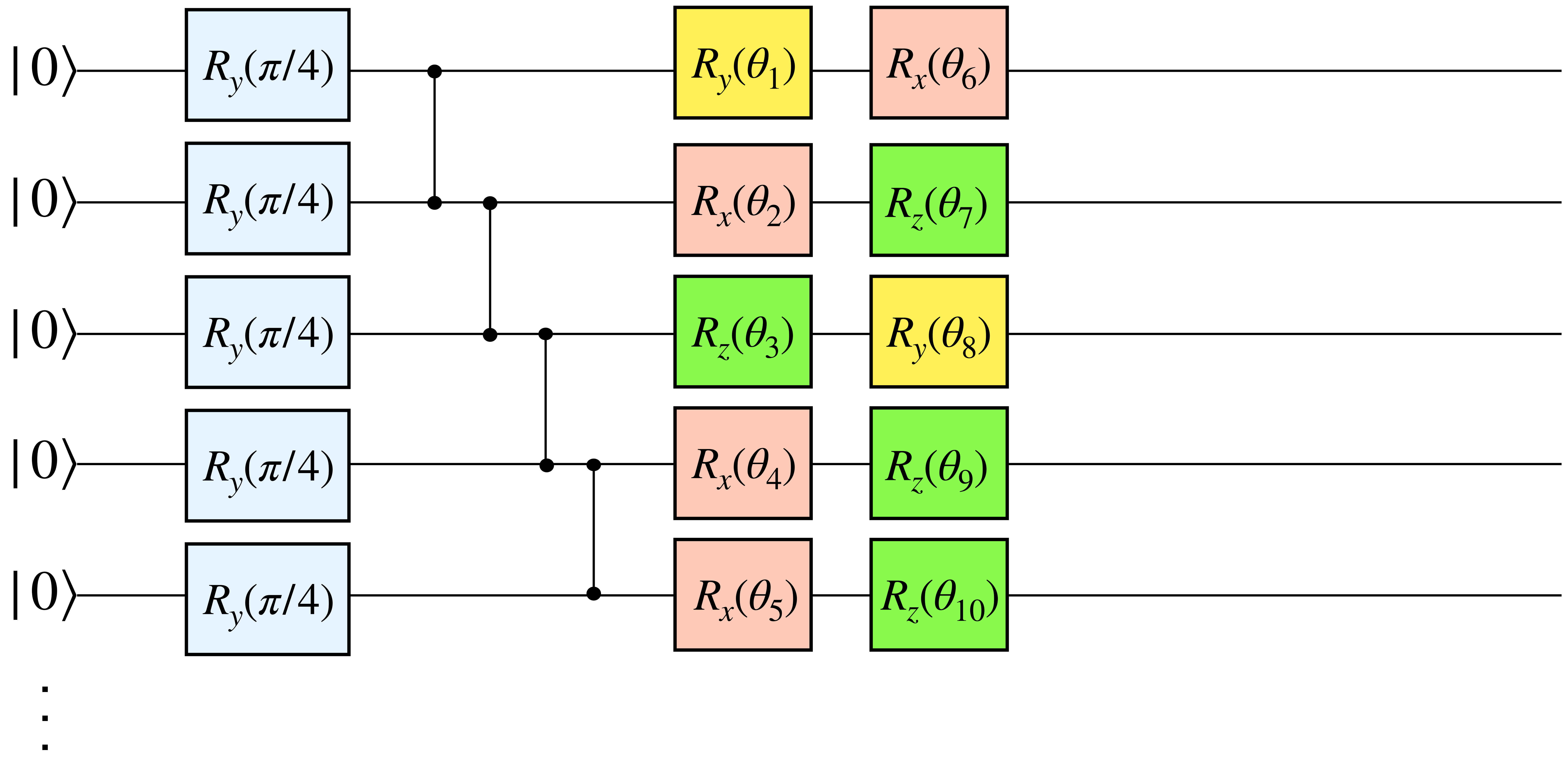
If θ is initialised randomly, gates are drawn randomly



Source: McClean, Jarrod R., et al. "Barren plateaus in quantum neural network training landscapes." Nature communications 9.1 (2018): 1-6.

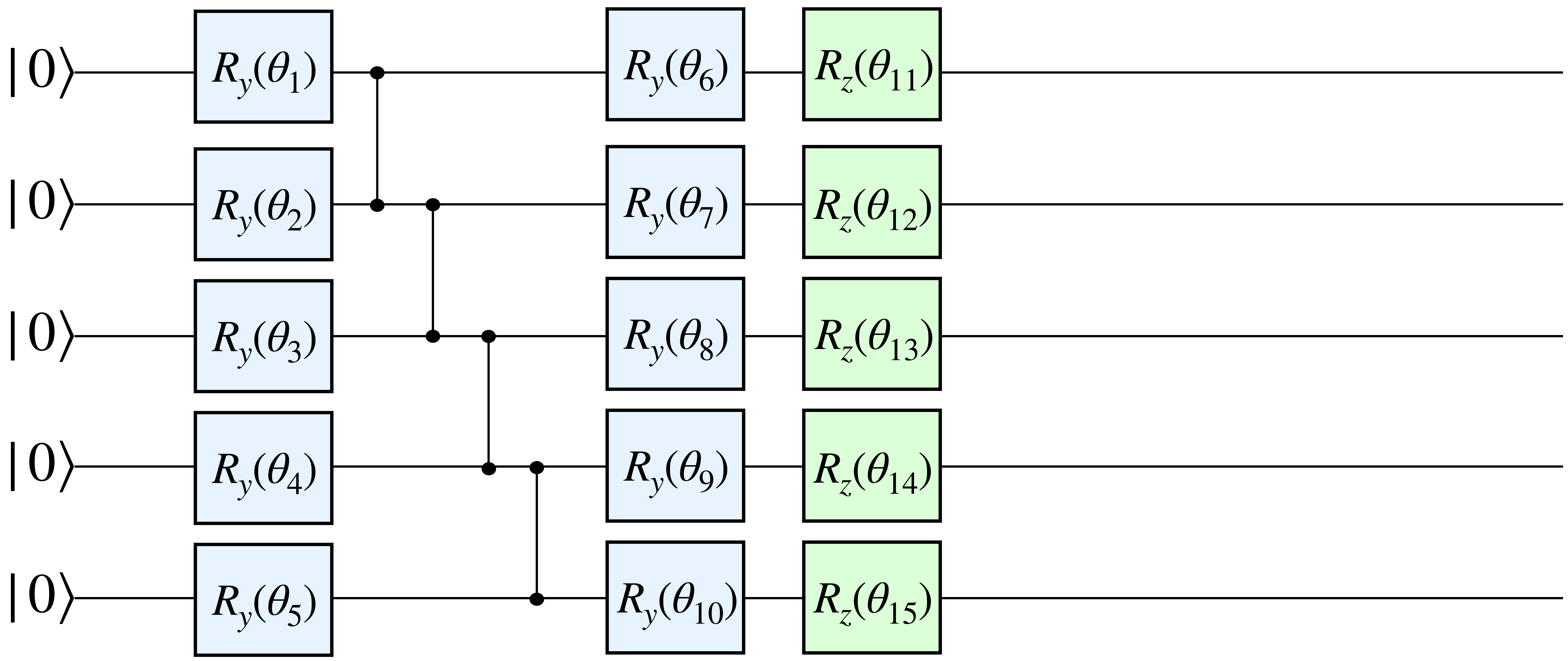
Random initialisation

If θ is initialised randomly, gates are drawn randomly

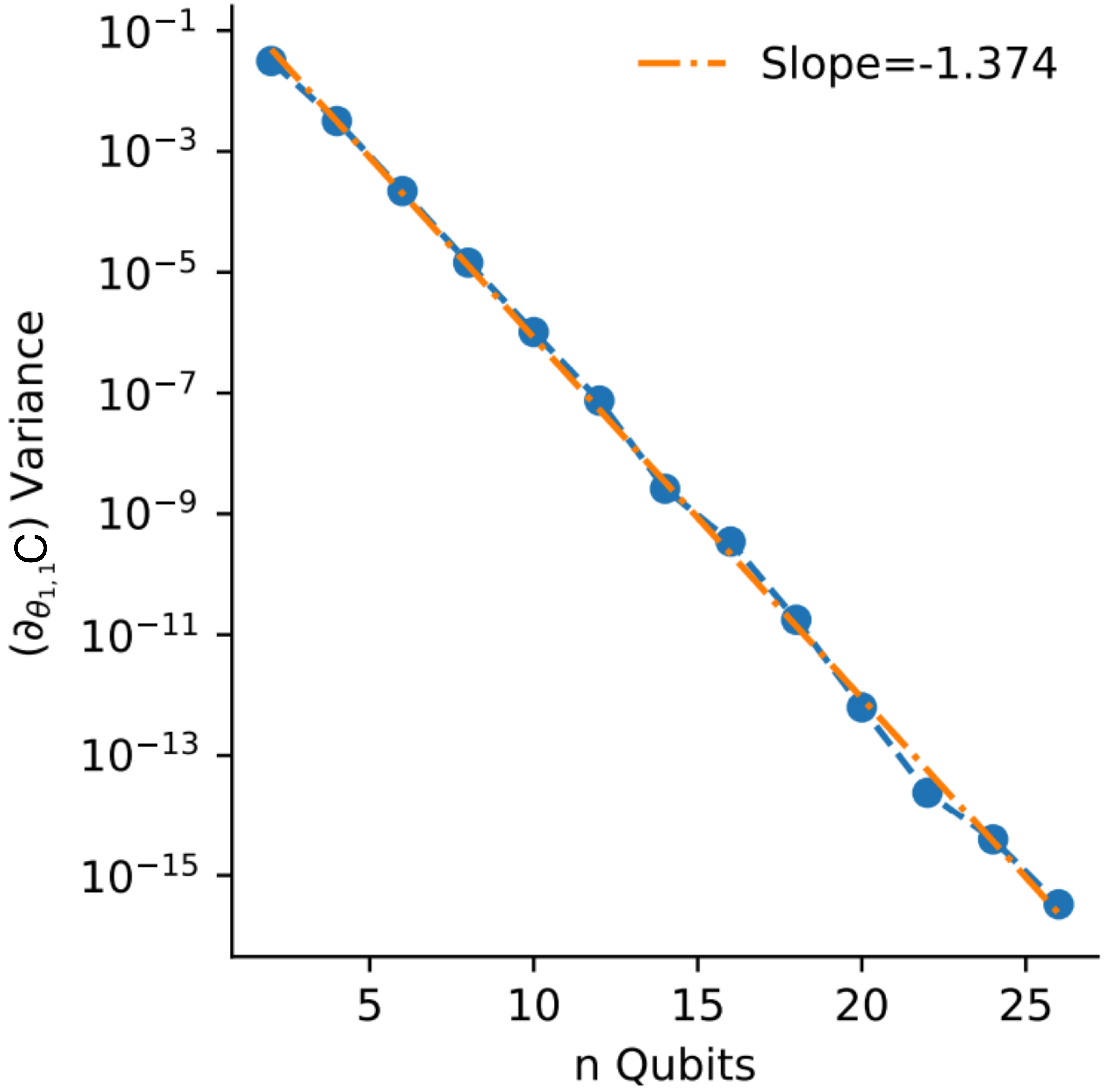


Source: McClean, Jarrod R., et al. "Barren plateaus in quantum neural network training landscapes." Nature communications 9.1 (2018): 1-6.

Random initialisation



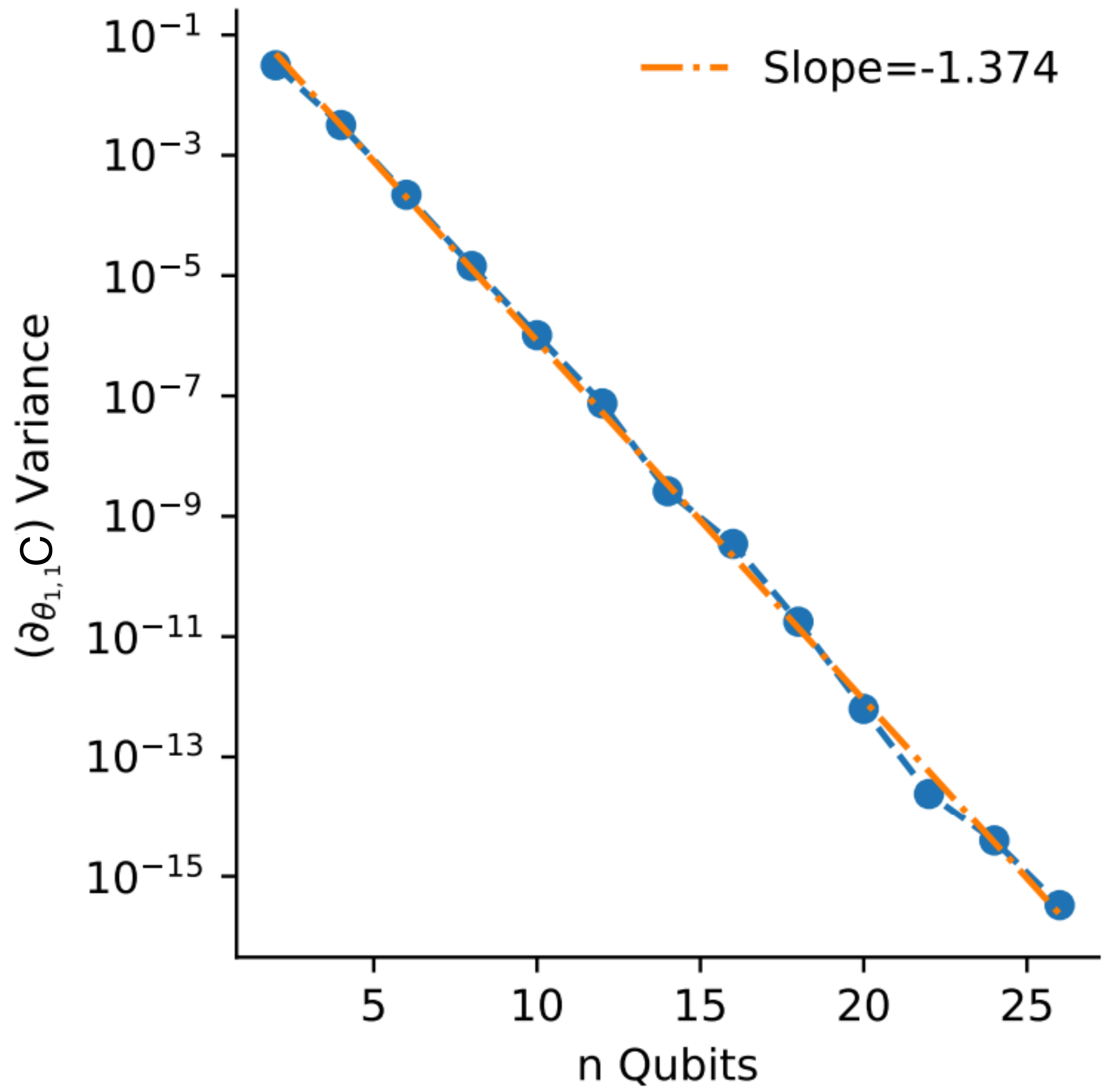
Random initialisation



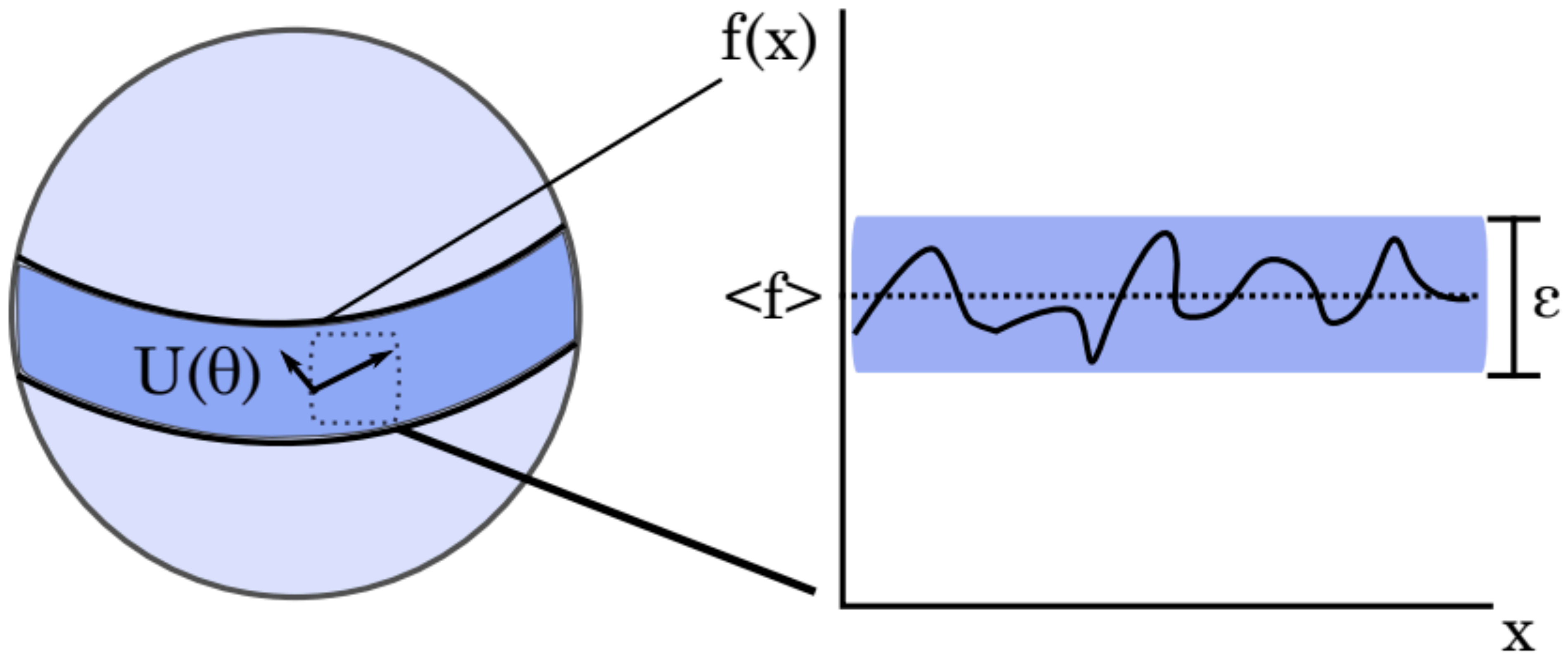
- If θ is initialised randomly, this can cause problems in variational circuits
- The variance of the gradient of the loss function vanishes

Source: McClean, Jarrod R., et al. "Barren plateaus in quantum neural network training landscapes." Nature communications 9.1 (2018): 1-6.

Random initialisation



- If θ is initialised randomly, this can cause problems in variational circuits
- The variance of the gradient of the loss function vanishes
- Gradients become concentrated around zero
- Require greater precision to estimate the gradients which removes quantum advantages



Source: McClean, Jarrod R., et al. "Barren plateaus in quantum neural network training landscapes." Nature communications 9.1 (2018): 1-6.

Barren plateaus

$$\theta_{\text{new}} = \theta - \eta \frac{\partial C(\hat{y}, y)}{\partial \theta}$$

- If θ is initialised randomly, this can cause problems in variational circuits
- The variance of the gradient of the loss function vanishes
- Gradients become concentrated around zero
- Require greater precision to estimate the gradients which removes quantum advantages

Source: McClean, Jarrod R., et al. "Barren plateaus in quantum neural network training landscapes." Nature communications 9.1 (2018): 1-6.

Methods to try avoid barren plateaus

Methods to try avoid barren plateaus

- Introduce structure into the circuit (encoding with data) or initialising layers to the identity
  Grant, Edward, et al. "An initialization strategy for addressing barren plateaus in parametrized quantum circuits." Quantum 3 (2019): 214.

Methods to try avoid barren plateaus

- Introduce structure into the circuit (encoding with data) or initialising layers to the identity

  Grant, Edward, et al. "An initialization strategy for addressing barren plateaus in parametrized quantum circuits." Quantum 3 (2019): 214.

- Use a local cost function

  Cerezo, M., et al. "Cost-function-dependent barren plateaus in shallow quantum neural networks." *arXiv preprint arXiv:2001.00550* (2020).

  https://www.eventbrite.ca/e/quantum-research-seminars-toronto-tickets-122394389915

Methods to try avoid barren plateaus

- Introduce structure into the circuit (encoding with data) or initialising layers to the identity
  Grant, Edward, et al. "An initialization strategy for addressing barren plateaus in parametrized quantum circuits." Quantum 3 (2019): 214.

- Use a local cost function
  Cerezo, M., et al. "Cost-function-dependent barren plateaus in shallow quantum neural networks." *arXiv preprint arXiv:2001.00550* (2020).

- Higher order optimisation methods
  Huembeli, Patrick, and Alexandre Dauphin. "Characterizing the loss landscape of variational quantum circuits." arXiv preprint arXiv:2008.02785 (2020).

Methods to try avoid barren plateaus

The Hessian matrix

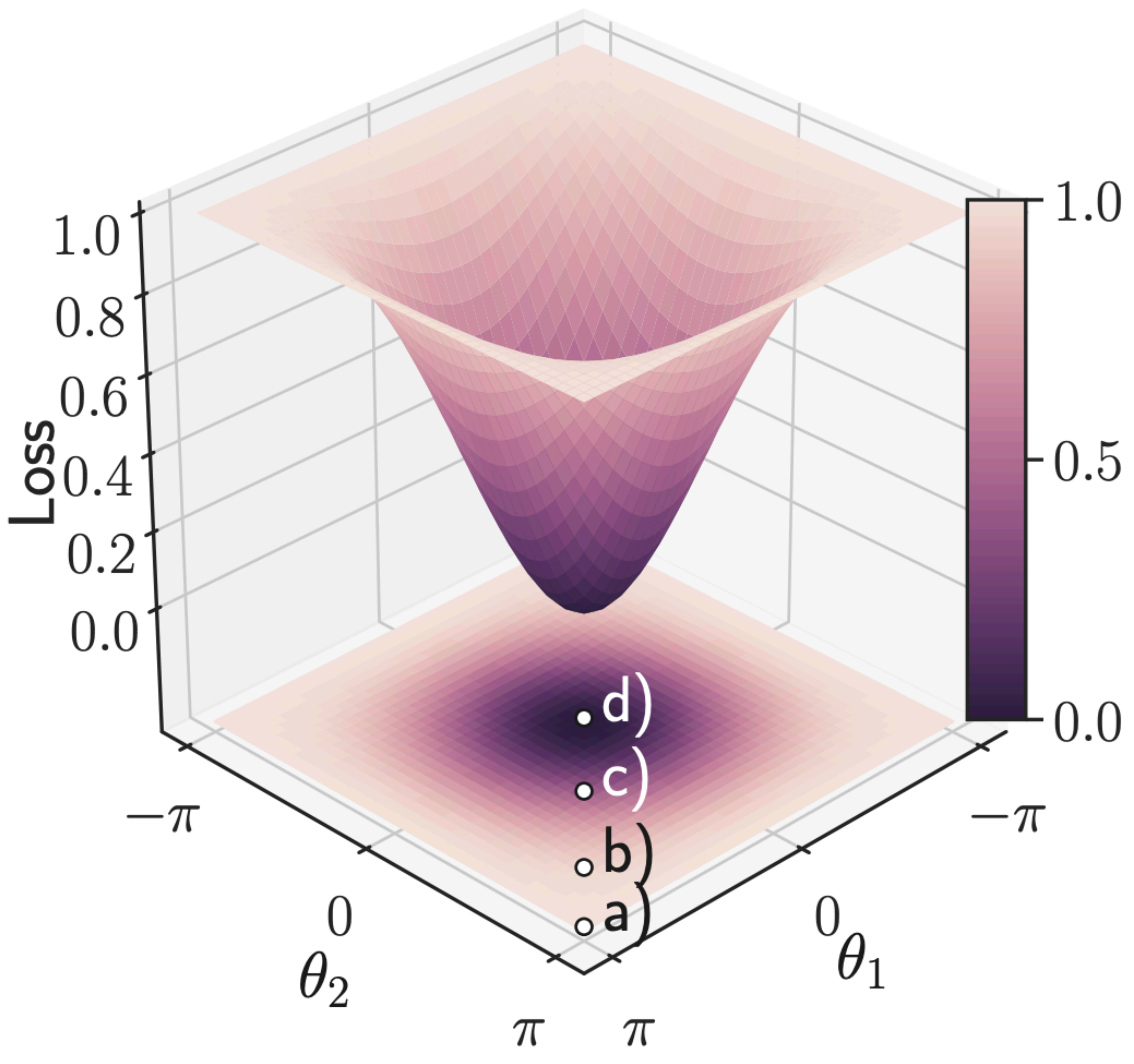
$$\frac{\partial C}{\partial \theta} = \begin{bmatrix} \frac{\partial C}{\partial \theta_1} \\ \frac{\partial C}{\partial \theta_2} \\ \vdots \end{bmatrix}$$

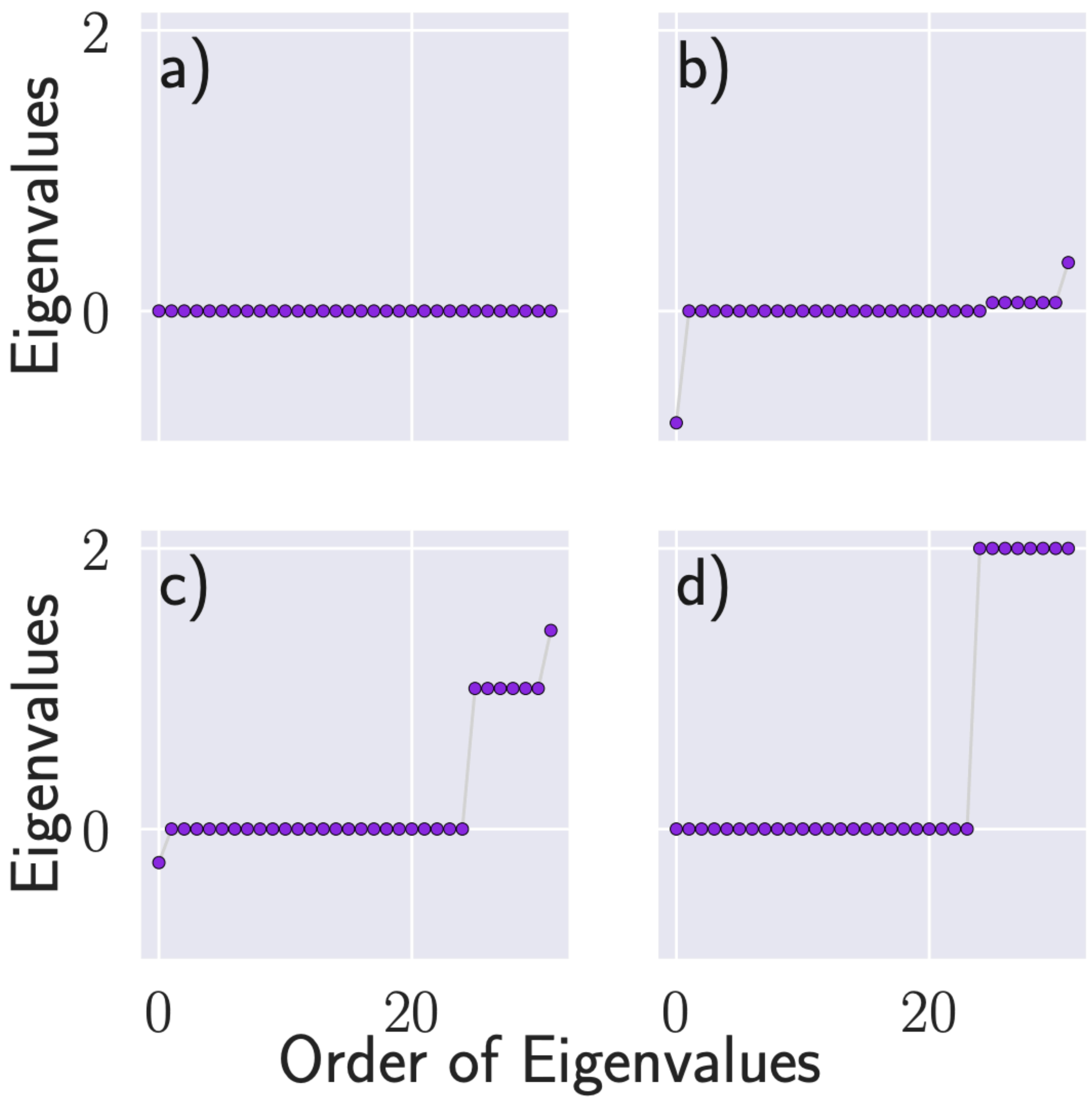
Methods to try avoid barren plateaus

The Hessian matrix

$$\frac{\partial C}{\partial \theta} = \begin{bmatrix} \frac{\partial C}{\partial \theta_1} \\ \frac{\partial C}{\partial \theta_2} \\ \vdots \end{bmatrix} \qquad \mathbf{H} = \begin{bmatrix} \frac{\partial^2 C}{\partial \theta_1^2} & \frac{\partial^2 C}{\partial \theta_1 \theta_2} & \cdots \\ \frac{\partial^2 C}{\partial \theta_2 \theta_1} & & \\ \vdots & & \end{bmatrix}$$

Loss landscape of quantum models





Source: Huembeli, Patrick, and Alexandre Dauphin. "Characterizing the loss landscape of variational quantum circuits." arXiv preprint arXiv:2008.02785 (2020).

Methods to try avoid barren plateaus

- Introduce structure into the circuit (encoding with data) or initialising layers to the identity
  Grant, Edward, et al. "An initialization strategy for addressing barren plateaus in parametrized quantum circuits." Quantum 3 (2019): 214.

- Use a local cost function
  Cerezo, M., et al. "Cost-function-dependent barren plateaus in shallow quantum neural networks." *arXiv preprint arXiv:2001.00550* (2020).

- Higher order optimisation methods
  Huembeli, Patrick, and Alexandre Dauphin. "Characterizing the loss landscape of variational quantum circuits." arXiv preprint arXiv:2008.02785 (2020).
  Cerezo, M., and Patrick J. Coles. "Impact of Barren Plateaus on the Hessian and Higher Order Derivatives." arXiv preprint arXiv:2008.07454 (2020).

Loss landscapes

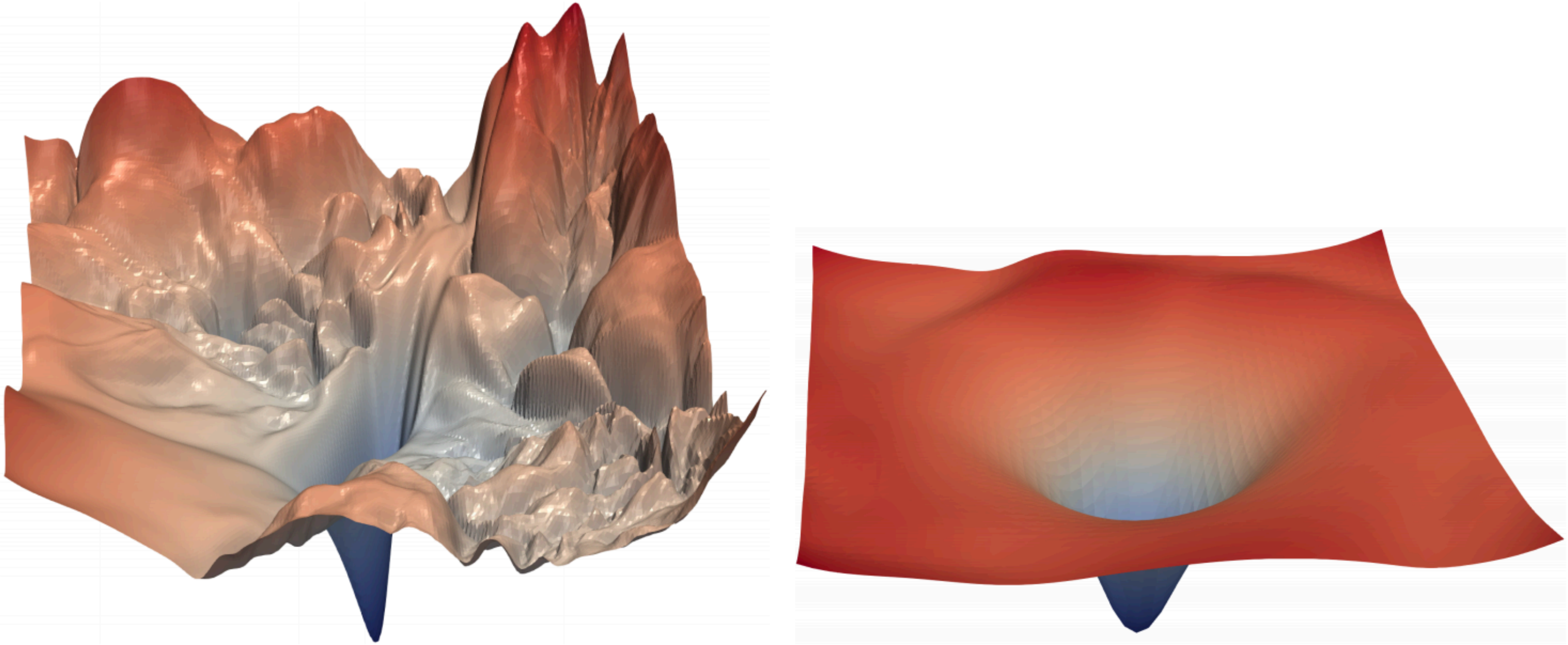


Image source: Li, Hao, et al. "Visualizing the loss landscape of neural nets." *Advances in Neural Information Processing Systems*. 2018.

Jump to the code

Thank you!

NITheP Mini-school

Amira Abbas
@AmiraMorphism
amyami187@gmail.com

UNIVERSITY OF KWAZULU-NATAL ™
INYUVESI YAKWAZULU-NATALI

NITheP
National Institute for Theoretical Physics