



Ronald Cortes
(ICTP)



Aravindh Shankar
(ICTP)



Marcello Dalmonte
(ICTP $\rightarrow$ U. of Bologna)



Roberto Verdel
(ICTP $\rightarrow$ Eniquantic)

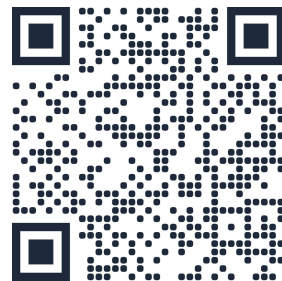
Basis dependence of neural quantum states

Speaker: Nils Niggemann (ICTP Trieste)

Interplay between neural and tensor networks, Grenoble, July 2026



R. S. C. Cortes, A. S. Shankar, M. Dalmonte, R. Verdel, **N. N.** SciPost Phys. Core 9, 027 (2026)



Introduction: Many body wavefunctions

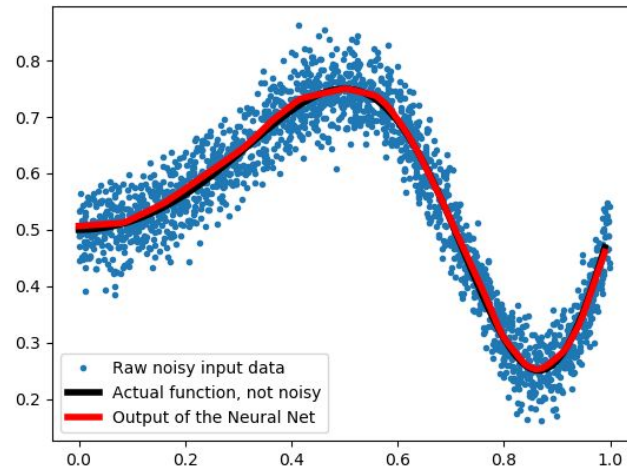
Ground states of many-body systems $H = \sum_{i\alpha} h_i^\alpha S_i^\alpha + \sum_{j<i, \alpha\beta} J_{ij}^{\alpha\beta} S_i^\alpha S_i^\beta + \dots$

Time evolution $|\psi(t)\rangle = e^{-iHt}|\psi(0)\rangle$

Exponential growth of Hilbert space,
entanglement

→ Highly complicated problem...

But neural networks can compress data very
efficiently



<https://robertfilter.net/blog/machine-learning/how-to-implement-a-simple-neural-network-as-interpolation-function-in-python-using-keras.html>

Introduction: Neural Quantum states

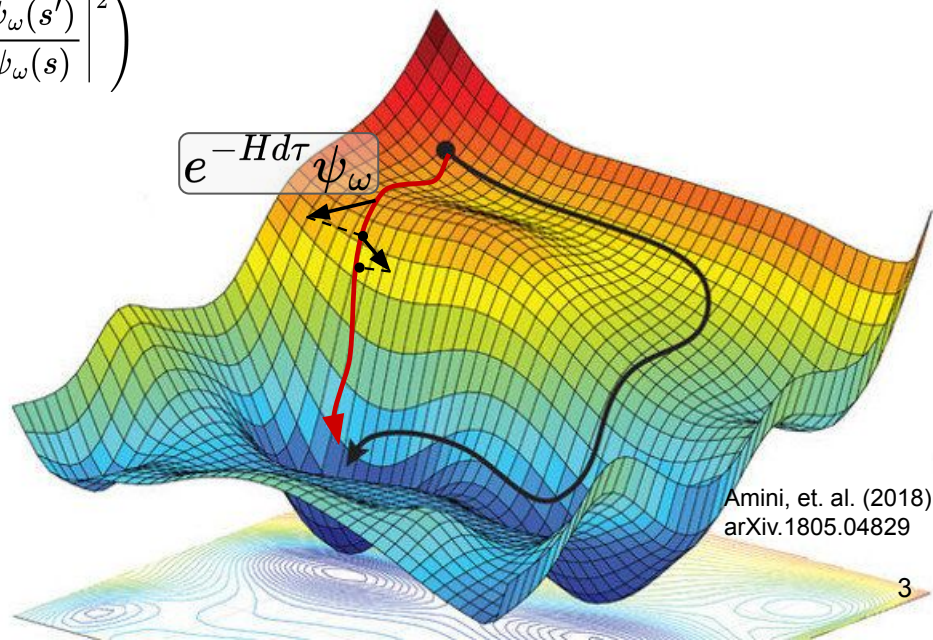
1. Function that evaluates $\psi_\omega(s) = \langle s | \psi_\omega \rangle$ for any configuration $s = (\uparrow, \downarrow, \dots, \uparrow)$
2. Start with some initial parameters $\omega = \omega_{\text{init}}$
3. Generate samples from the distribution given by $|\psi_\omega(s)|^2$

Metropolis rule: $A(s \rightarrow s') = \min \left(1, \left| \frac{\psi_\omega(s')}{\psi_\omega(s)} \right|^2 \right)$

4. Estimate variational energy

$$E_\omega = \langle \psi_\omega | H | \psi_\omega \rangle \approx \frac{1}{N_{\text{samples}}} \sum_{s \in P(|\psi_\omega|^2)} \sum_{s'} H_{ss'} \frac{\psi_\omega(s')}{\psi_\omega(s)}$$

5. Update ω to minimize E_ω
 - Stochastic gradient descent
 - **Stochastic Reconfiguration**

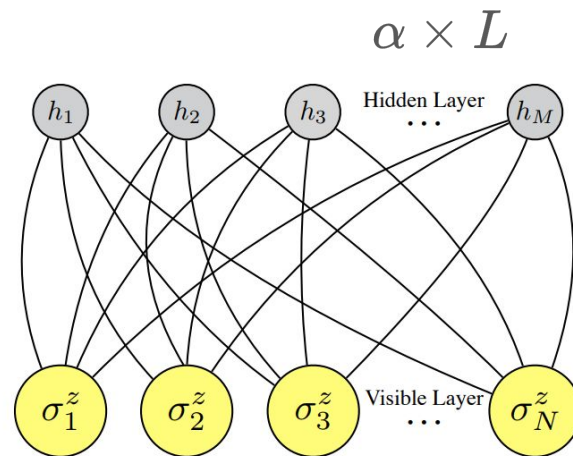


Amini, et. al. (2018).
arXiv.1805.04829

Restricted Boltzmann Machine

$$\begin{aligned}\psi_{\text{RBM}}(s) &= \sum_{\{h\}} e^{\sum_i a_i s_i + \sum_j (b_j h_j + \sum_i W_{ij} s_i h_j)} \\ &= e^{\sum_i a_i s_i} \prod_{j=1}^M 2 \cosh\left(\sum_i W_{ij} s_i + b_j\right)\end{aligned}$$

- Total parameters = $\alpha L^2 + (\alpha + 1)L$
- Universal approximator for large enough α
- Correlations intrinsically non-local
- Mappings to matrix product states, stabilizer states ...



Carleo, Troyer *Science* **355**,602-606(2017)

Introduction: The spectrum of Many-Body approaches

Memory required



$ \psi\rangle$	Computational basis $\sum_{\{s\}} \psi_{s_1, \dots, s_L} s_1, \dots, s_L\rangle$	Matrix Product States $\sum_{\{s\}} A_{s_1}^{(1)} \dots A_{s_L}^{(L)} s_1, \dots, s_L\rangle$	Neural Quantum States $\sum_{\{s\}} \psi_{\omega}(s_1, \dots, s_L) s_1, \dots, s_L\rangle$	Quantum Monte Carlo $\sum_{\{s\}} P(s_1, \dots, s_L) s_1, \dots, s_L\rangle$
basis dependence	- 2^L amplitudes	- $L \times D^2$ matrix elements	- N_{params} weights	- $\sim L$ configurations
	- exp with size L	- exp with entanglement	- ? polynomial(*) ?	- power-law in L (*)
	- Equivalent in (nearly) any basis	- Local basis choice is arbitrary (*)	- accuracy basis dependent (**)	- impossible in most bases!
		(*) DMRG speed may be affected: Schollwöck, Phys. Rev. B 58, 81941998 1998	(*) Rende et. al. arXiv:2606.02794v1 (2026) (**) Moreno et. al. 2023, arXiv:2302.11588v2	(*) Assuming ergodicity!

Basis dependence

Lessons learned from Tensor networks:

- DMRG's Breakthrough: Understanding MPS & Entanglement
 - Understanding **limitations** drives **improvement**!
- Important difference: Optimization via DMRG is extremely reliable
 - MPS: Much focus on **representative** power
 - NQS: Limitations driven substantially by **optimization** procedure

When do neural quantum states (not) work?

Universal function approximation theorems:

- Error scaling may be power-law or exponential, depending on intrinsic details of the function!

Empirical observations: NQS become inaccurate if

- there is a severe **sign problem** (ground states)
 - large **evolution times** (time evolution)
- Wang, Wu³, He, Zhong-Yi Lu 2024
 - Szabó, Castelnovo 2020
 - Lin, Pollmann, 2021
 - ...

Takeaway from Tensor networks: **Entanglement** is a measure of **complexity**

→ Immensely **useful!**

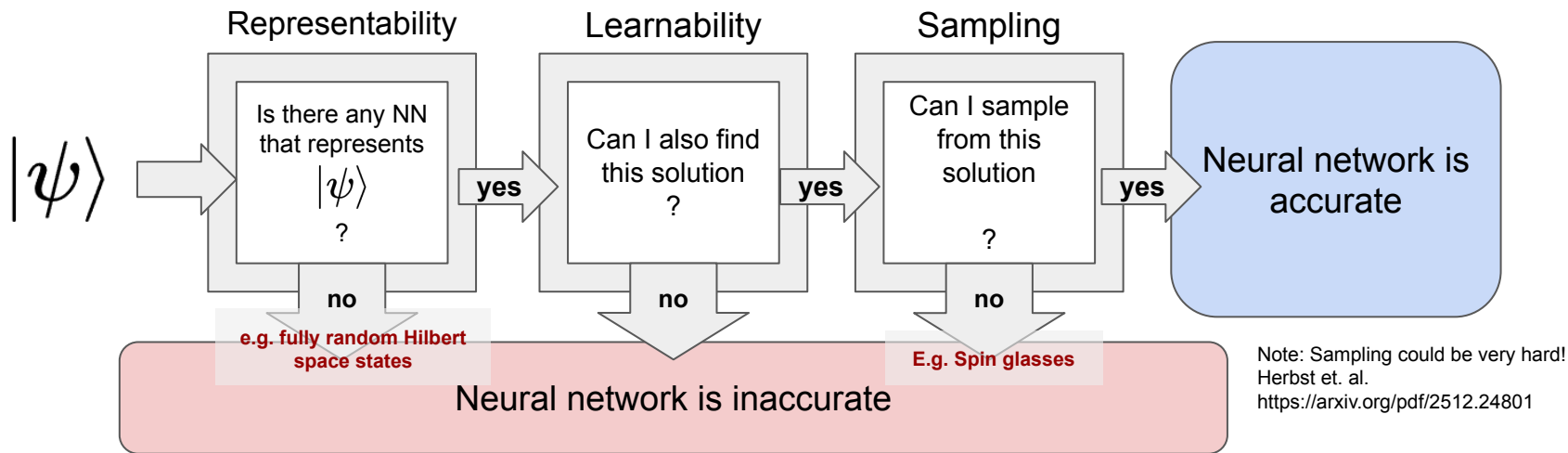
(very ambitious) Goal: Can we define a physical quantity that measures complexity?

Can we define a physical quantity that measures complexity?

- Example MPS: High Entanglement = High Complexity
- Natural question: Does entanglement also limit NQS?

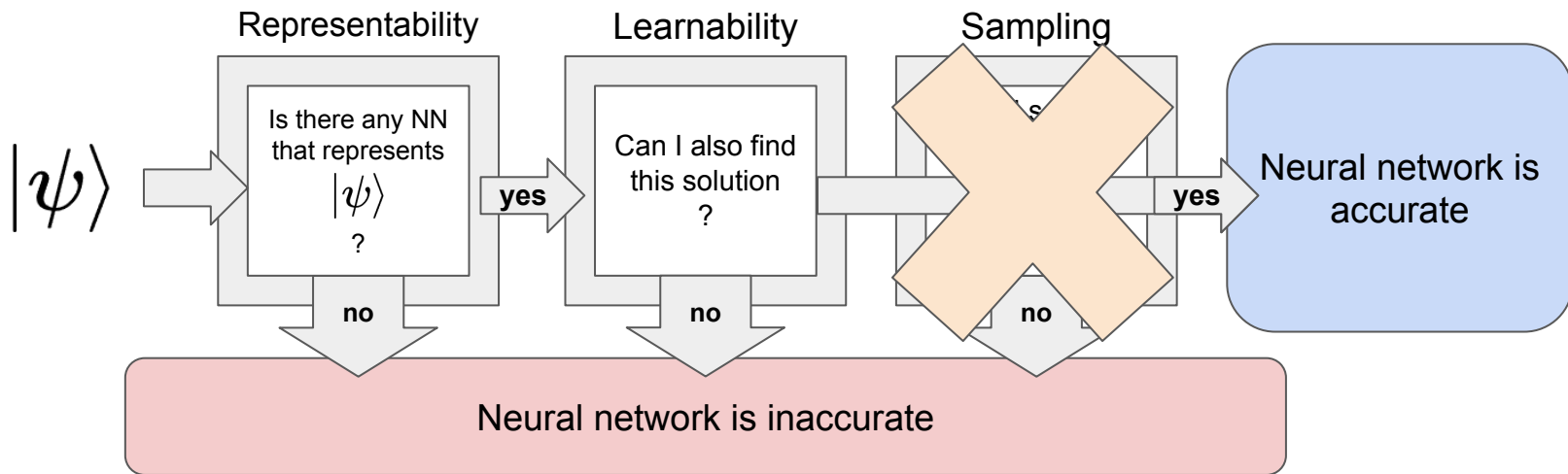
In general: **NO!** Denis, Sinibaldi, Carleo, Phys. Rev. Lett. 134, 079701 (2025)

We want: For any given **accuracy**: Polynomial **time** & polynomial **memory**



Simplification: exact summation

$$E_\omega = \langle \psi_\omega | H | \psi_\omega \rangle = \sum_{ss'} H_{ss'} |\psi_\omega(s)|^2 \frac{\psi_\omega(s')}{\psi_\omega(s)}$$



no Monte Carlo sampling
focus only on representability / learning

Noteworthy sampling-related work:

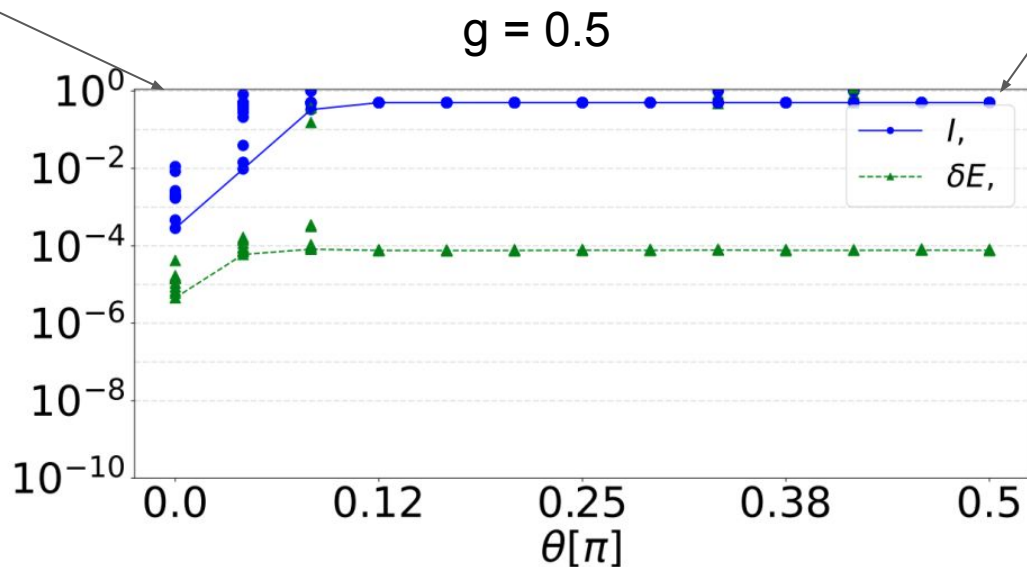
Chen et. al. arXiv:2606.23657v1 (2026)

- Experiment: Modify H without affecting the physics!

transverse field Ising model
$$H(\theta) = - \sum_{i=1}^{L-1} \tilde{\sigma}_i^z(\theta) \tilde{\sigma}_{i+1}^z(\theta) - g \sum_{i=1}^L \tilde{\sigma}_i^x(\theta)$$

$$H(0) = - \sum_{i=1}^{L-1} Z_i Z_{i+1} - g \sum_{i=1}^L X_i$$

$$H(\pi/2) = - \sum_{i=1}^{L-1} X_i X_{i+1} - g \sum_{i=1}^L Z_i$$

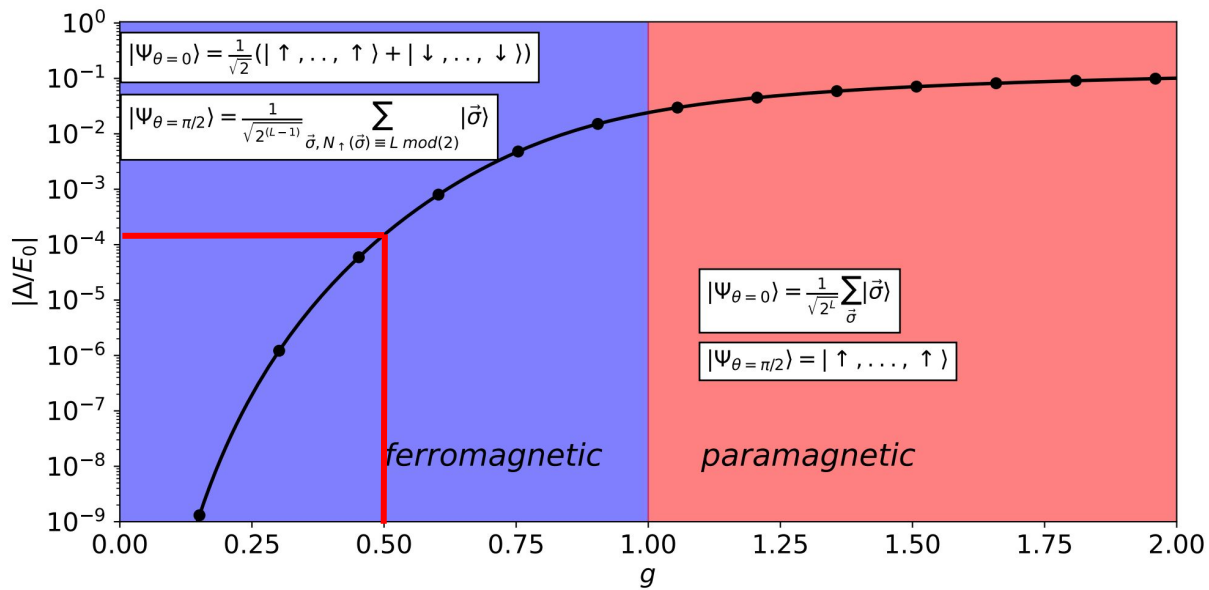


$$\delta E = \left| \frac{E_{\text{RBM}} - E_0}{E_0} \right|$$

$$\mathcal{I} = 1 - |\langle \Psi_{\text{RBM}} | \Psi_0 \rangle|^2$$

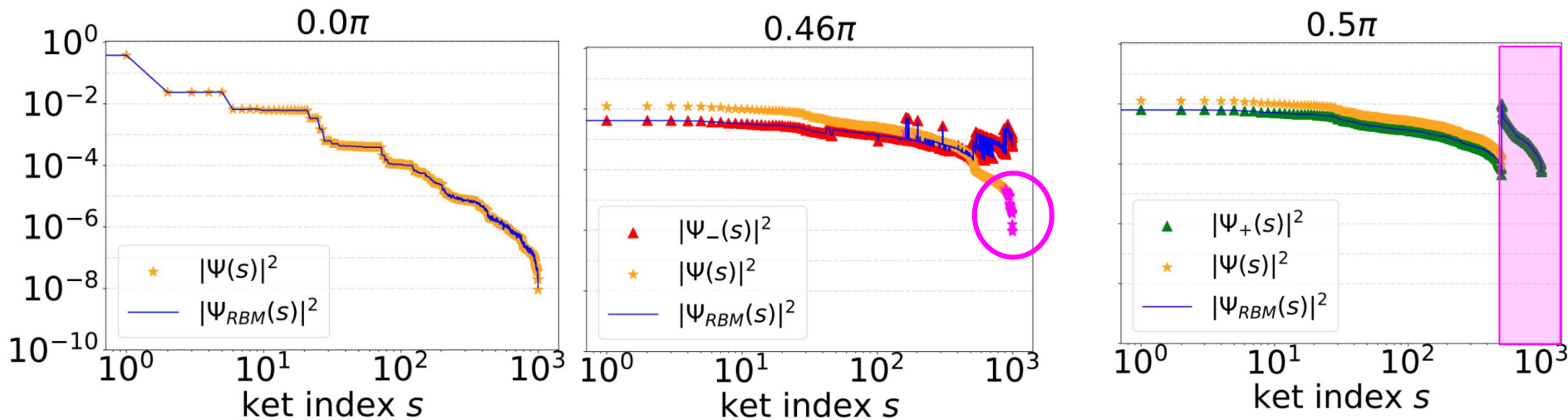
Bad infidelity despite good energy?

To leading order, energy and infidelity are proportional $\Delta E \approx \Delta \times \mathcal{I}$



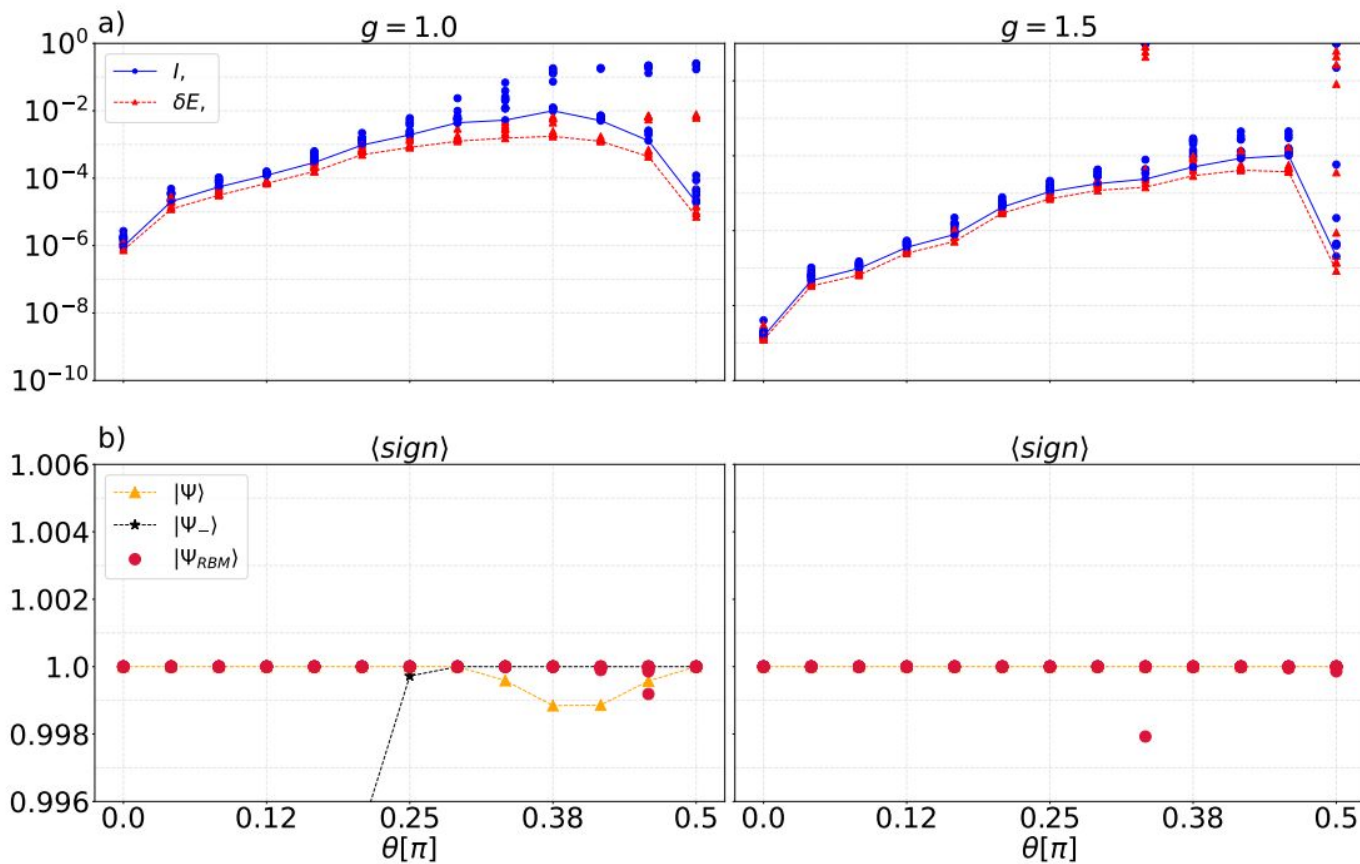
What is the simplest state to learn?

Neural networks = lazy students (“How to pass the test with minimum effort?”)

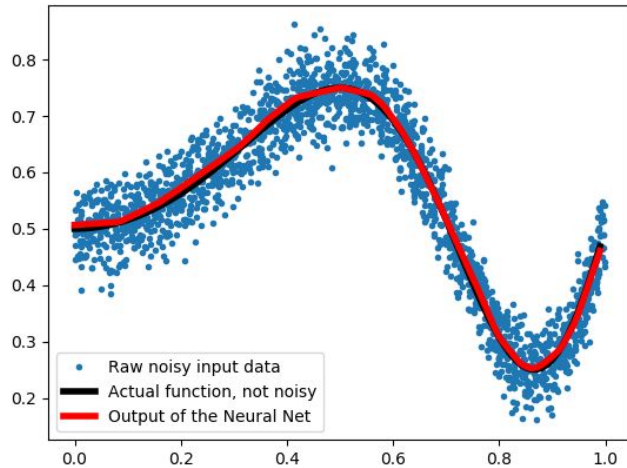


amplitudes	- smoothly varying	- smoothly varying	- parity constraint
signs	- uniform	- non-uniform	- uniform

Observation: Basis still matters if there is a finite (size) gap!

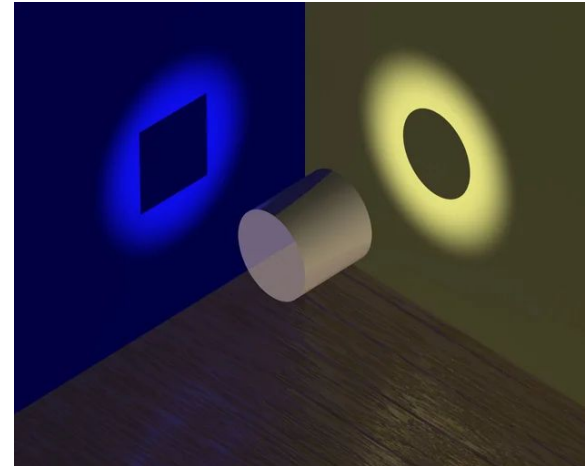


In which basis is learning harder?



<https://robertfilter.net/blog/machine-learning/how-to-implement-a-simple-neural-network-as-interpolation-function-in-python-using-keras.html>

- Neural networks try to **fit a function:**
 $f(s) = \log(\psi(s))$
- Shape of $f(s)$ depends on the basis



- C.f. : “spikiness” (Paul Nisarga’s talk yesterday)

When is the shape of a function complicated?

- Cumulant expansion: characterize importance of many-body correlations

$$\begin{aligned}\Phi(s) &= \exp \left[\log(\mathcal{N}) + \sum_i c_i s_i + \sum_{i < j} c_{ij} s_i s_j + \sum_{i < j < k} c_{ijk} s_i s_j s_k + \cdots \right] \\ &\equiv \exp \left[\sum_{A \subseteq \mathcal{P}(\{1, \dots, L\})} c_A(\Phi) S_A(s) \right],\end{aligned}$$

Reminder: NQS are actually fitting $\log(\psi)$!

- Bijective: 2^L terms
- compare: binary Fourier (Hadamard-Walsh) transform:

$$c_A = (H^{\otimes L} \log \Phi)_A$$

Borin, Abanin, Physical Review B 101(19), 195141 (2020)
Glasse et. al. Physical Review X 8(1), 011006 (2018)
Valenti et. al., Physical Review Research, 4(1), (2022)
Döschl, Bohrdt, arxiv: 2508.14152 (2025)
Schurov et. al., arXiv:2508.09870 (2025)

Truncated cumulant expansion

- ▶ Compute *all* 2^L cluster coefficients $c_A(\Psi)$ of the **exact** ground state $\Psi(s)$. $\mathcal{O}(L 2^L)$ complexity through FWHT of $\log \Psi$
- ▶ Find the set $\{A\}_{\text{largest}}$ corresponding to the indices of the N_{trunc} **by magnitude largest** coefficients $|c_A(\Psi)|$.
- ▶ Define reconstructed state **$\Phi^{(N_{\text{trunc}})}(s)$** $\equiv \exp\left(\sum_{A \in \{A\}_{\text{largest}}} c_A(\Phi) S_A(s)\right)$

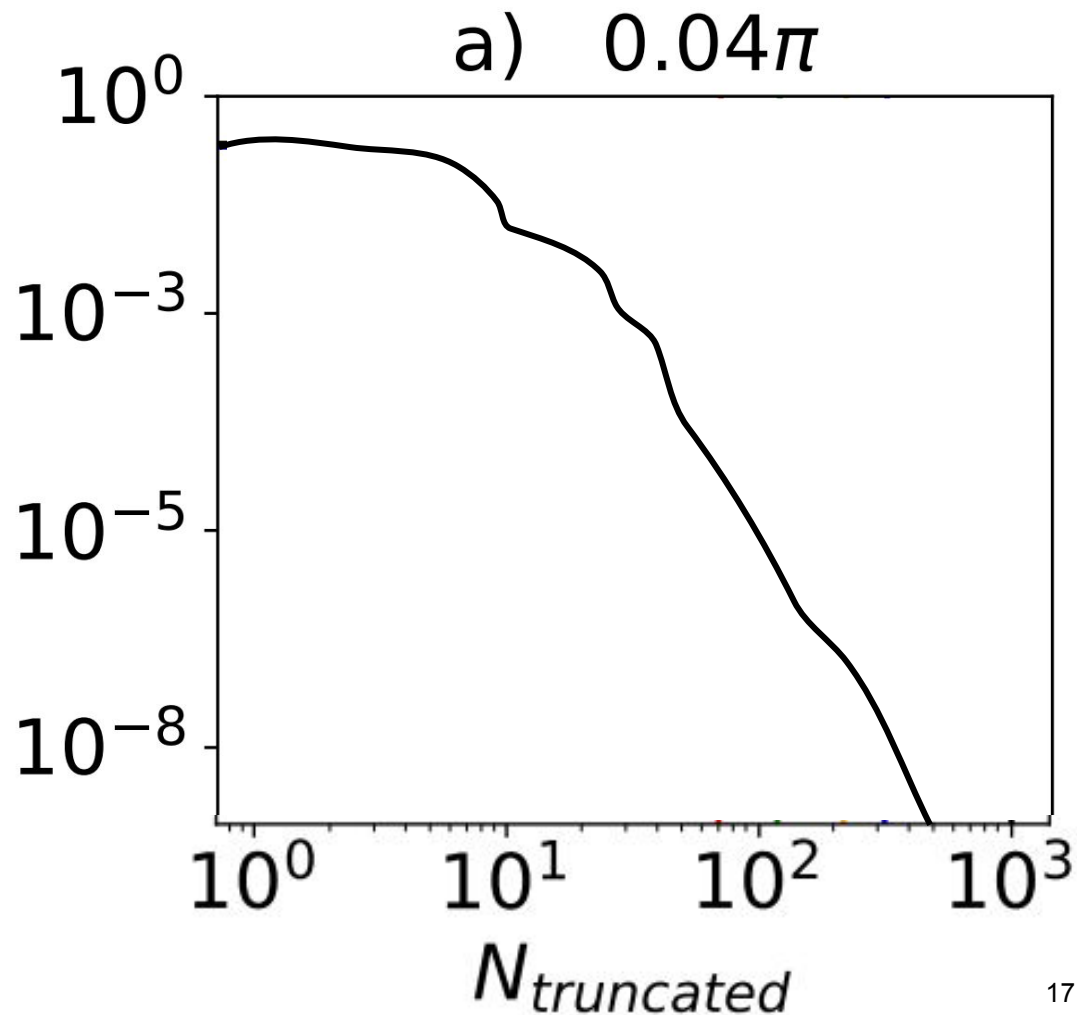
example: $\Phi^{(3)}(s) = \exp(c_{138}s_1s_3s_8 + c_1s_1 + c_{16789}s_1s_6s_7s_8s_9)$

Relation to NQS

$$I(\psi_0^{(N_{\text{trunc}})}, \psi_0)$$

Infidelity of the truncation of the exact state

$$g = 1.5$$



Relation to NQS

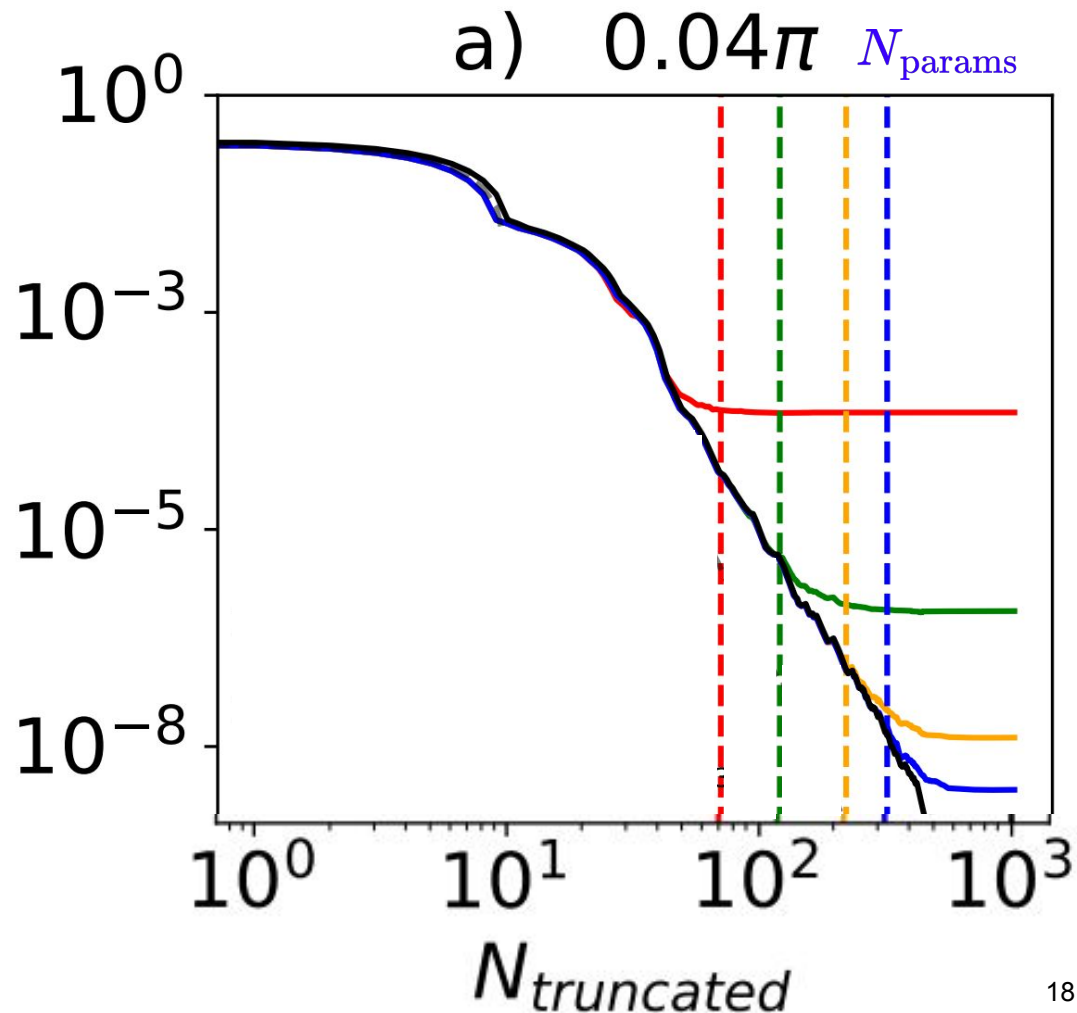
$$I(\psi_0^{(N_{\text{trunc}})}, \psi_0)$$

Infidelity of the truncation of the exact state

$$g = 1.5$$

Infidelity of the truncation of the RBM

$$\begin{aligned} &I(\psi_{\alpha=0.5}^{(N_{\text{trunc}})}, \psi_0) \quad I(\psi_{\alpha=2}^{(N_{\text{trunc}})}, \psi_0) \\ &I(\psi_{\alpha=1}^{(N_{\text{trunc}})}, \psi_0) \quad I(\psi_{\alpha=3}^{(N_{\text{trunc}})}, \psi_0) \end{aligned}$$



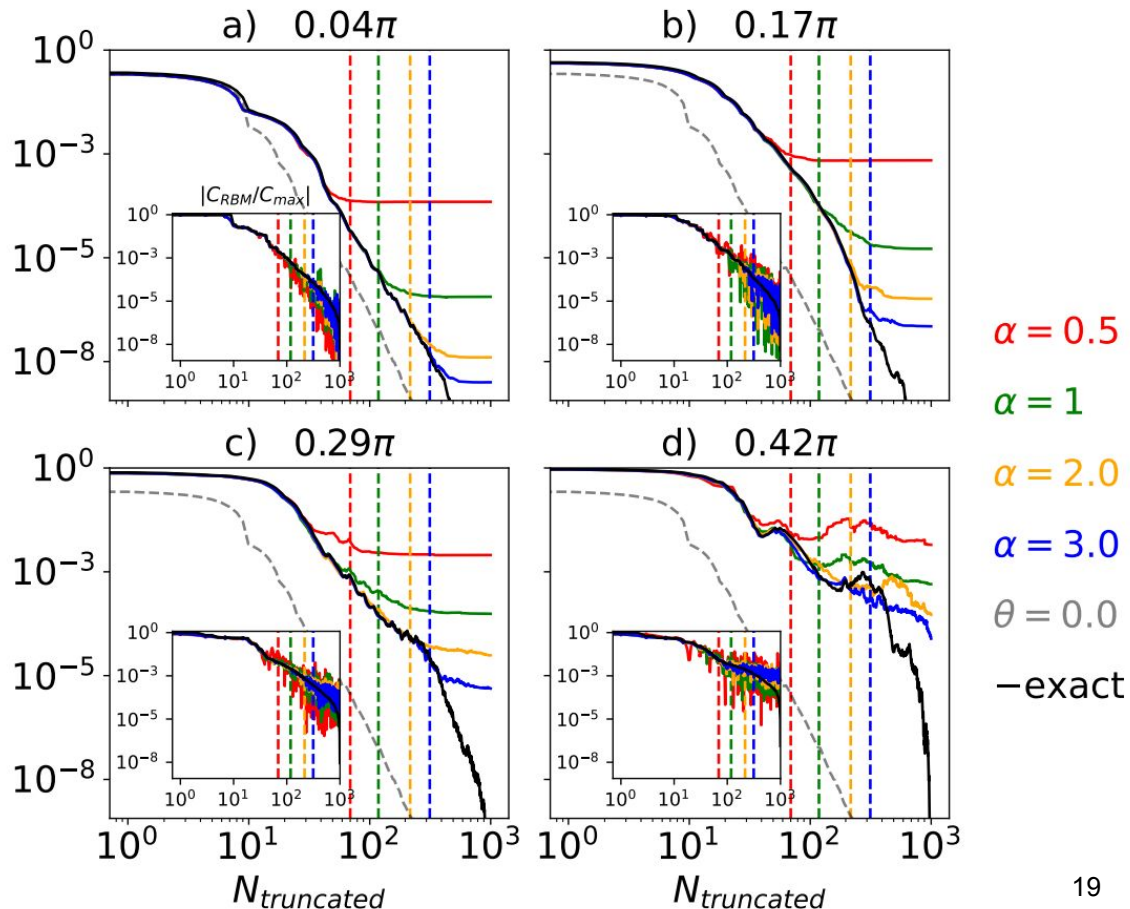
$$g = 1.5$$

Relation to NQS

Interpretation

- RBM's perform (about) as well as a truncated cumulant expansion **with the same number of parameters**
- If expansion **converges slowly**, many coefficients are needed $\rightarrow$ hard to approximate
- Cumulant expansion is **basis dependent!**

Infidelity (I)



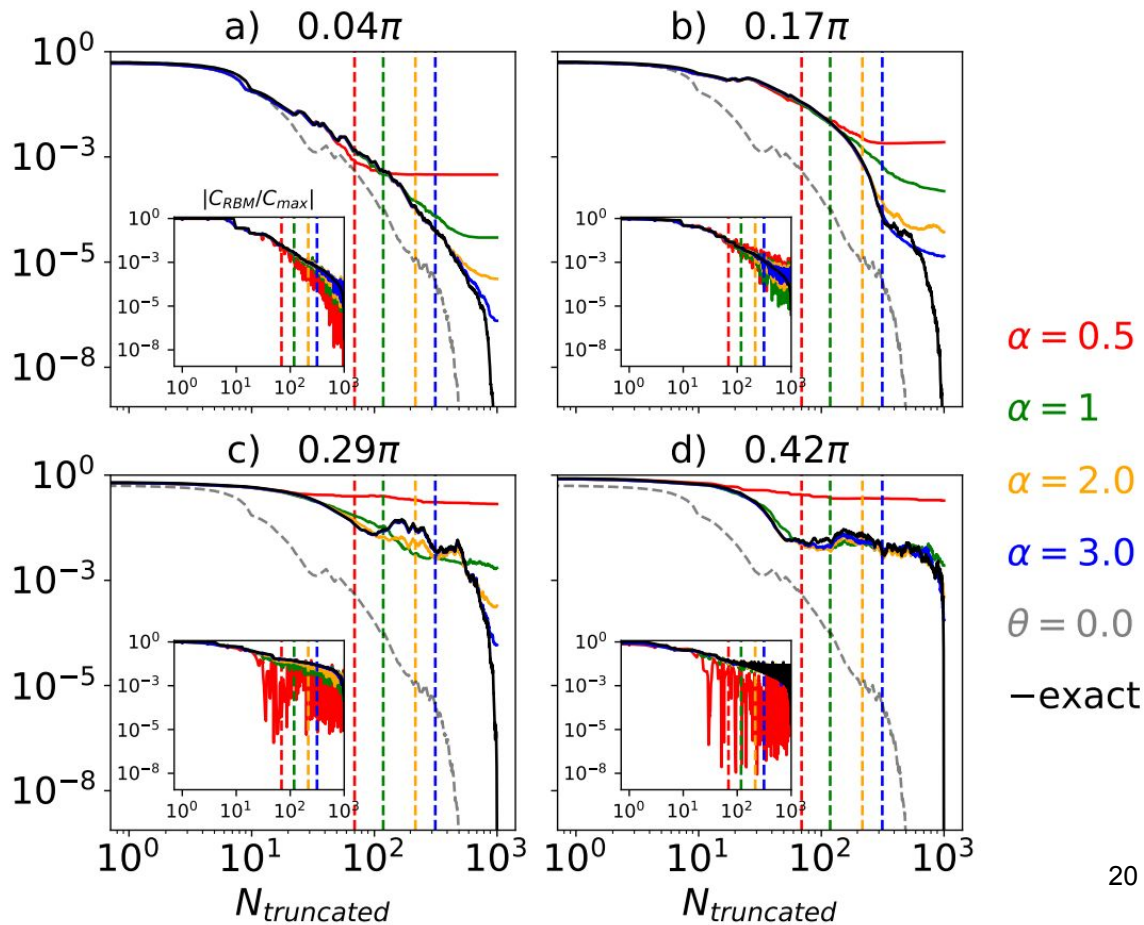
$$g = 1.0$$

Relation to NQS

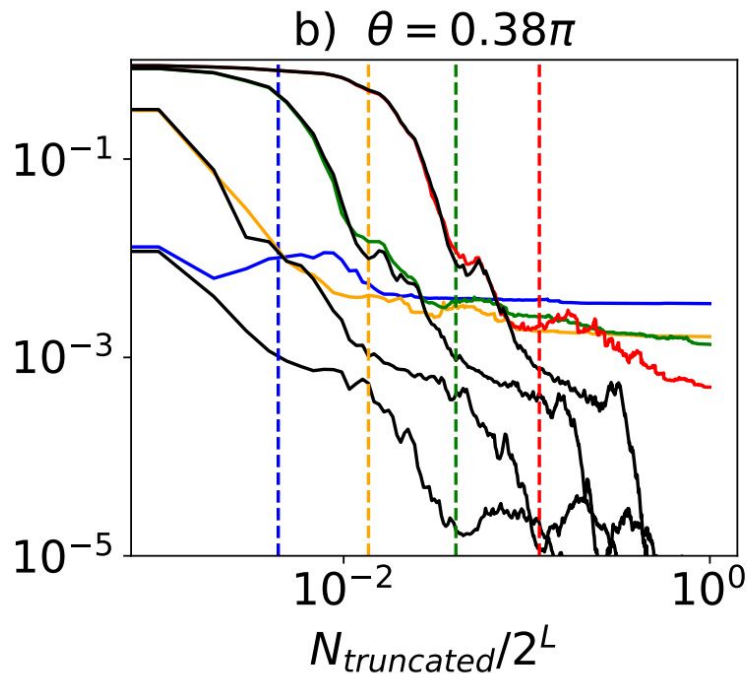
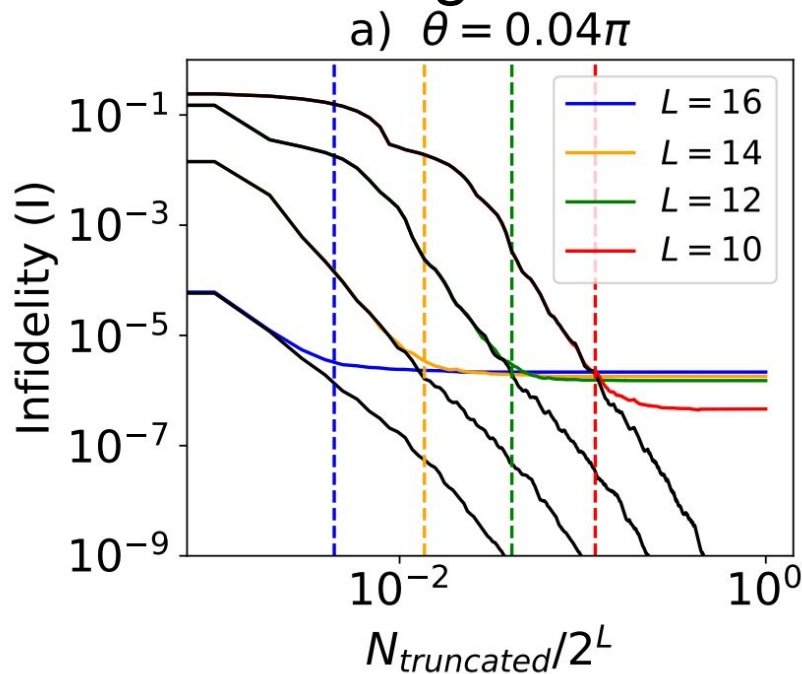
Further observations

- Cumulant expansion convergence continues to be highly indicative of NQS performance!
- In some case, RBM's can also “outperform” cumulant expansion

Infidelity (I)



Finite size scaling



Results are consistent with system size,
despite exponential growth.

Outlook: What to we learn from this?

Optimize the basis before training!

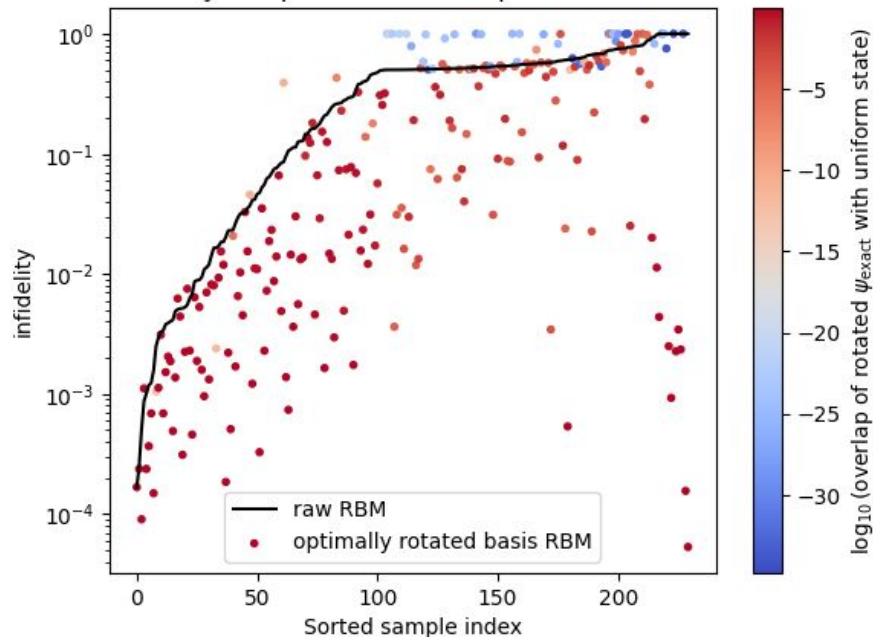
(i) Exact-diagonalize a small cluster. Check for near-degeneracies.

(ii) Compute the cumulant expansion. Expect $\text{RBM} \approx \Psi^{(N_{\text{var}})}$.

(iii) Find the basis where the cumulant expansion converges fastest.

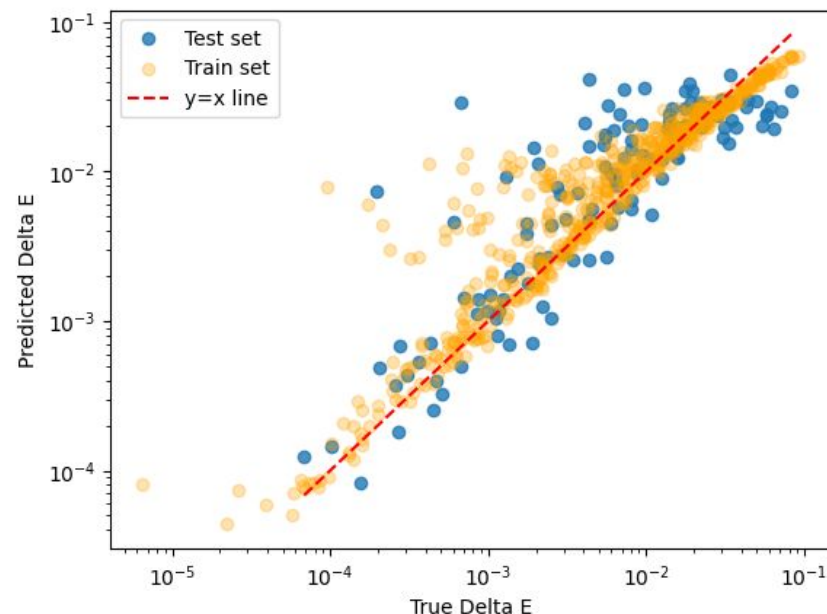
$$H = \sum_{i=1}^N \sum_{k=1}^K \sum_{\alpha, \beta \in \{X, Y, Z\}} S_{i, \alpha} J_{\alpha \beta}^{(k)} S_{i+k, \beta} + \sum_{i, \alpha} h_{\alpha} S_{i, \alpha},$$

infidelity comparison: raw vs optimal basis RBM



Outlook: Is this all?

- Cumulant expansion is highly correlated to performance, but there are also other performance indicators
- Can we “predict” the performance of Neural nets just from the ground state?
- Answer: Maybe (need more testing)

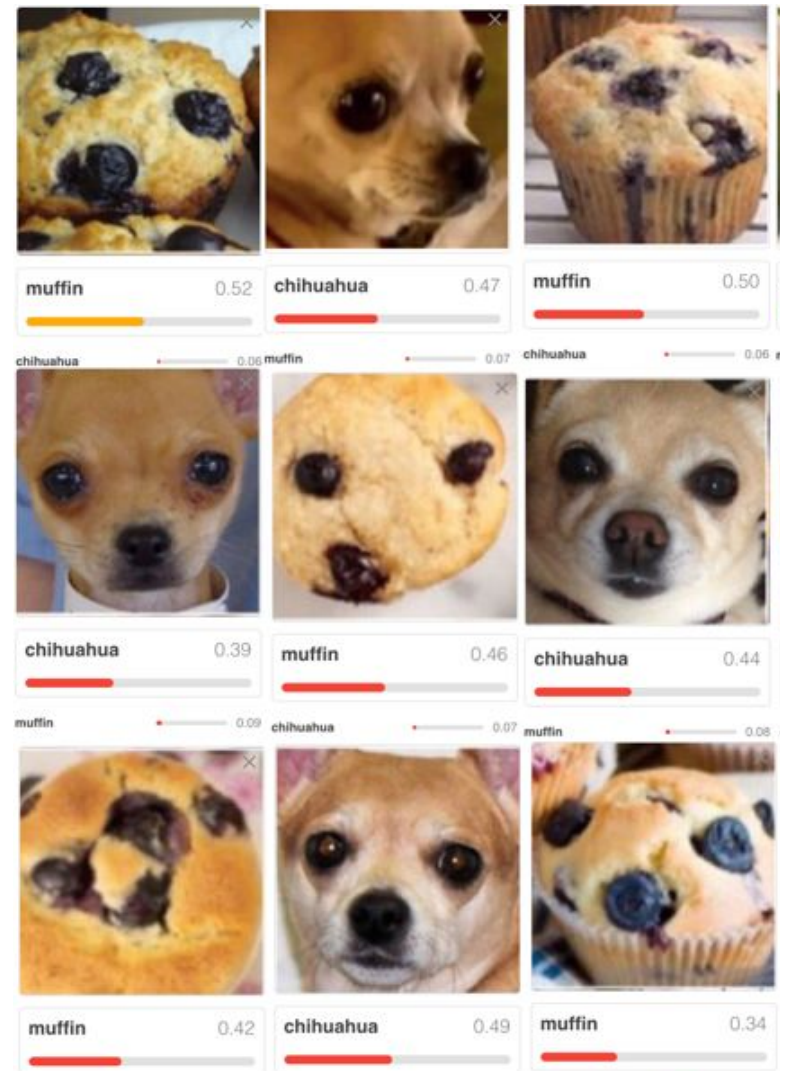


Summary

- Need to distinguish **representability, learnability, sampling**
- Local **basis rotations** influence **accuracy** significantly
- Neural networks **prefer** to learn **uniform** states if variational energy allows it
- **Cumulant expansion** captures the **difficulty** of approximating a given state
- Some bases require **more correlations** to describe than others
- Finding a measure of complexity would possibly allow many **improvements** with minimal overhead

Thank you for your attention!

(here is a neural network trying to distinguish chihuahuas from muffins)



Bonus: real time dynamics (quench from para - ferro)

