Centre for Quantum Technologies

NUS National University of Singapore | Computing

Post-variational quantum neural networks

Po-Wei Huang, Patrick Rebentrost

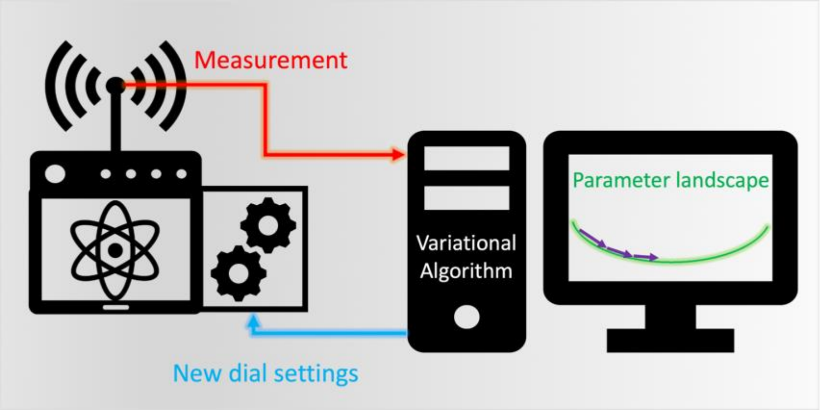
arXiv:2307.10560

Nov 22, 2023, QTML, Geneva, Switzerland

QTML2023

Methods in quantum machine learning

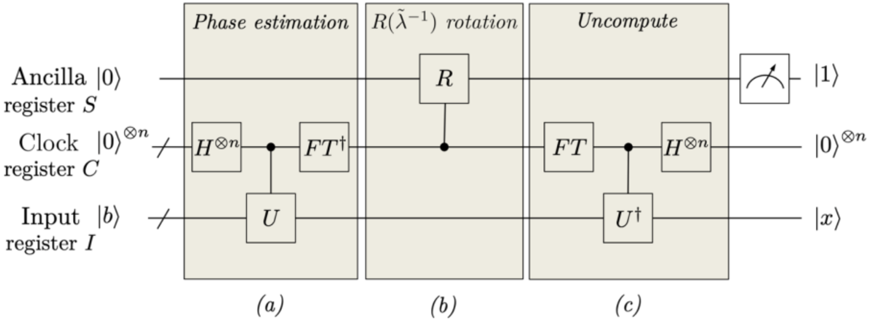
Variational quantum algorithms



https://quantum-journal.org/papers/q-2019-10-07-191/

Barren plateaus

Fault tolerant algorithms

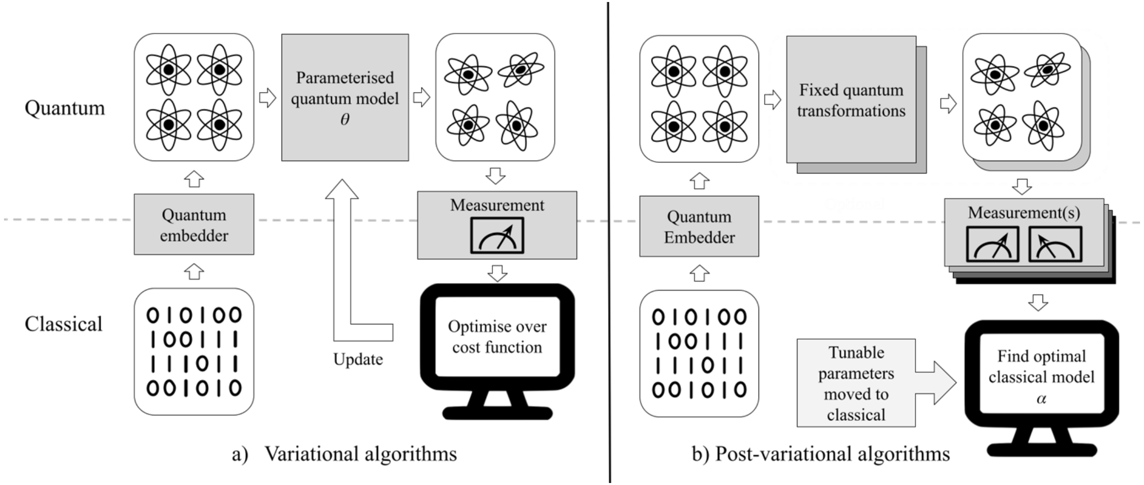


https://arxiv.org/abs/1802.08227

Limited application on current hardware

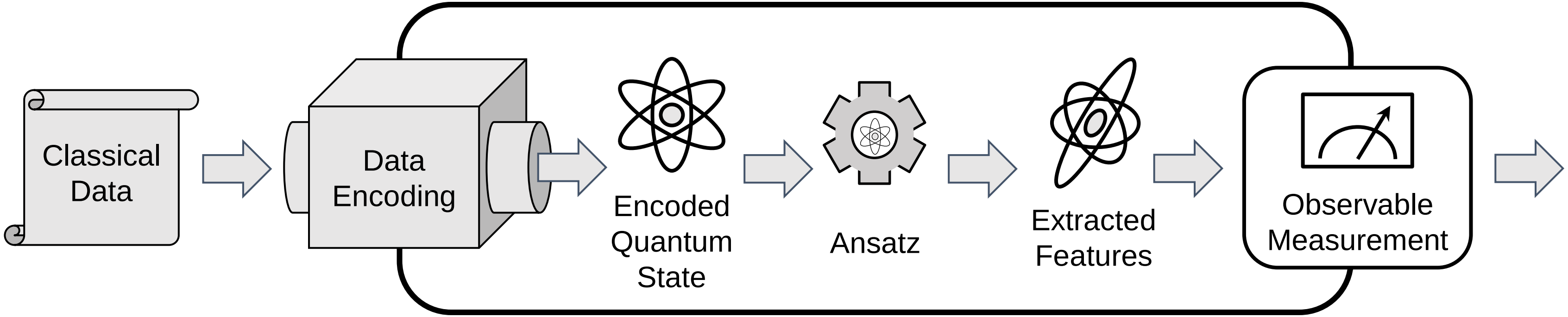
Post-variational strategies

- Convert problem from optimization with quantum circuits to a classical convex optimization problem of combination of quantum circuits.
  - Use quantum computers to compute values from fixed Ansätze/observables and use classical computer to find optimal combination.



a) Variational algorithms b) Post-variational algorithms

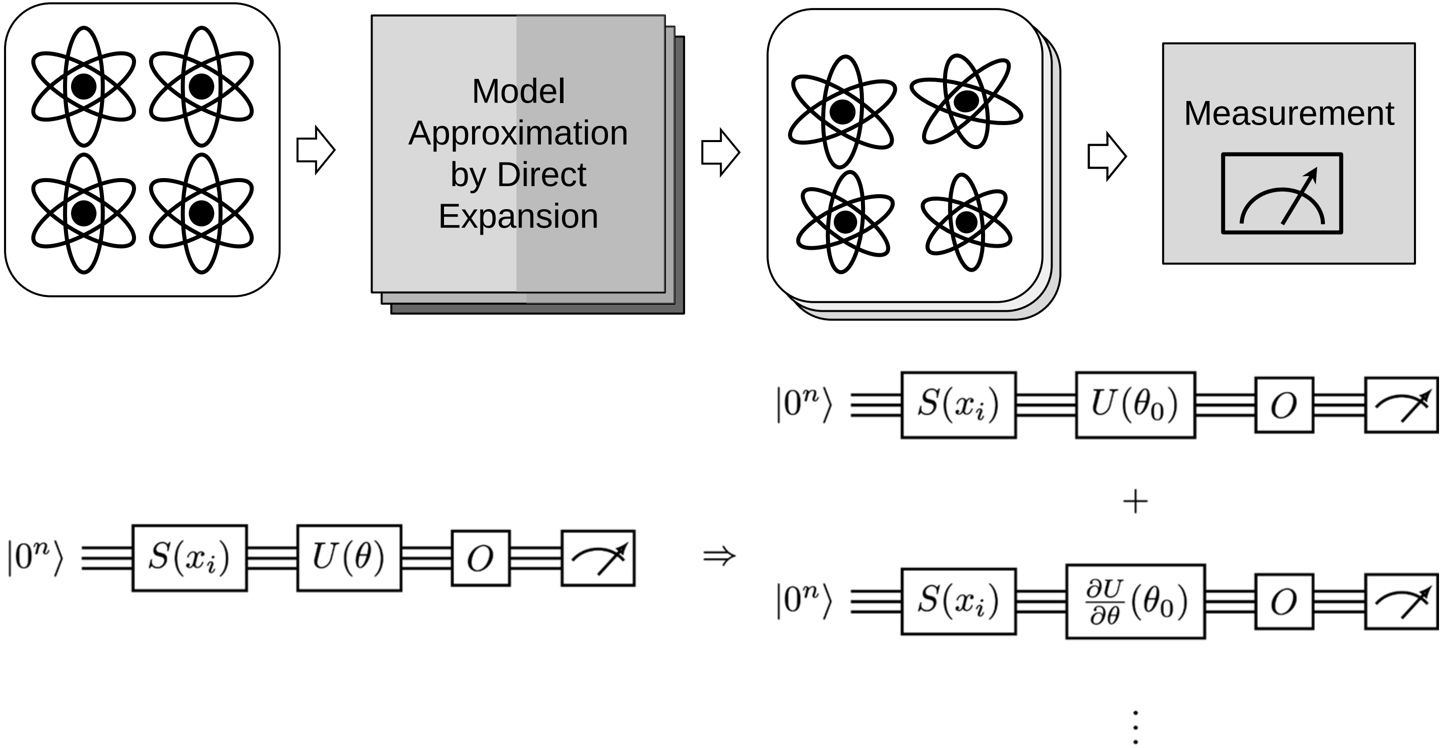
From variational to post-variational



$$\mathbb{E}[X] = \langle 0^n | S^\dagger(x) \underbrace{U^\dagger(\theta) O U(\theta)}_{\mathcal{O}(\theta)} S(x) | 0^n \rangle$$

$$\sum U_j \qquad \sum \mathcal{O}_j$$

Ansatz expansion strategy

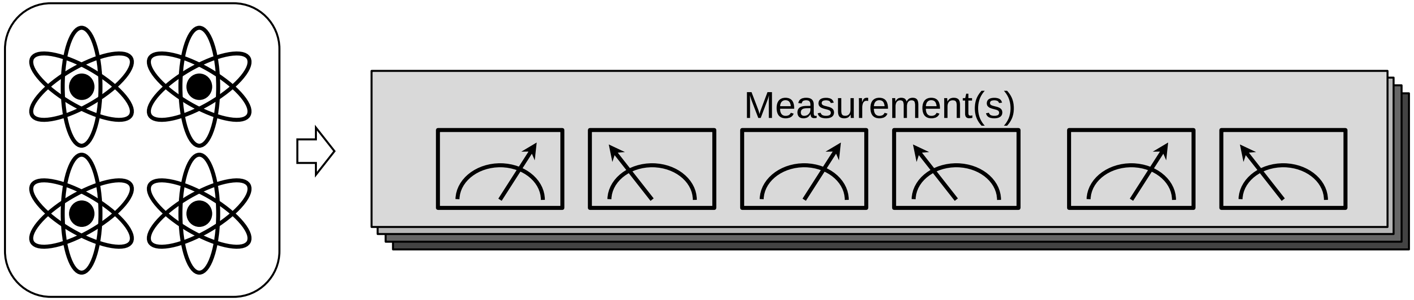


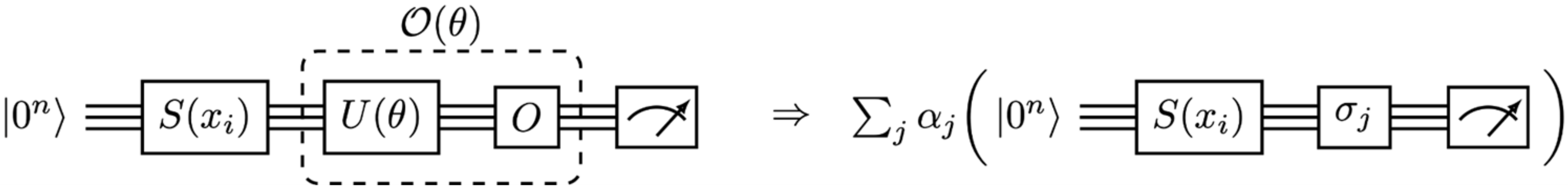
Ansatz expansion strategy

- Construct gradients and higher-order derivatives with parameter shift rules.
  - Number of measurements required are exponential to the derivative order to be taken.
  - A low order truncation is required but may not be an effective enough estimation.
- Strategies to avoid exponentially small gradients:
  - Initialization of parameters such that Ansatz blocks eval to identity → Avoid formation of 2-designs [GWOB19]

[GWOB19] E. Grant, L. Wossnig, M. Ostaszewski, and M. Benedetti, An initialization strategy for addressing barren plateaus in parametrized quantum circuits, Quantum 3, 214 (2019).

Observable construction strategy





Observable construction strategy

- Build an arbitrary observable using linear combination of Pauli strings.
  - To build a *k*-local Hamiltonian, Pauli strings exponential to *k* is needed.
- Theoretical guarantees with low degree approximations [HCP22]
  - For any distribution over *n*-qubits invariant under single qubit Clifford gate
  - Observable truncated at $k = \left\lceil \log_{1.5}\left(\frac{1}{\epsilon}\right) \right\rceil$ is epsilon close to optimal loss such that

$$\mathbb{E}_{\rho \in \mathcal{D}} \left| \mathrm{tr}(O\rho) - \mathrm{tr}(O^{(k)}\rho) \right|^2 \leq (2/3)^k \|O\|^2$$

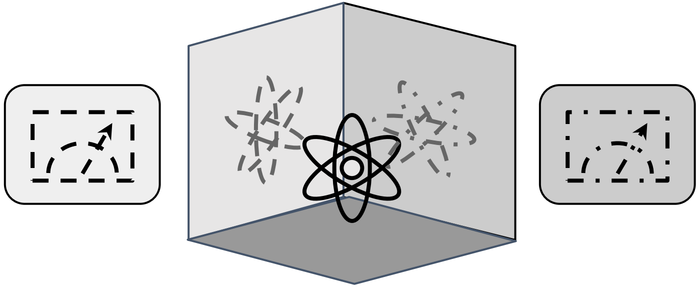
[HCP22] H.-Y. Huang, S. Chen, and J. Preskill, Learning to predict arbitrary quantum processes (2023), arXiv:2210.14894 [quant-ph].

Observable construction strategy

- Classical shadow tomography [HKP20] estimation
  - Use global random measurement on Pauli basis to estimate local observables
  - Dependency exponential to locality of observables and logarithmic to number of observables

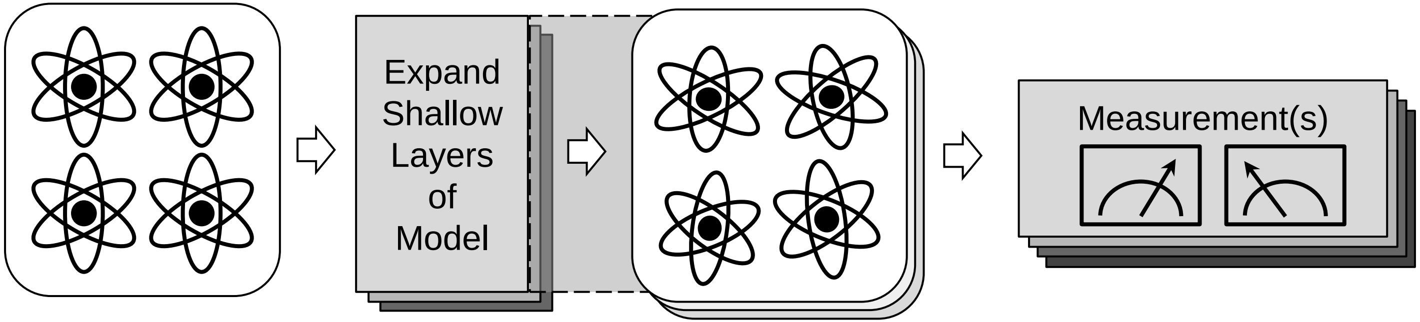
$$O(4^n) \rightarrow \begin{cases} O((3n)^{(k+1)}(k+1)\log n/\epsilon^2) & \text{direct measurement} \\ O(3^k k \log n/\epsilon^2) & \text{classical shadows} \end{cases}$$

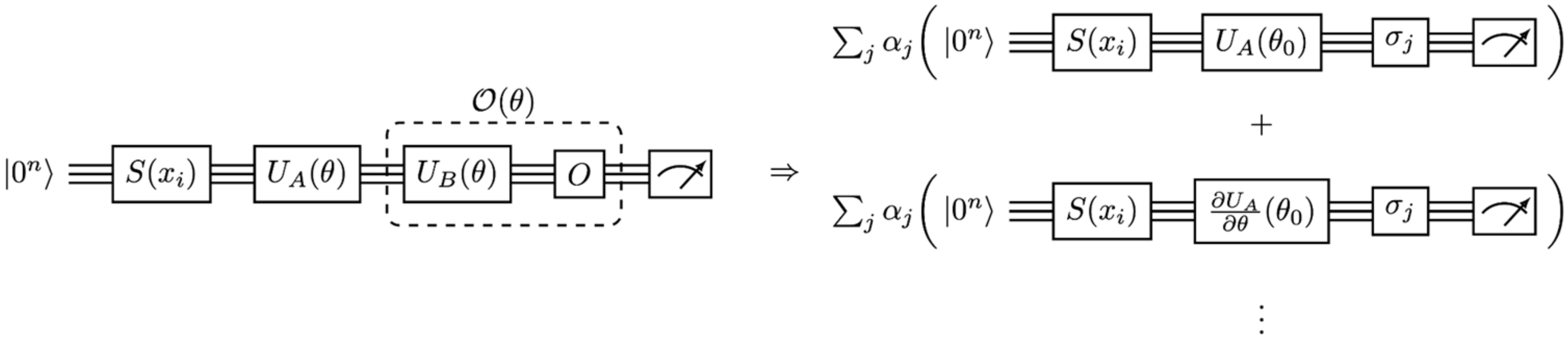
Reduction by $O(n^k)$



[HKP20] H.-Y. Huang, R. Kueng, and J. Preskill, Predicting many properties of a quantum system from very few measurements, Nature Physics 16, 1050 (2020).

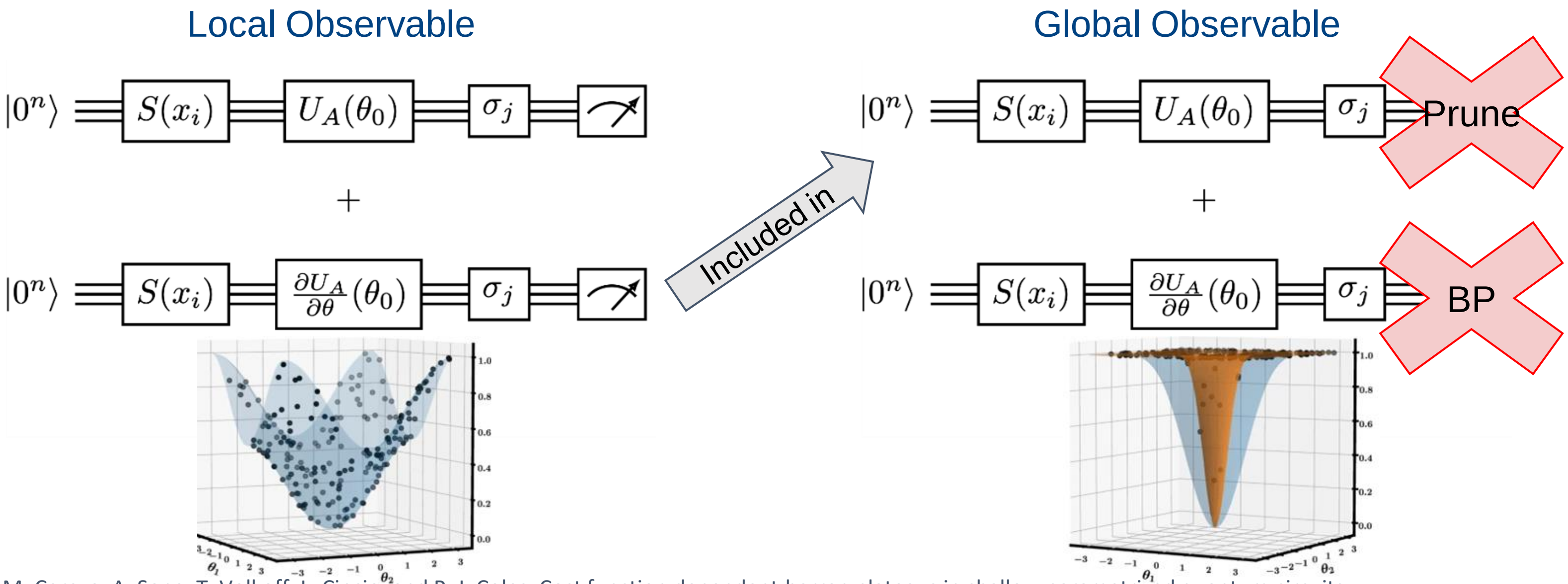
Hybrid strategy





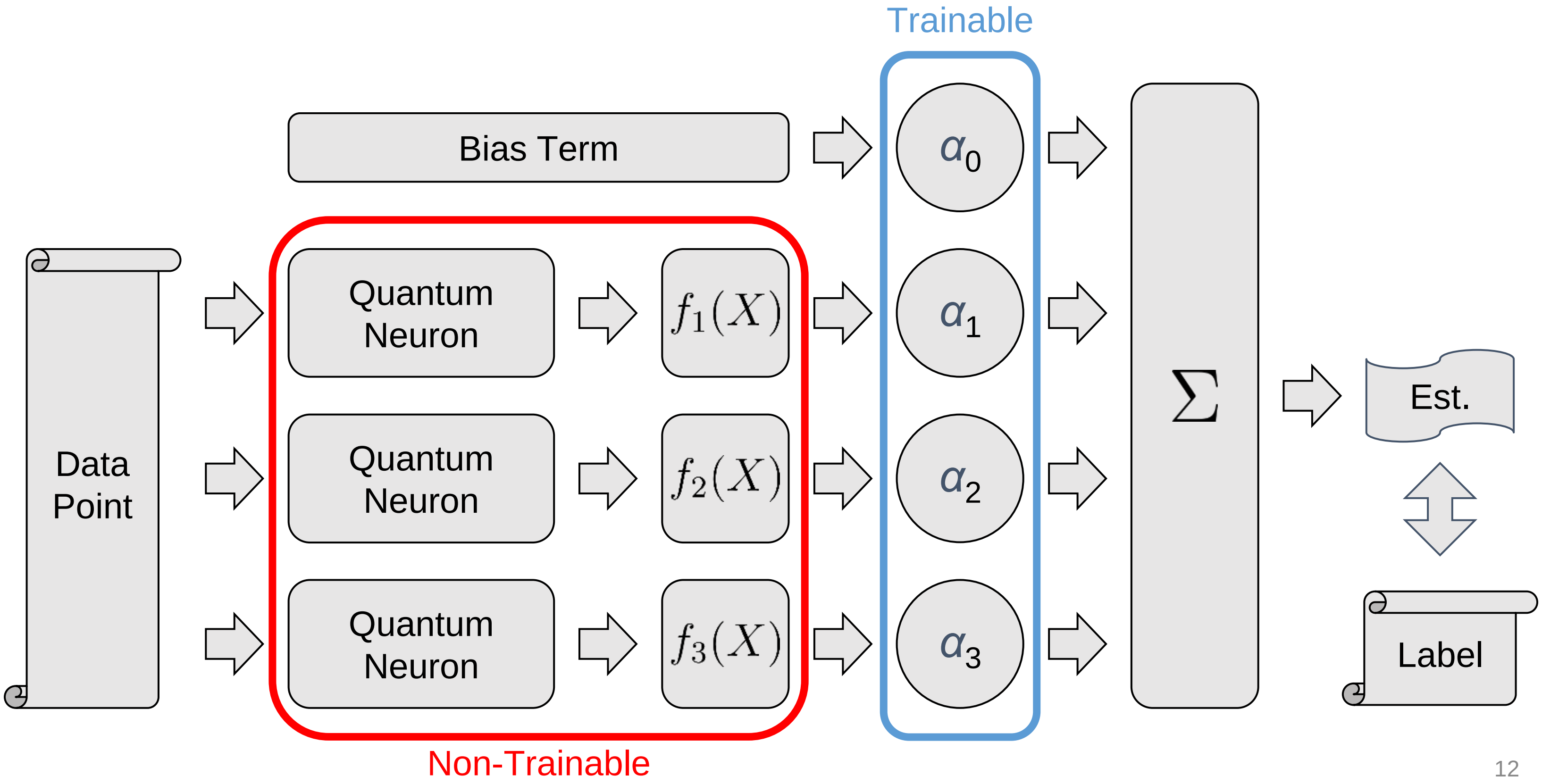
Hybrid strategy

- Construct gradients plus build an arbitrary observable.

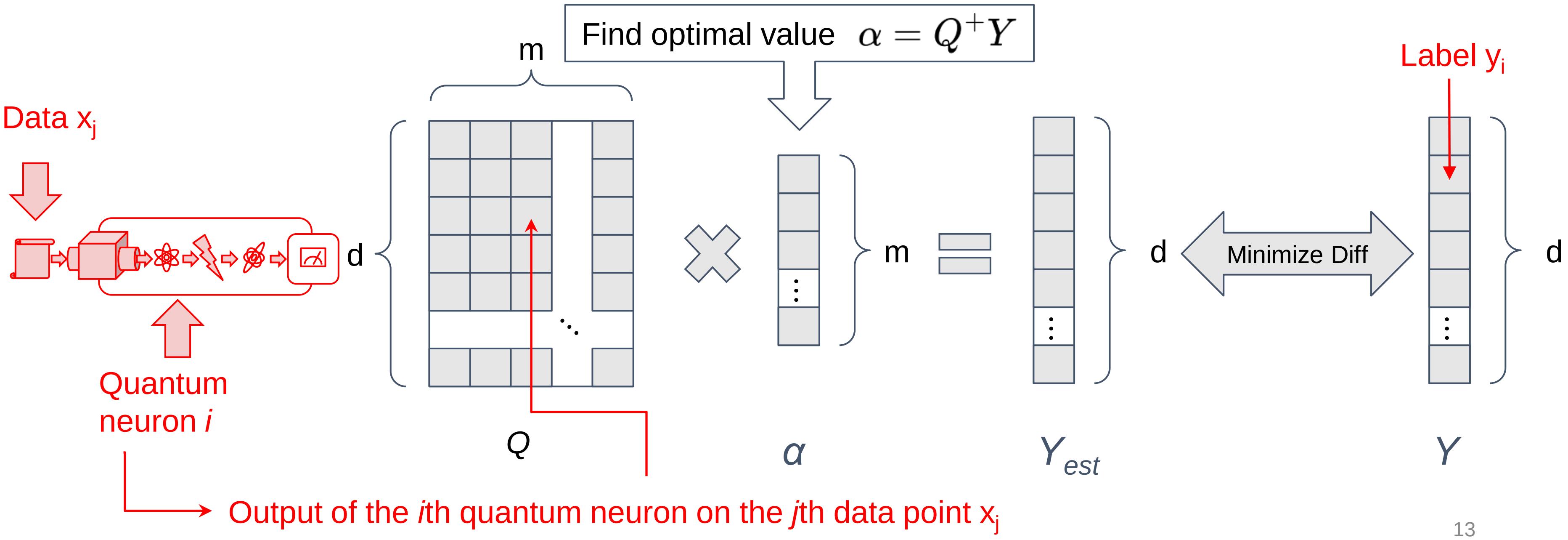


[CSV+21] M. Cerezo, A. Sone, T. Volkoff, L. Cincio, and P. J. Coles, Cost function dependent barren plateaus in shallow parametrized quantum circuits, Nature Communications 12 (2021).

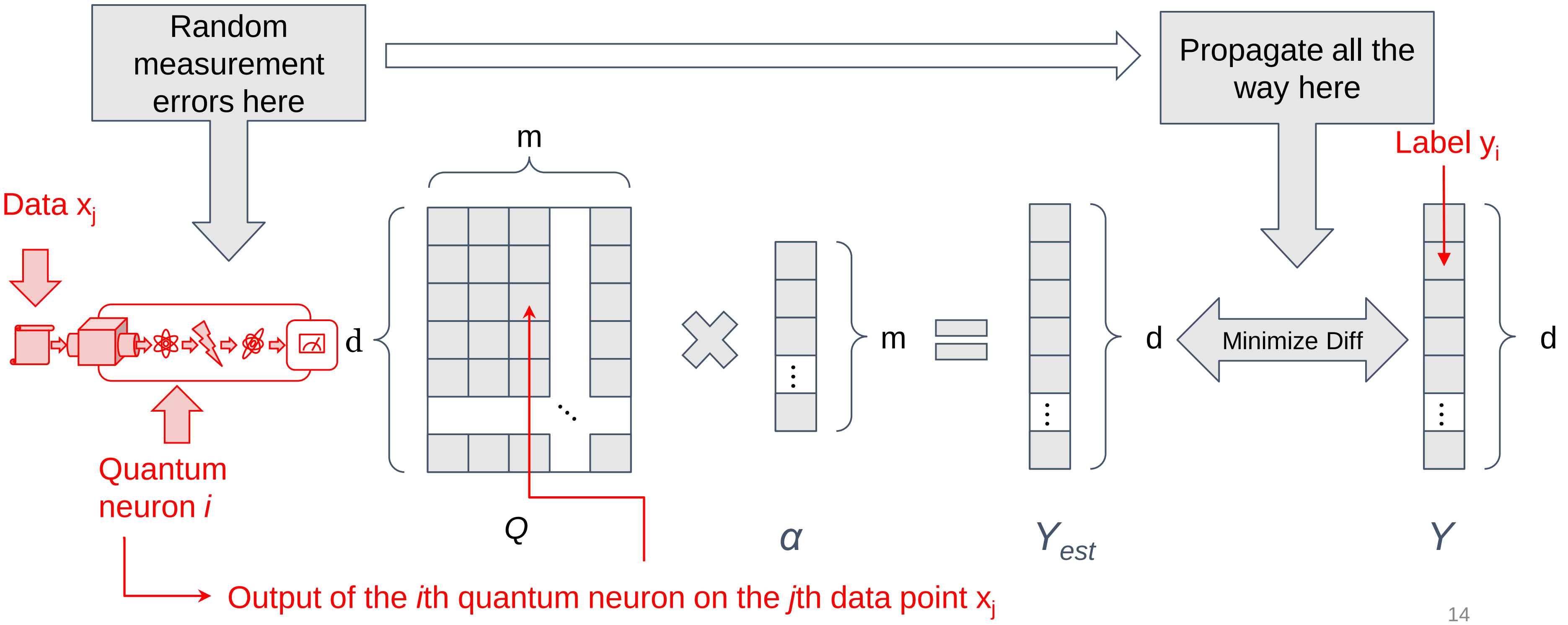
Neural network system



Classical optimization



Classical optimization…with errors



Measurements related to total loss

- Given the computed coefficients α̃ from Q̃ constructed from measurements, fix ϵ such that for root mean square error (RMSE) loss

$$\left| \|Q\alpha - Y\|_2 - \|\tilde{Q}\alpha - Y\|_2 \right| < \epsilon$$

- The total number of measurements for post-variational neural networks is

$$\mathbf{O}\left(\frac{m^3 d}{\epsilon} \log \frac{md}{\delta}\right)$$

Or, with classical shadows estimation,

$$\mathbf{O}\left(\frac{m^2 pd \max \|O\|_s}{\epsilon} \log \frac{md}{\delta}\right)$$

Experiment results

	Validation			Testing		
Strategy	Loss	Accuracy	F1	Loss	Accuracy	F1
Ansatz	0.70	0.49	0.57	0.69	0.50	0.66
Observable	0.40	0.83	0.82	0.42	0.82	0.82
Hybrid	0.15	0.94	0.94	0.20	0.91	0.90
Variational [7]	-	0.97	-	-	-	-