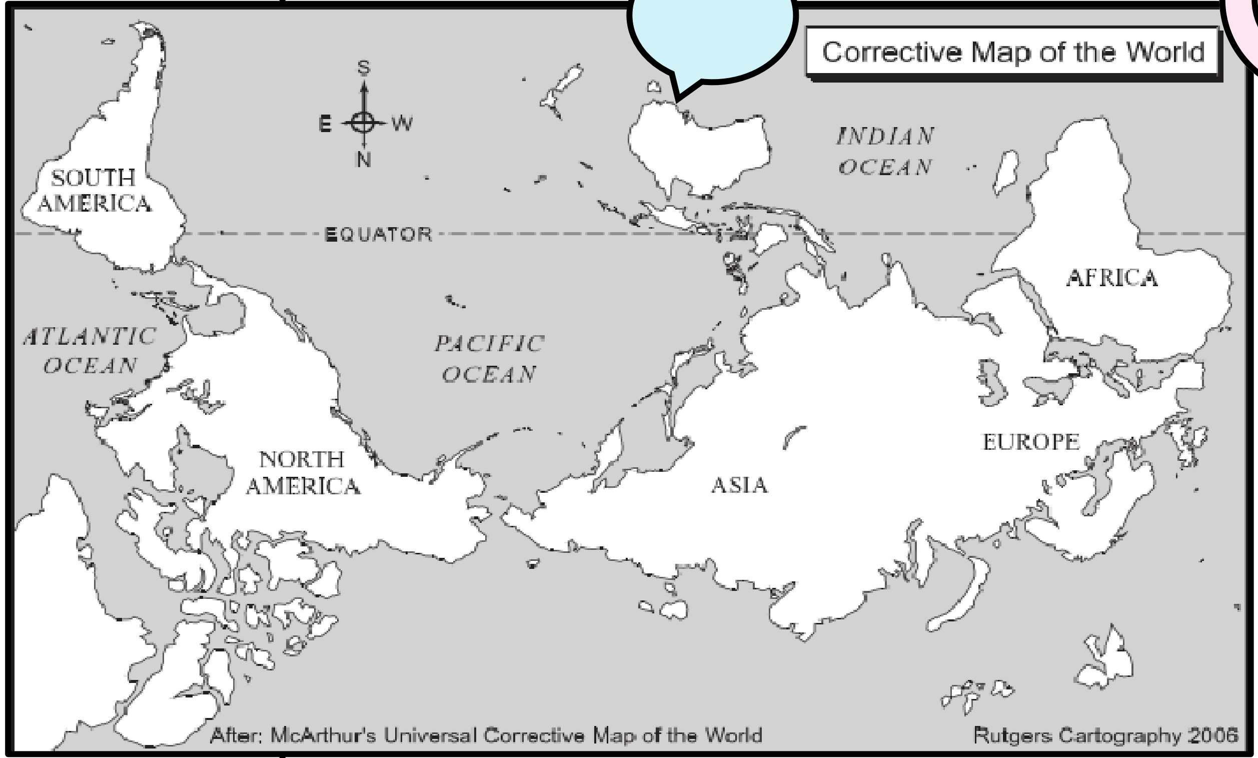
# THINK GLOBAL, ACT LOCAL





**CHRIS FERRIE**

UTS CENTRE FOR QUANTUM SOFTWARE AND INFORMATION (QSI)

# Takeaway

Global: **bad**

Local: **good**

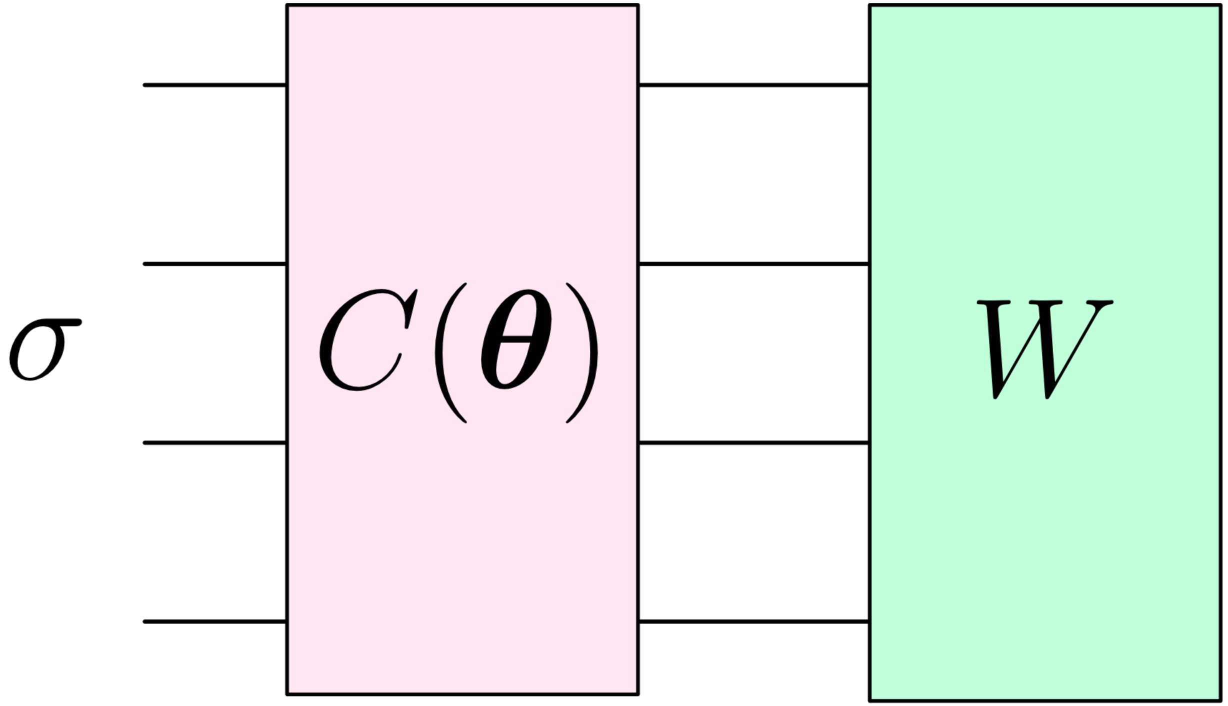
Why: **sparsity**

# On the Trainability and Classical Simulability of Learning Matrix Product States Variationally

**Afrad Basheer*[1], Yuan Feng[2], Christopher Ferrie[1], Sanjiang Li[1] and Hakop Pashayan[3]**

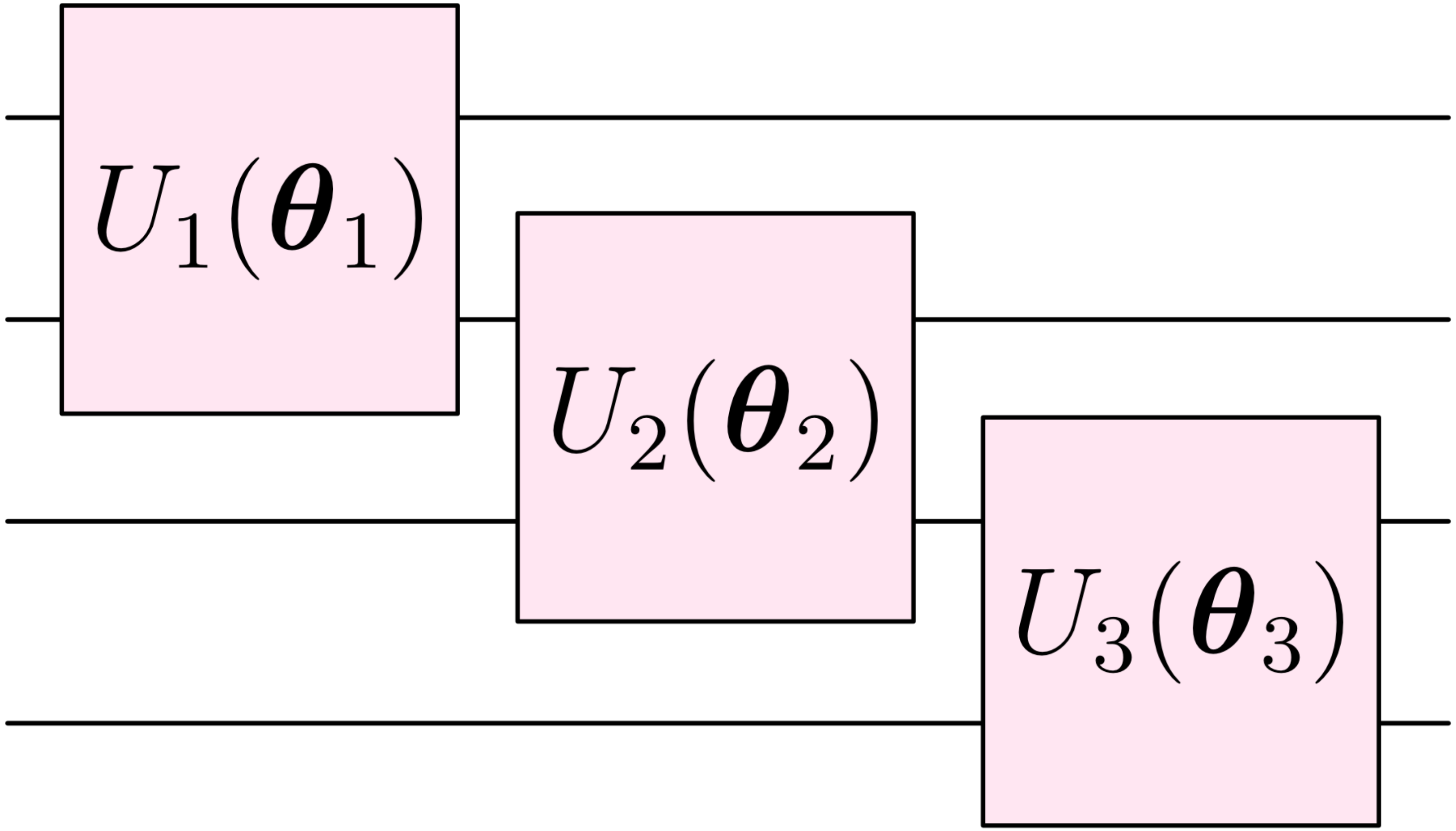
[1]Centre for Quantum Software and Information, University of Technology Sydney, NSW 2007, Australia
[2] Department of Computer Science and Technology, Tsinghua University, Beijing 100084, China
[3] Dahlem Center for Complex Quantum Systems, Freie Universität Berlin 14195, Germany

arxiv:2409.10055

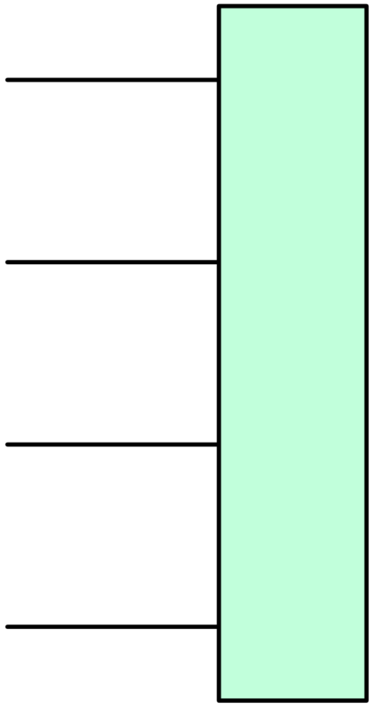


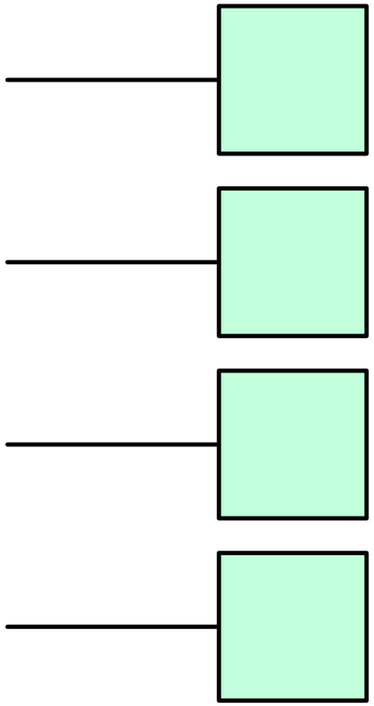

$$f_{\sigma,W,C}(\boldsymbol{\theta}) = \mathrm{tr}(WC(\boldsymbol{\theta})\sigma C(\boldsymbol{\theta})^\dagger)$$

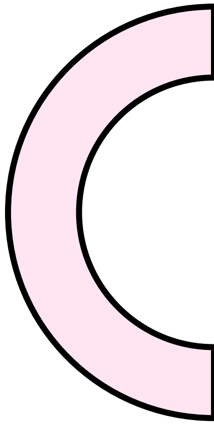
$$C_k^{(n)}(\boldsymbol{\theta}) = \prod_{p=1}^{n-k-1} \mathbb{I}^{\otimes n-k-p+1} \otimes U_p(\boldsymbol{\theta}_p) \otimes \mathbb{I}^{\otimes p-1}$$

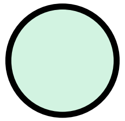



$$C_2^{(4)}(\boldsymbol{\theta})$$





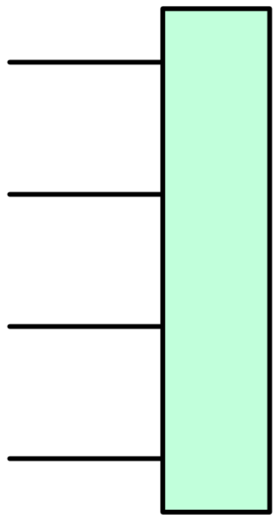




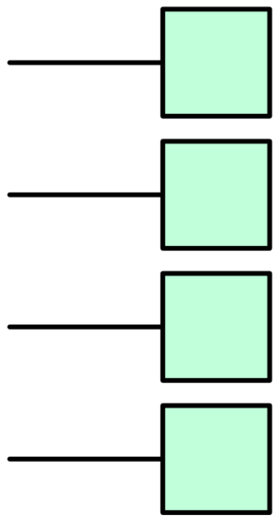
$$|\mathbf{0}\rangle\langle\mathbf{0}|$$

$$\bigotimes_{i=1}^{n} |0\rangle\langle 0|_i$$

$$f_{\sigma,|\mathbf{0}\rangle\langle\mathbf{0}|,C}(\boldsymbol{\theta})$$



$$f_{\sigma,\sum_i |0\rangle\langle 0|_i,C}(\boldsymbol{\theta})$$



**Definition 2.** *Let $\sigma \in \mathbb{D}_n$ and let $W \in \mathbb{H}_n$. For any ansatz $C(\boldsymbol{\theta}) = \prod_{p=1}^{t} U_p(\boldsymbol{\theta}_p)$, where $U_p(\boldsymbol{\theta}_p) = \prod_{q=1}^{m} e^{-i\theta_{pq}H_{pq}}$, $\boldsymbol{\theta}_p = [\theta_{p1} \dots \theta_{pm}]$, $H_{pq} \in \mathbb{H}_n$ and $\boldsymbol{\theta} = \boldsymbol{\theta}_1 \oplus \cdots \oplus \boldsymbol{\theta}_t$, and for any $p, q$, define*

$$U_p^{(L,q)}(\boldsymbol{\theta}_p) = \prod_{j=1}^{q-1} e^{-i\theta_{pj}H_{pj}}, \quad (6)$$

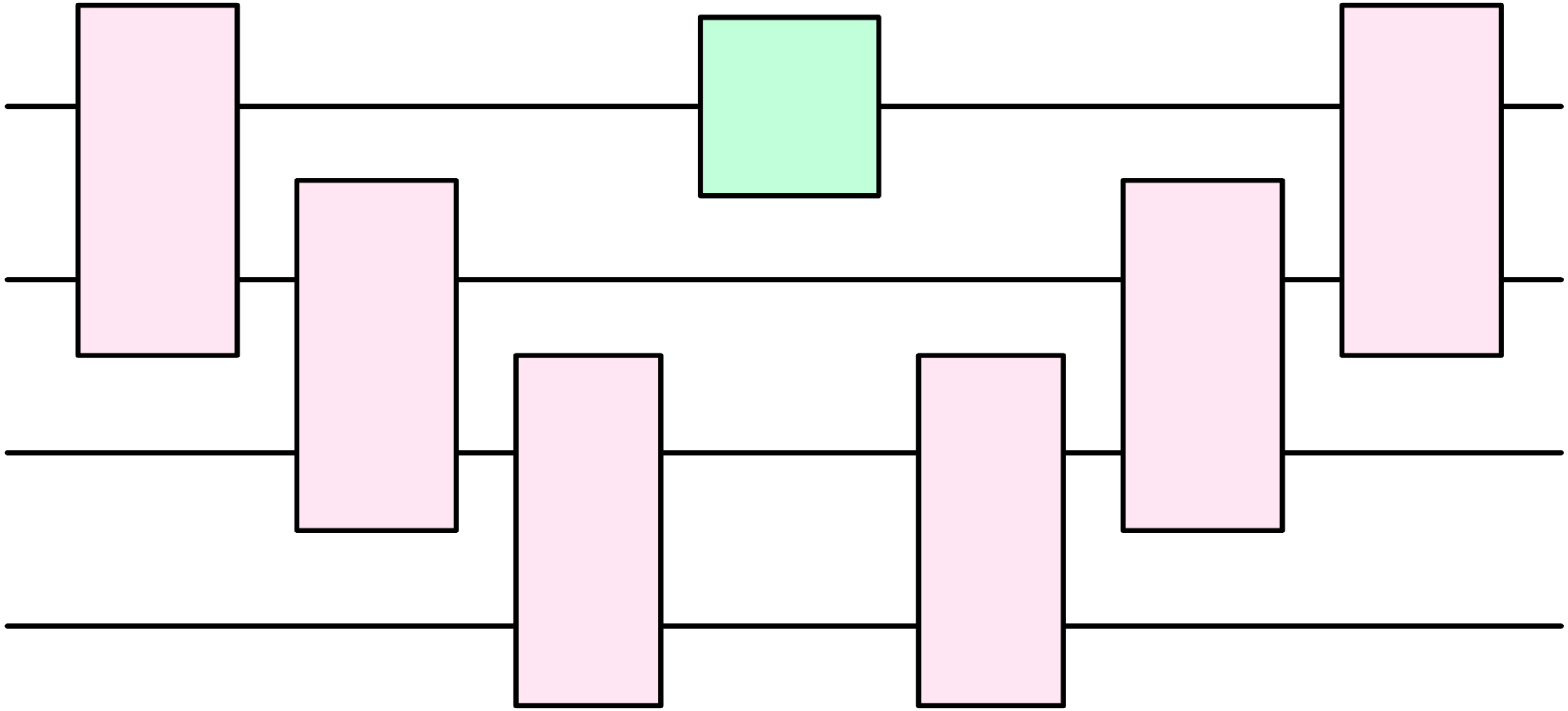
$$U_p^{(R,q)}(\boldsymbol{\theta}_p) = \prod_{j=q+1}^{m} e^{-i\theta_{pj}H_{pj}}. \quad (7)$$

*Then, $f_{\sigma,W}$ exhibits a barren plateau if $\forall\, p, q$ satisfying $1 \leq p \leq t, 1 \leq q \leq m$, we have*

$$Var_{\boldsymbol{\theta}}\left(\partial_{\theta_{pq}} f_{\sigma,W}(\boldsymbol{\theta})\right) \in \mathcal{O}\left(\frac{1}{b^n}\right), \quad (8)$$

*for some constant $b > 1$, where $\partial_{\theta_{pq}} f_{\sigma,W}(\boldsymbol{\theta})$ is it's partial derivative with respect to $\theta_{pq}$ and $U_1, \dots U_{p-1}, U_{p+1}, \dots U_t$, along with one of $U_p^{(L,q)}$ or $U_p^{(R,q)}$ are distributed according to the Haar measure and $\theta_{pq}$ is distributed uniformly.*

$$\mathrm{Var}_{\boldsymbol{\theta}}\left(f_{\sigma,W,C}(\boldsymbol{\theta})\right) \in \mathcal{O}\left(\frac{1}{b^n}\right)$$



$$C^{\dagger}(\boldsymbol{\theta})WC(\boldsymbol{\theta})$$

Sparsity

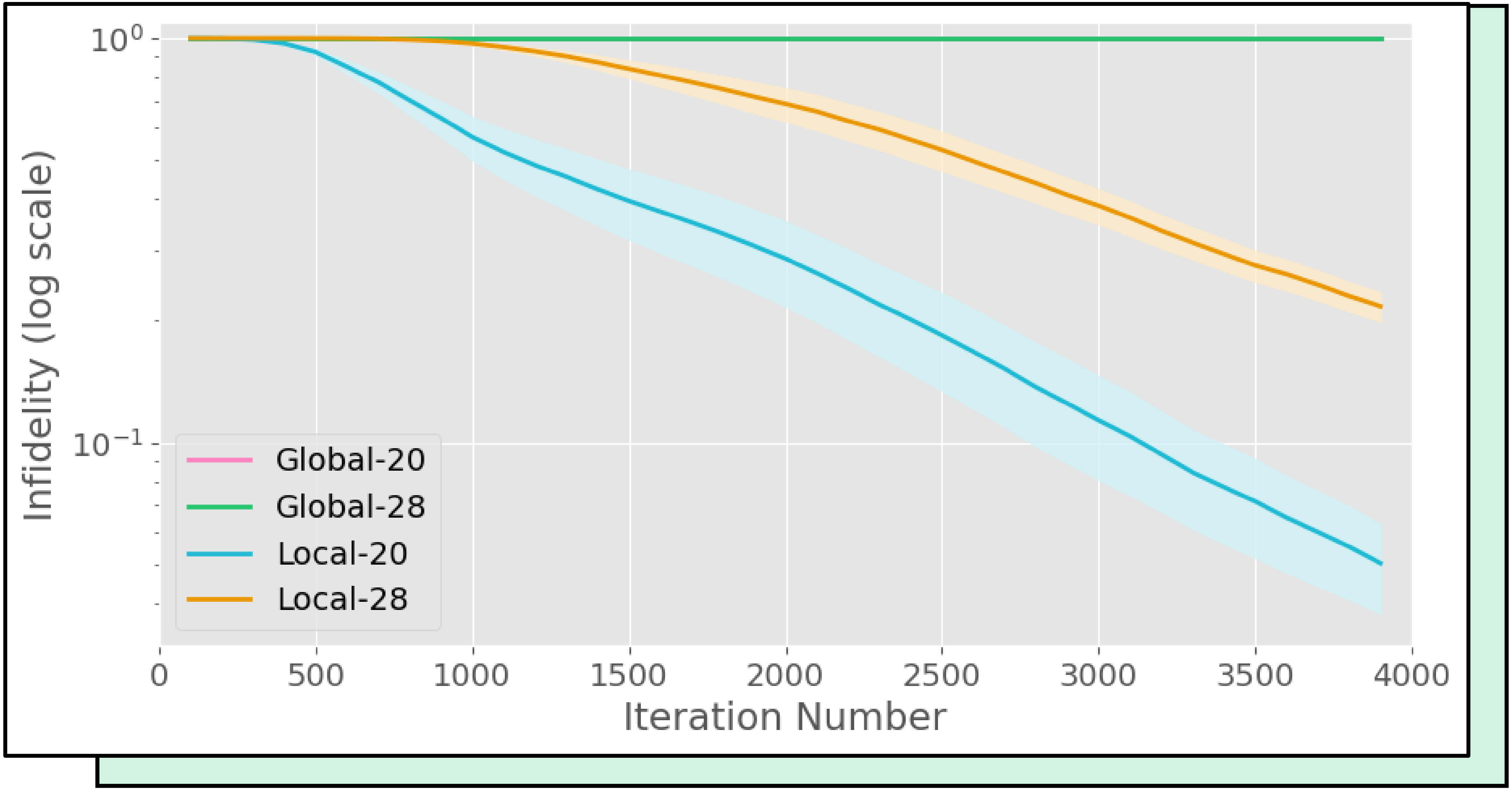
$$h_1(\sigma) = \min_{V_1,\ldots,V_n} \|\sigma_{V_1\otimes\cdots\otimes V_n}\|_1^2$$

$$h_2(\sigma) = \min_{\rho_1,\ldots,\rho_n} \|\rho_1 \otimes \cdots \otimes \rho_n - \sigma\|_{\mathrm{tr}}$$

Product-ness

$$\mathrm{Var}_{\boldsymbol{\theta}}\left(f_{\sigma,|\mathbf{0}\rangle\langle\mathbf{0}|,C_k^{(n)}}(\boldsymbol{\theta})\right) \leq \frac{h_1(\sigma)}{4^{n-k-1}}$$

$$\mathrm{Var}_{\boldsymbol{\theta}}\left(f_{\sigma,\sum_i |\mathbf{0}\rangle\langle\mathbf{0}|_i,C_k^{(n)}}(\boldsymbol{\theta})\right) \geq \frac{1}{n(2^{2k+1}+4)} - \frac{h_2(\sigma)}{2n}$$



Infidelity (log scale)
$10^0$
$10^{-1}$
0
500
1000
1500
2000
2500
3000
3500
4000
Iteration Number
Global-20
Global-28
Local-20
Local-28

**Does provable absence of barren plateaus imply classical simulability?**
**Or, why we need to rethink variational quantum computing**

M. Cerezo,[1,2,*] Martin Larocca,[3,4] Diego García-Martín,[1] N. L. Diaz,[1,5] Paolo Braccia,[3] Enrico Fontana,[6] Manuel S. Rudolph,[7] Pablo Bermejo,[8,1] Aroosa Ijaz,[3,9,10] Supanut Thanasilp,[7,11] Eric R. Anschuetz,[12,13] and Zoë Holmes[7]

arxiv:2312.09121

$$\mathcal{P}_{\boldsymbol{\theta},W}(K_j) = \frac{f_{K_j,W,C}(\boldsymbol{\theta})^2}{\|W\|_2^2}$$

$$
\begin{aligned}
f_{\sigma,W,C}(\boldsymbol{\theta}) &= \operatorname{tr}\left(WC(\boldsymbol{\theta})\sigma C(\boldsymbol{\theta})^\dagger\right) \\
&= \sum_{\boldsymbol{K}\in\mathbb{K}} \operatorname{tr}\left(\boldsymbol{K} C(\boldsymbol{\theta})^\dagger W C(\boldsymbol{\theta})\right) \operatorname{tr}(\boldsymbol{K}\sigma) \\
&= \|W\|_2^2 \sum_{K\in\mathbb{K}} \mathcal{P}_{\boldsymbol{\theta},W}(K)\,\mathbf{tr}(\boldsymbol{K\sigma})
\end{aligned}
$$

# Alternating Layered Variational Quantum Circuits Can Be Classically Optimized Efficiently Using Classical Shadows

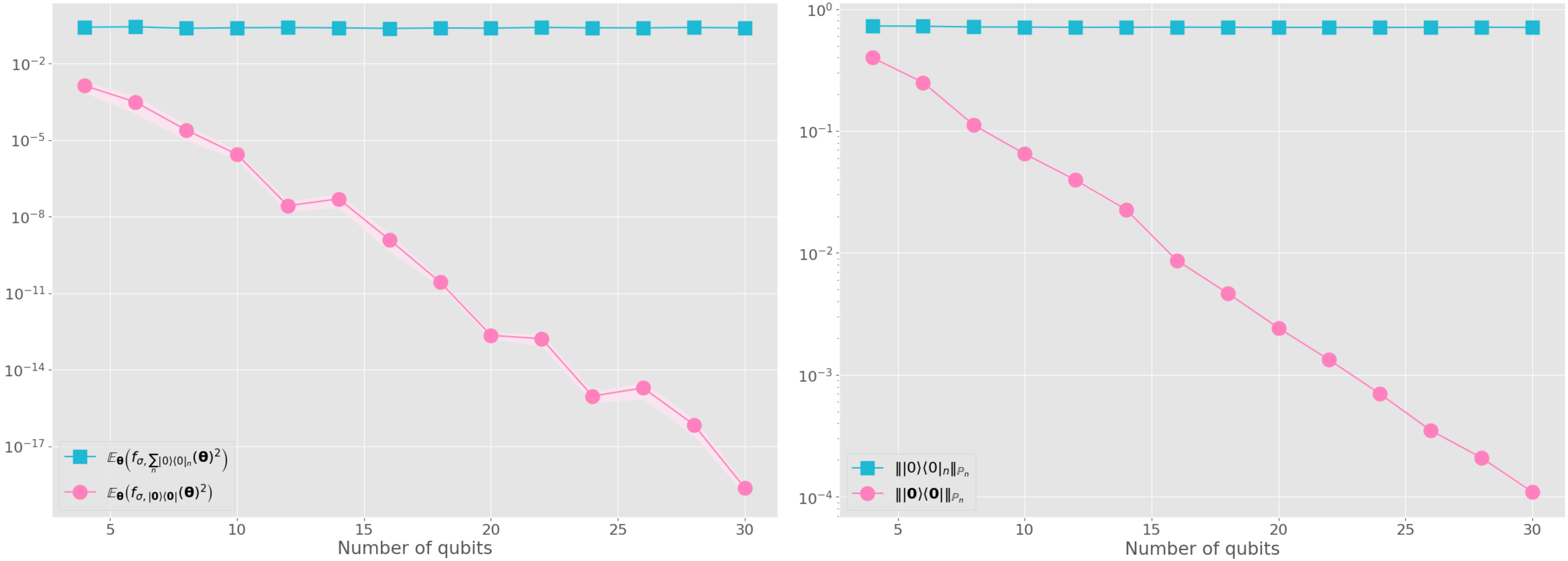
**Afrad Basheer***,[1] **Yuan Feng,** [1] **Christopher Ferrie,** [1] **Sanjiang Li** [1]

[1] Centre for Quantum Software and Information, University of Technology Sydney, NSW 2007, Australia

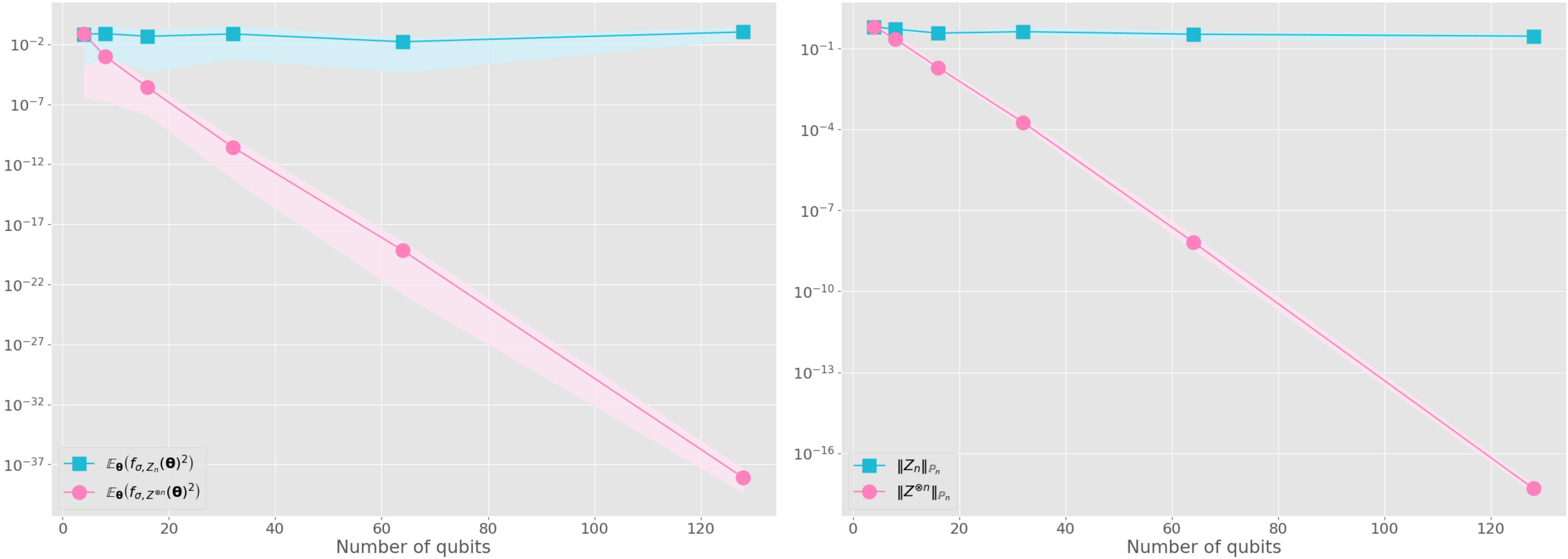
arxiv:2208.11623

$$\mathcal{P}_{\boldsymbol{\theta},W}(K_j) = \frac{f_{K_j,W,C}(\boldsymbol{\theta})^2}{\|W\|_2^2}$$

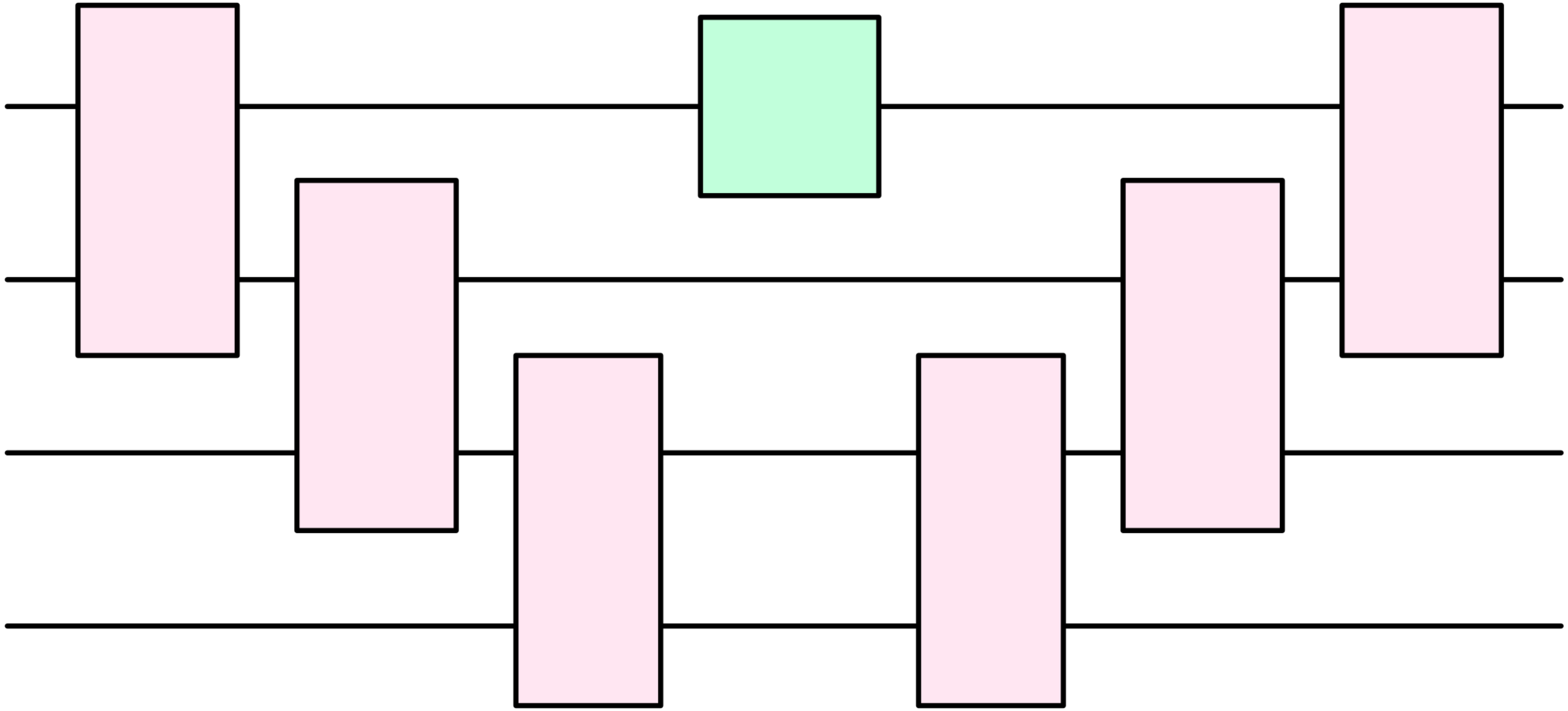
$$\|W\|_{\mathbb{K}} = \frac{1}{\|W\|_2} \left[ \sum_{K \in \mathbb{K}} \operatorname{tr}(KW)^4 \right]^{1/4}$$



$10^{-2}$
$10^{-5}$
$10^{-8}$
$10^{-11}$
$10^{-14}$
$10^{-17}$
$\mathbb{E}_{\boldsymbol{\theta}}\left(f_{\sigma,\sum_n |0\rangle\langle 0|_n}(\boldsymbol{\theta})^2\right)$
$\mathbb{E}_{\boldsymbol{\theta}}\left(f_{\sigma,|\mathbf{0}\rangle\langle\mathbf{0}|}(\boldsymbol{\theta})^2\right)$
5
10
15
20
25
30
Number of qubits
$10^{0}$
$10^{-1}$
$10^{-2}$
$10^{-3}$
$10^{-4}$
$\||0\rangle\langle 0|_n\|_{\mathbb{P}_n}$
$\||\mathbf{0}\rangle\langle\mathbf{0}|\|_{\mathbb{P}_n}$
5
10
15
20
25
30
Number of qubits



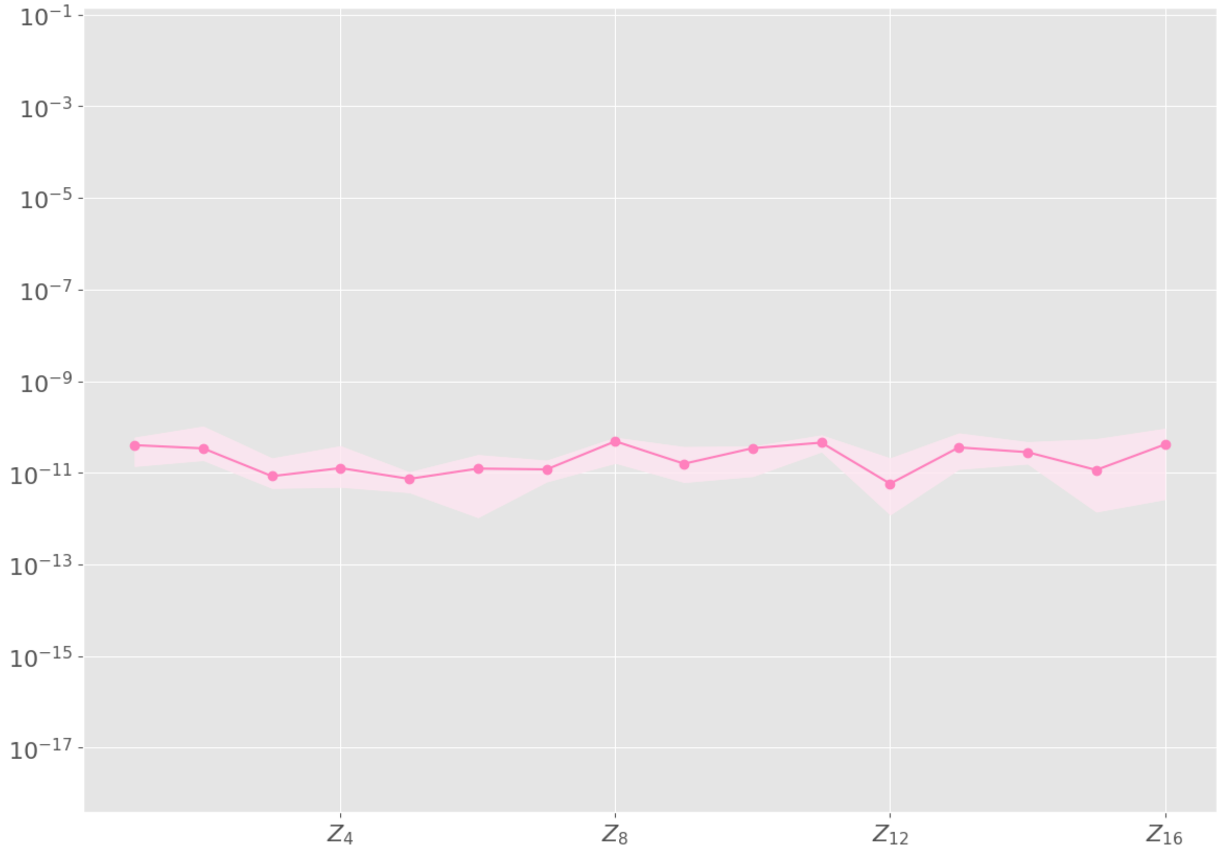
$\mathbb{E}_{\boldsymbol{\theta}}(f_{\sigma,Z_n}(\boldsymbol{\Theta})^2)$
$\mathbb{E}_{\boldsymbol{\theta}}(f_{\sigma,Z^{\otimes n}}(\boldsymbol{\Theta})^2)$
$10^{-2}$
$10^{-7}$
$10^{-12}$
$10^{-17}$
$10^{-22}$
$10^{-27}$
$10^{-32}$
$10^{-37}$
0
20
40
60
80
100
120
Number of qubits
$\|Z_n\|_{\mathbb{P}_n}$
$\|Z^{\otimes n}\|_{\mathbb{P}_n}$
$10^{-1}$
$10^{-4}$
$10^{-7}$
$10^{-10}$
$10^{-13}$
$10^{-16}$
0
20
40
60
80
100
120
Number of qubits

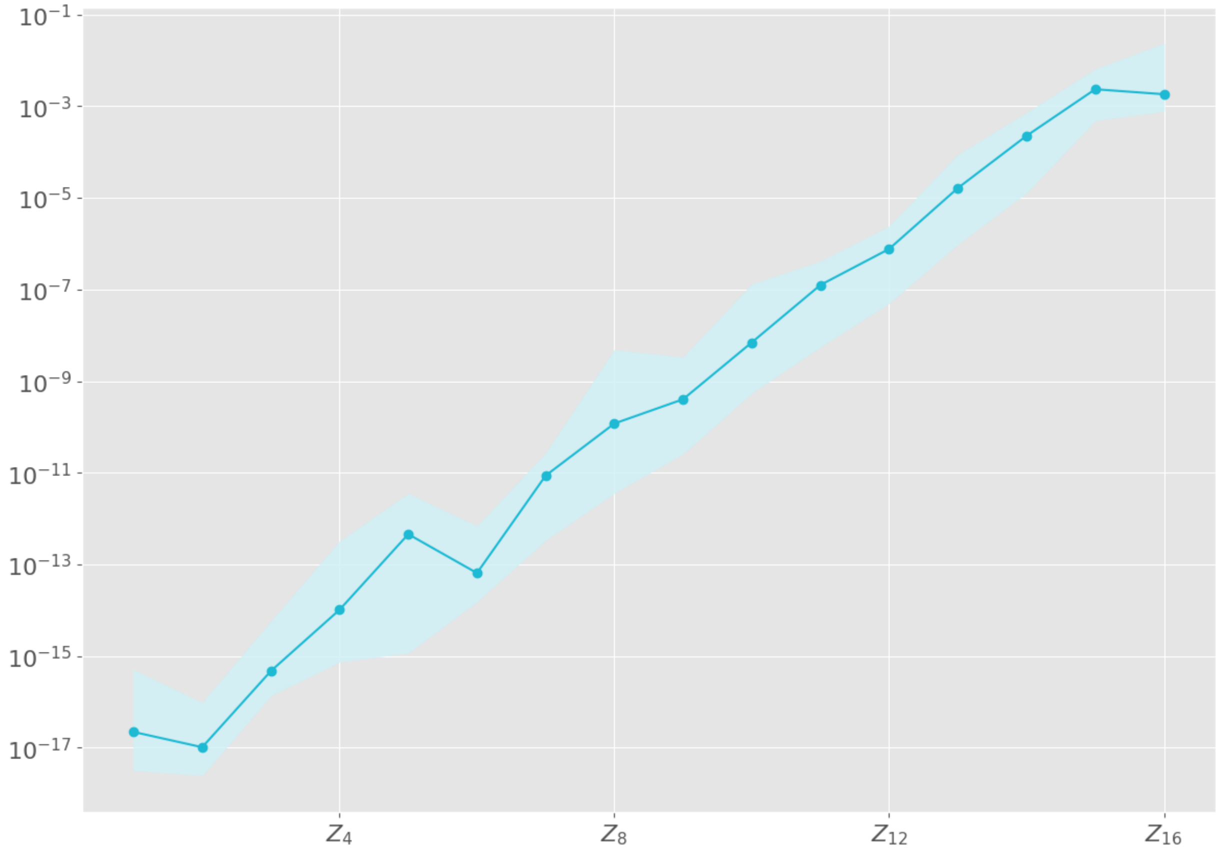


$$C^{\dagger}(\boldsymbol{\theta})WC(\boldsymbol{\theta})$$

$$|\mathbf{0}\rangle\langle\mathbf{0}|$$

$$\bigotimes_{i=1}^{n} |0\rangle\langle 0|_i$$



# On the Trainability and Classical Simulability of Learning Matrix Product States Variationally

**Afrad Basheer*[1], Yuan Feng[2], Christopher Ferrie[1], Sanjiang Li[1] and Hakop Pashayan[3]**

[1]Centre for Quantum Software and Information, University of Technology Sydney, NSW 2007, Australia
[2] Department of Computer Science and Technology, Tsinghua University, Beijing 100084, China
[3] Dahlem Center for Complex Quantum Systems, Freie Universität Berlin 14195, Germany

arxiv:2409.10055

# Local quantum surrogate models

**Sreeraj Nair**

- Today!
- Poster 51

# Riemannian-geometric generalizations of fidelity

**Afham**

- Today!
- Poster 56