The power and limitations of learning quantum dynamics incoherently

Sofiene Jerbi, Joe Gibbs, Manuel Rudolph, Matthias Caro,
Patrick Coles, Hsin-Yuan Huang, Zoë Holmes

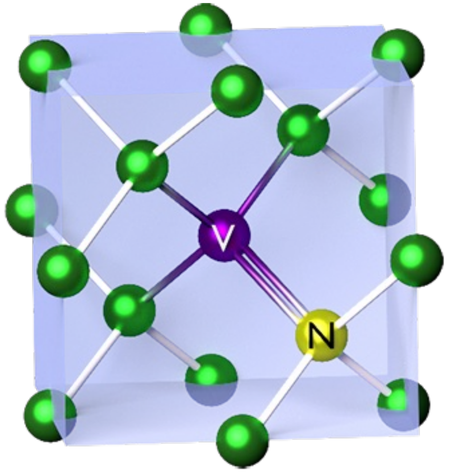
universität innsbruck

Los Alamos NATIONAL LABORATORY EST. 1943

arXiv: 2303.12834

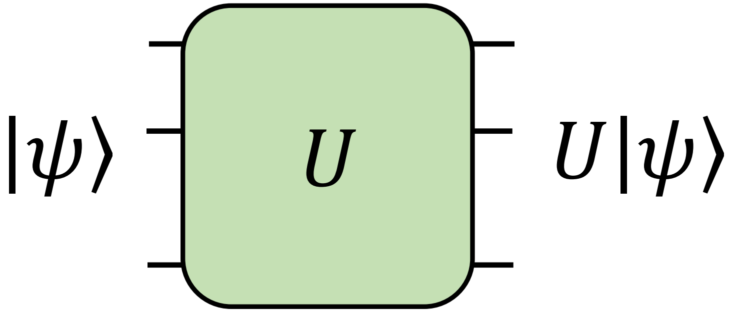
QTML 2023
20/11/2023

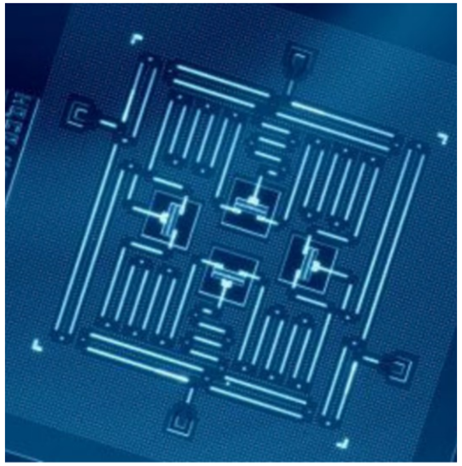
Learning quantum dynamics



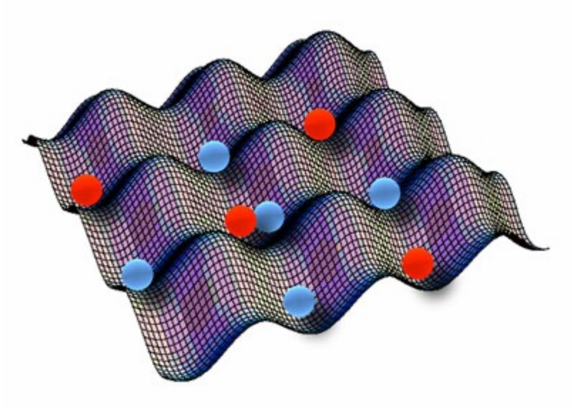
NV centers

Quantum dynamics described by unitary evolution





Superconducting qubits

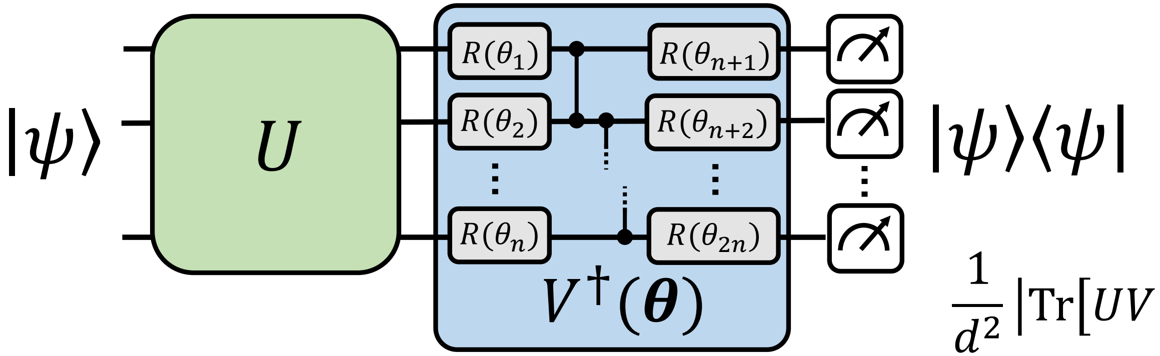


Ultracold atoms

Quantum compiling *in the wild*

- Most popular scenario: variational quantum compiling

Generally assumes a lot of knowledge about U



$$\frac{1}{d^2}\left|\mathrm{Tr}\left[UV^\dagger(\boldsymbol{\theta})\right]\right|^2$$

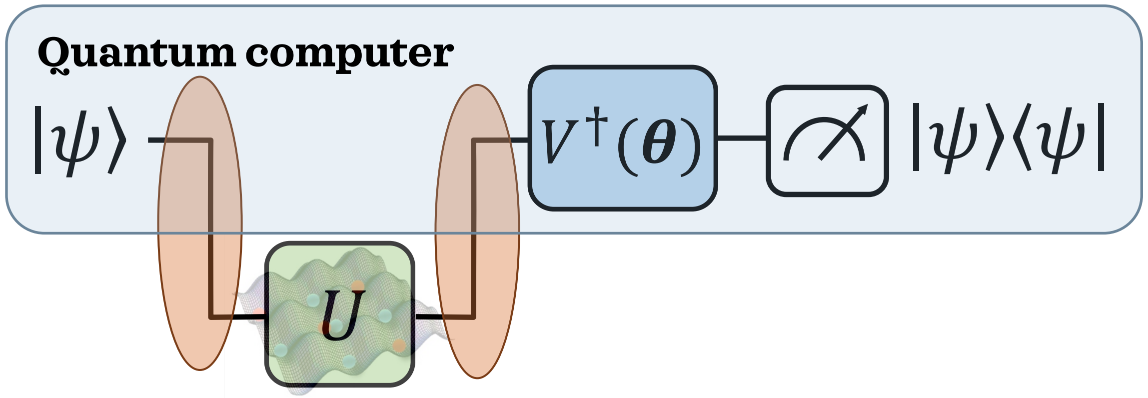
U can be:

- implemented on a Q. computer (e.g., QFT),
- simulated classically (e.g., short-time Hamiltonian simulation)

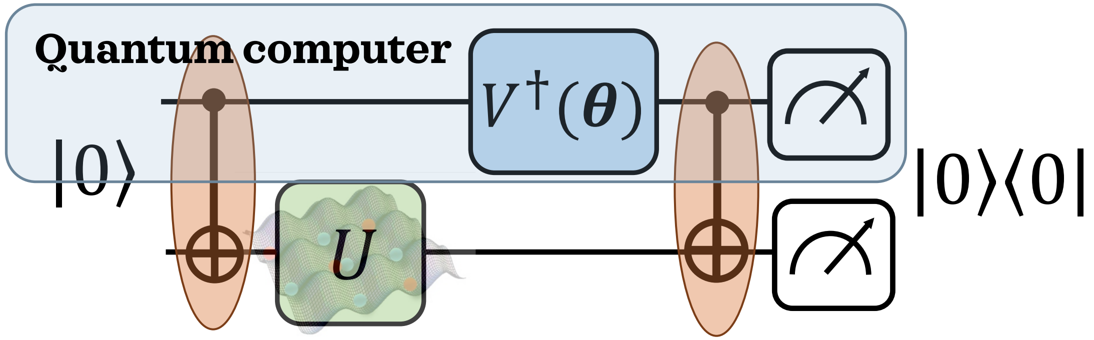
<u>We want to compress the implementation.</u>

- But what about learning totally unknown dynamics?

Assumes powerful interaction with system

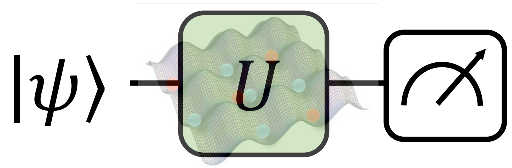


or

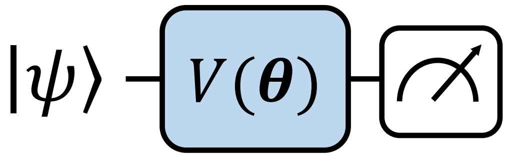


What can be done *incoherently*?

Step 1:

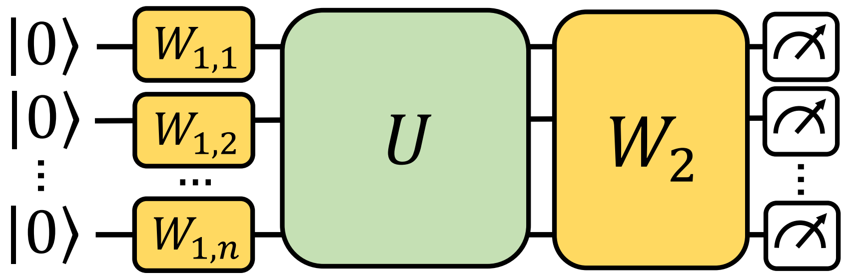


Step 2:



- We investigate conditions that allow to *simulate* coherent learning, and their limitations

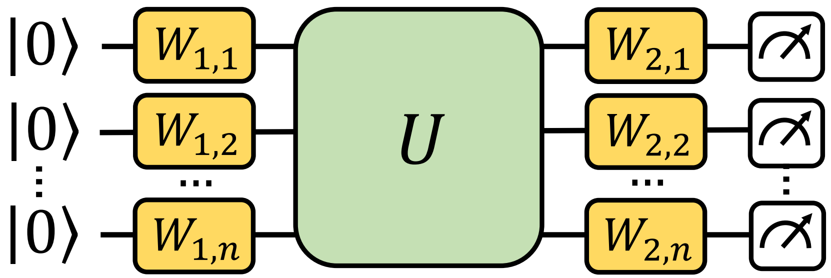
Deep measurements



Power: Can simulate coherent learning for **efficiently representable** $V(\boldsymbol{\theta})$.

Limitations: Not computationally efficient

Shallow measurements



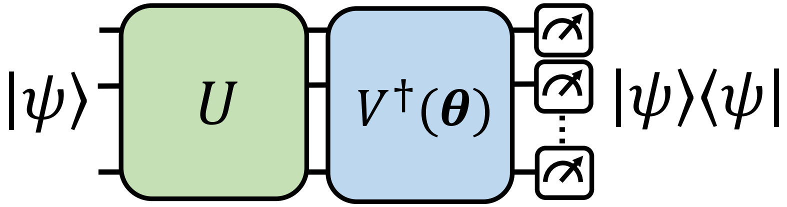
Power: Can simulate coherent learning for **shallow** $V(\boldsymbol{\theta})$.

Limitations: Needs **exponentially** many measurements to learn $\mathcal{O}(n)$-depth U.

What can be done *coherently*?

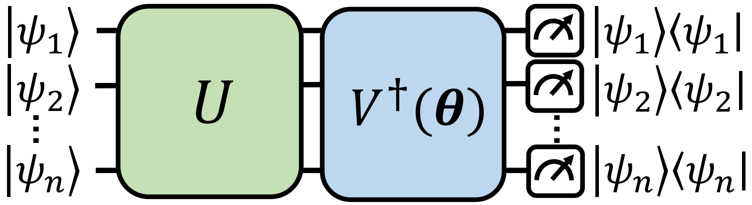
Out-of-distribution generalization

Global Haar



$$1 - \mathcal{C}_1(\boldsymbol{\theta}) = \mathbb{E}_{|\psi\rangle\sim Haar}\left[\left|\langle\psi|UV^\dagger(\boldsymbol{\theta})|\psi\rangle\right|^2\right] \propto \frac{1}{d^2}\left|\mathrm{Tr}\left[UV^\dagger(\boldsymbol{\theta})\right]\right|^2$$

Local Haar

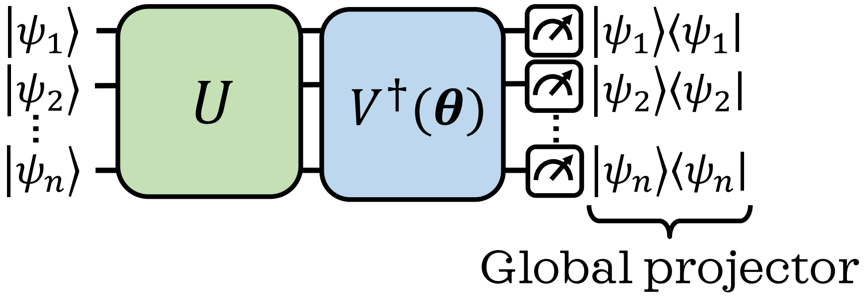


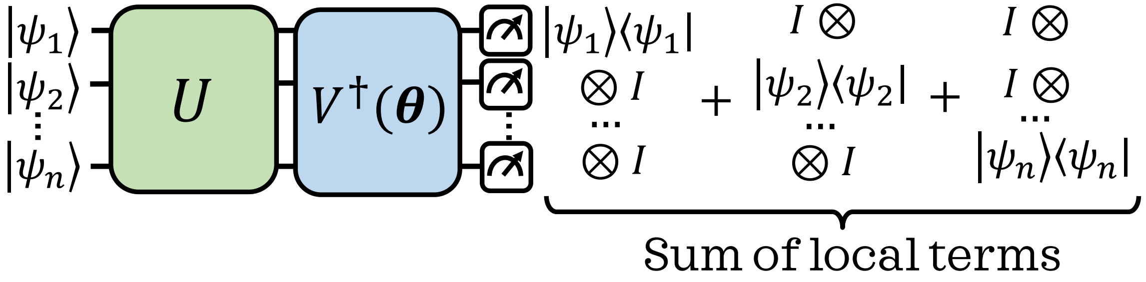
$$1 - \mathcal{C}_2(\boldsymbol{\theta}) = \mathbb{E}_{|\psi\rangle\sim Haar^{\otimes n}}\left[\left|\langle\psi|UV^\dagger(\boldsymbol{\theta})|\psi\rangle\right|^2\right]$$

$$\frac{1}{2}\mathcal{C}_1(\boldsymbol{\theta}) \leq \mathcal{C}_2(\boldsymbol{\theta}) \leq 2\mathcal{C}_1(\boldsymbol{\theta})$$

[1] Caro *et al.*, Nature Communications (2023)

Local costs





$$\mathcal{C}_{local}(\boldsymbol{\theta}) \leq \mathcal{C}_{global}(\boldsymbol{\theta}) \leq n\mathcal{C}_{local}(\boldsymbol{\theta})$$

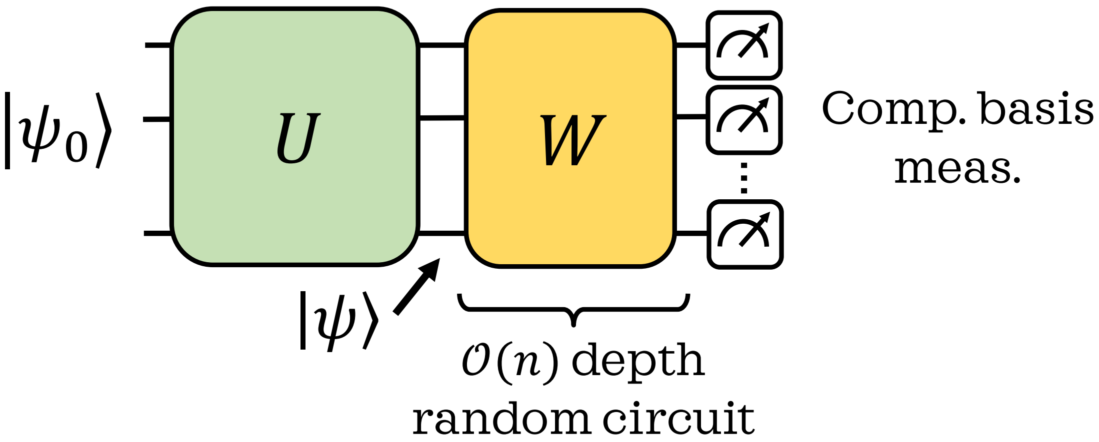
[2] Khatri *et al.*, Quantum (2019)

[3] Cerezo *et al.*, Nat. Comm. (2021)

Deep measurements: *power*

- Can simulate the coherent setting for **efficiently representable** $V(\boldsymbol{\theta})$.

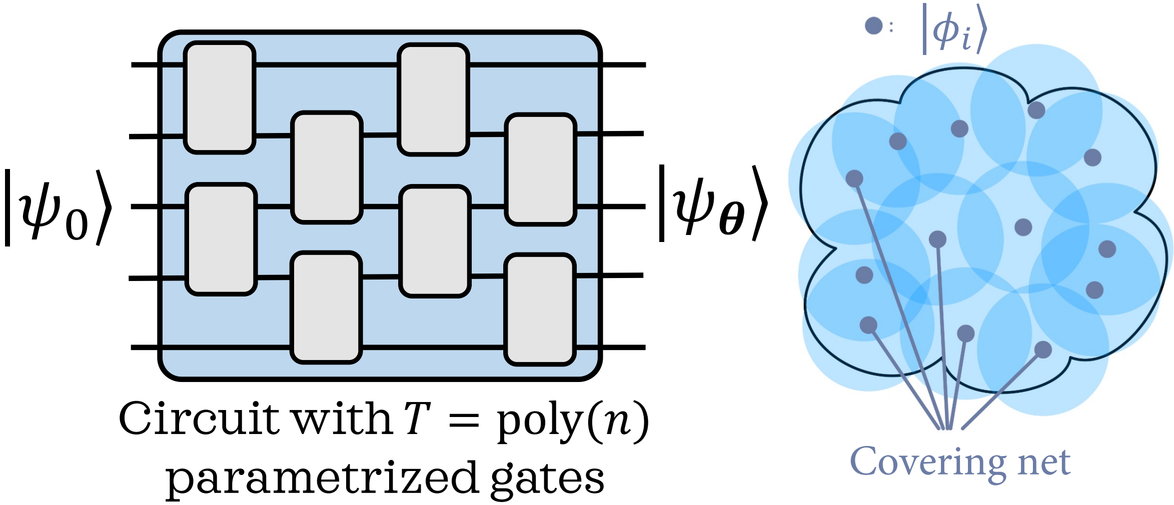
Clifford shadows



$\mathcal{O}\left(\frac{\log(M)}{\varepsilon^2}\right)$ measurements allow to estimate $\{|\langle\psi|\phi_1\rangle|^2, \dots, |\langle\psi|\phi_M\rangle|^2\}$ to prec. ε

[3] Huang *et al.*, Nature Physics (2020)

Covering nets



$\mathcal{O}\left(\exp(T \log \frac{T}{\varepsilon})\right)$ states $|\phi_i\rangle$ are sufficient to guarantee: $\forall \boldsymbol{\theta}, \exists i \ / \ \||\psi_{\boldsymbol{\theta}}\rangle - |\phi_i\rangle\| \leq \varepsilon$

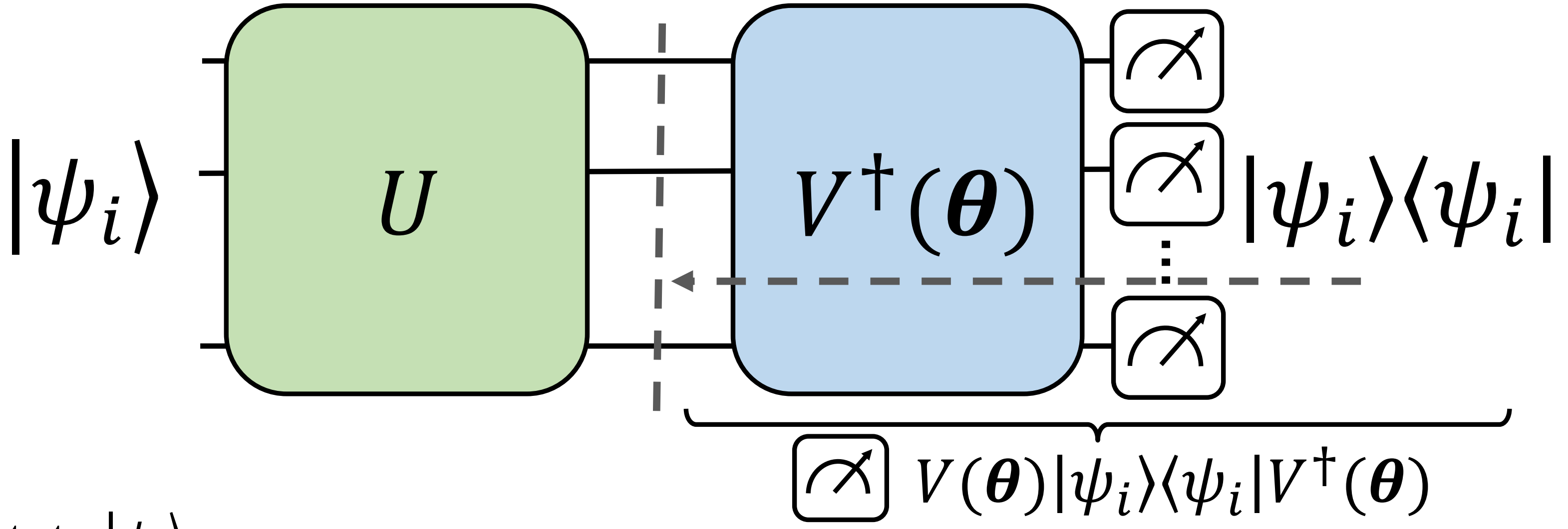
[4] Caro *et al.*, Nature Communications (2022)

Deep measurements: *power*

- Can simulate the coherent setting for **efficiently representable** $V(\boldsymbol{\theta})$.

$\mathcal{O}\left(\frac{\log(M)}{\varepsilon^2}\right)$ measurements allow to estimate $\{|\langle\psi|\phi_1\rangle|^2, \ldots, |\langle\psi|\phi_M\rangle|^2\}$ to prec. ε

$\mathcal{O}\left(\exp(T\log\frac{T}{\varepsilon})\right)$ states $|\phi_i\rangle$ are sufficient to guarantee: $\forall\boldsymbol{\theta}, \exists i \,/\, \| |\psi_{\boldsymbol{\theta}}\rangle - |\phi_i\rangle \| \leq \varepsilon$



For each training state $|\psi_i\rangle$:

$\mathcal{O}\left(T\log\frac{T}{\varepsilon}\varepsilon^{-2}\right)$ measurements of $U|\psi_i\rangle$ allow to estimate any $|\langle\psi_i|U^\dagger|\psi_{\boldsymbol{\theta}}\rangle|^2 = |\langle\psi_i|U^\dagger V(\boldsymbol{\theta})|\psi\rangle|^2$ to prec. ε

Deep measurements: *power*

Simulating coherent learning

$\mathcal{O}\left(T \log\frac{T}{\varepsilon}\varepsilon^{-2}\right)$ measurements of $U|\psi_i\rangle$

estimate $\left|\langle\psi_i|UV^\dagger(\boldsymbol{\theta})|\psi_i\rangle\right|^2$

$\Rightarrow \tilde{\mathcal{O}}(NT\varepsilon^{-2})$ meas. to estimate $\widetilde{\mathcal{C}_2}(\boldsymbol{\theta})$

$= 1 - \sum_{i=1}^{N}\left|\langle\psi_i|UV^\dagger(\boldsymbol{\theta})|\psi_i\rangle\right|^2$

&

Generalization from few training samples

$$\mathcal{C}_1(\boldsymbol{\theta}) \leq 2\widetilde{\mathcal{C}_2}(\boldsymbol{\theta}) + \mathcal{O}\left(\sqrt{\frac{T\log T}{N}}\right)$$

Expected error on global Haar | Training error on local Haar | Generalization gap

$$\mathcal{C}_1(\boldsymbol{\theta}) = 1 - \frac{1}{d^2}\left|\mathrm{Tr}\left[UV^\dagger(\boldsymbol{\theta})\right]\right|^2$$
$$\mathcal{C}_2(\boldsymbol{\theta}) = 1 - \mathbb{E}_{|\psi\rangle\sim Haar^{\otimes n}}\left[\left|\langle\psi|UV^\dagger(\boldsymbol{\theta})|\psi\rangle\right|^2\right]$$

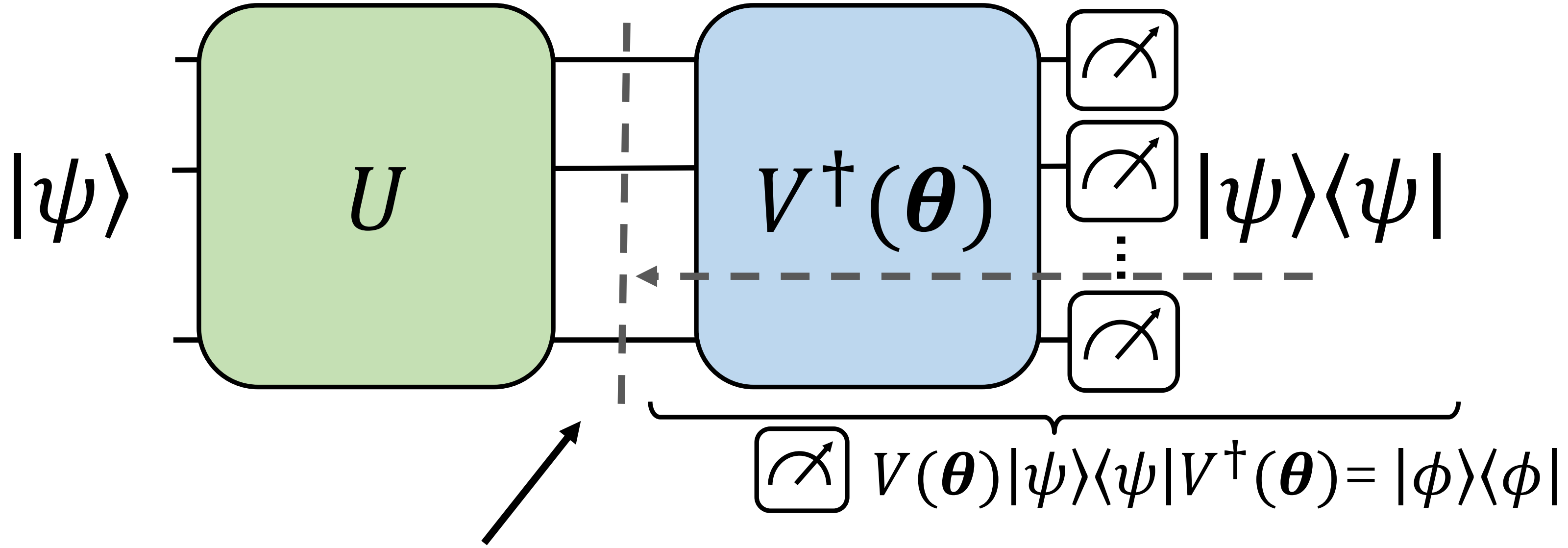
(& brute-force search 🤫)

$\Rightarrow$ Any $poly(n)$-sized unitary U can be learned within Hilbert-Schmidt error ε using $\tilde{\mathcal{O}}\left(\frac{poly(n)}{\varepsilon^4}\right)$ calls to U.

[4] Caro *et al.*, Nat. Comm. (2022) [1] Caro *et al.*, Nat. Comm. (2023)

Deep measurements: *limitations*

- Deep flaw of **Clifford/global** classical shadows: $|\langle\psi|U^\dagger|\phi\rangle|^2 \longrightarrow \mathrm{Tr}[\hat{\rho}|\phi\rangle\langle\phi|]$
  - sample-complexity efficient but **computationally hard**

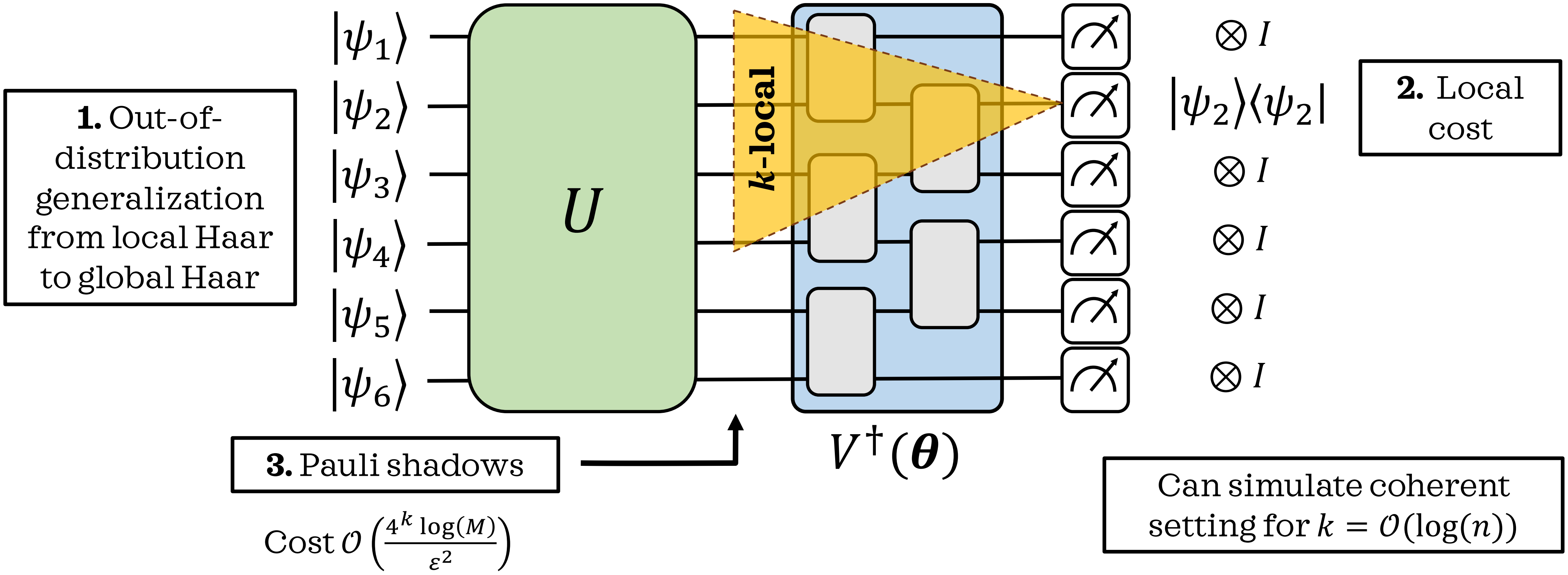


Shadow: $\hat{\rho} = (2^n + 1)|\mathrm{Stab}_i\rangle\langle\mathrm{Stab}_i| - \mathbb{I}$

Precise estimation only efficient for $|\phi\rangle$ a stabilizer state (or close)

Shallow measurements: *power*

- Can simulate coherent learning for **shallow** $V(\boldsymbol{\theta})$.
  - Combination of **3 results**



Shallow measurements: *power*

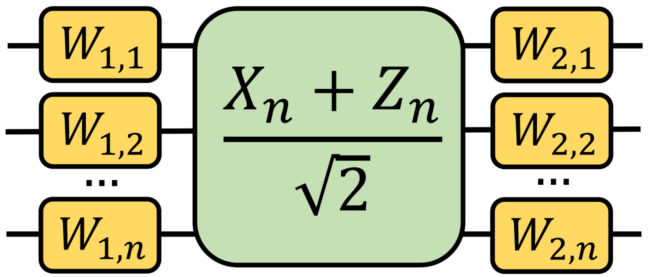
Simulating coherent learning & **Generalization from few training samples**

(& brute-force search 🤫)

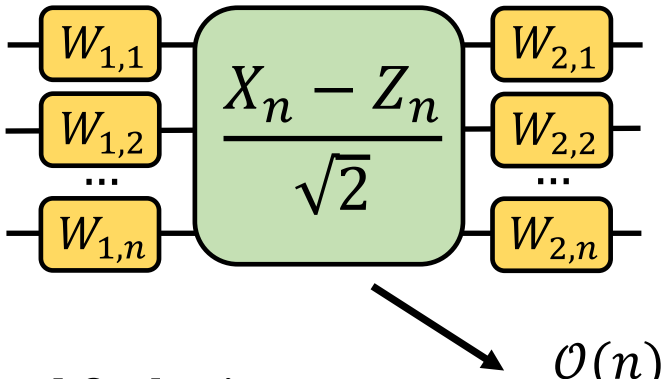
$\Rightarrow$ Any $log(n)$-depth (1D) unitary U can be learned within Hilbert-Schmidt error ε using $\tilde{O}\left(\frac{poly(n)}{\varepsilon^4}\right)$ calls to U.

Shallow measurements: *limitations*

- Need **exponentially** many measurements to learn $\mathcal{O}(n)$ depth unitaries.



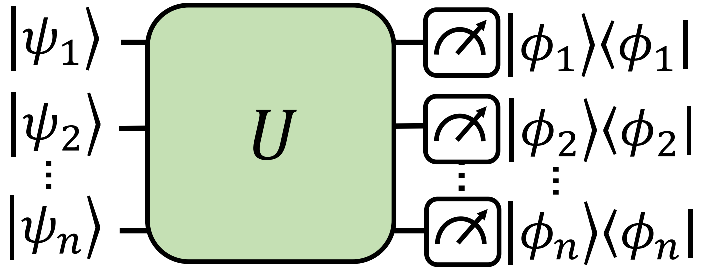
v.s.



$\mathcal{O}(n)$ depth GHZ-like

$W_{i,j}$: sampled from local 2-design

Allowed measurements:

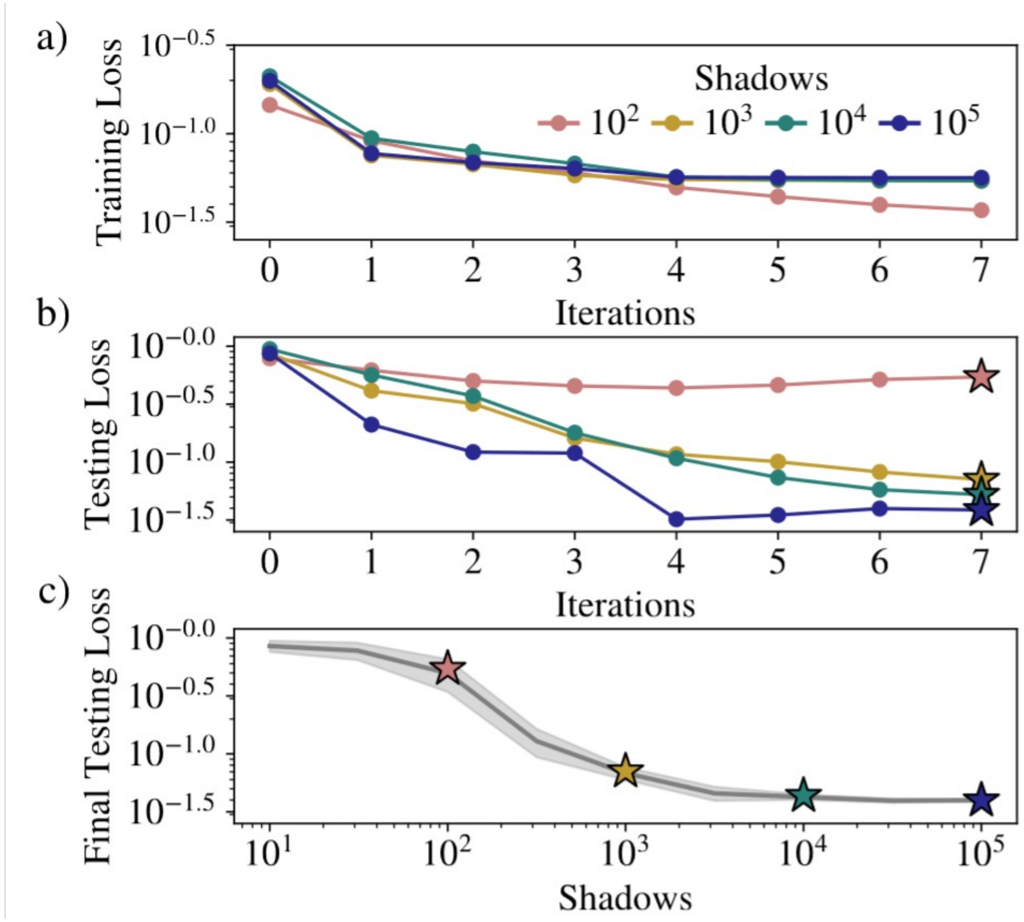


For T measurements:

$$\mathbb{E}_W[d_{TV}(p_+, p_-)] \leq \frac{T}{2^{\mathcal{O}(n)}}$$

Holds also for adaptive measurements

Experimental demonstration (shallow meas.)



Transverse-Field Ising Model

$$H_{\text{Ising}} = \sum_{i=0}^{14} Z_i Z_{i+1} + \sum_{i=0}^{15} \alpha_i X_i$$

$\alpha_i \sim \mathcal{N}(0{,}0.5)$

`ibmq_kolkata`

U: Trotterization for time $\Delta t = 0.1$

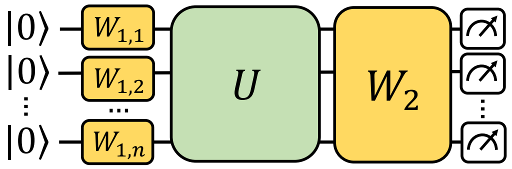
Pauli shadows of 2 training states $\{U|\psi_i\rangle\}_{i=0}^{1}$

$V(\boldsymbol{\theta})$: variational Trotter-circuit

Conclusions & Outlook

- We investigate conditions that allow to *simulate* coherent learning, and their limitations

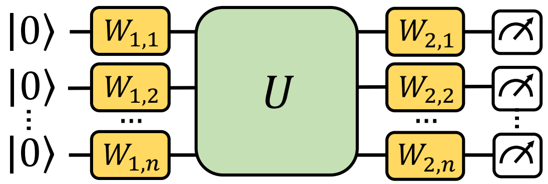
Deep measurements



Power: Works for **efficiently representable** unitaries.

Limitations: Not computationally efficient

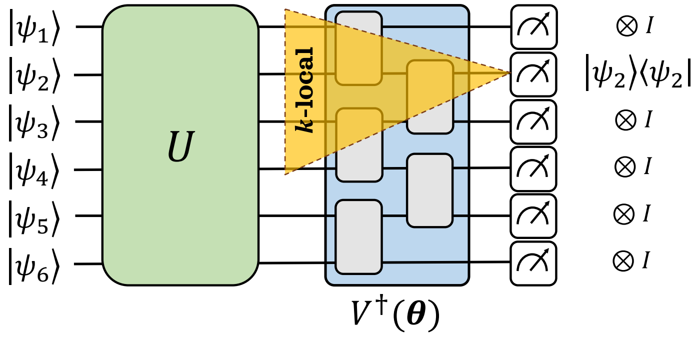
Shallow measurements



Power: Works for **shallow** unitaries.

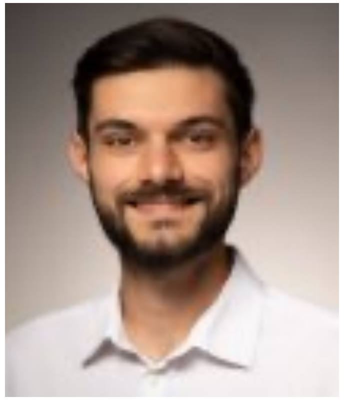
Limitations: Needs **exponentially** many measurements to learn $\mathcal{O}(n)$-depth U.

- Rigorous limitations of deep measurements setting?
  - **Pseudo-random** unitaries: efficiently representable but computationally undistinguishable from Haar random
- Can we go beyond $\mathcal{O}(\log(n))$ depth with shallow measurements? (Approximate locality)



Special thanks

Joe Gibbs



Manuel Rudolph

Matthias Caro

Patrick Coles

Hsin-Yuan Huang

Zoë Holmes