Simulating Quantum Many-Body Systems with Language Models

Stef Czischek

June 12, 2024

uOttawa

APRIQuOt

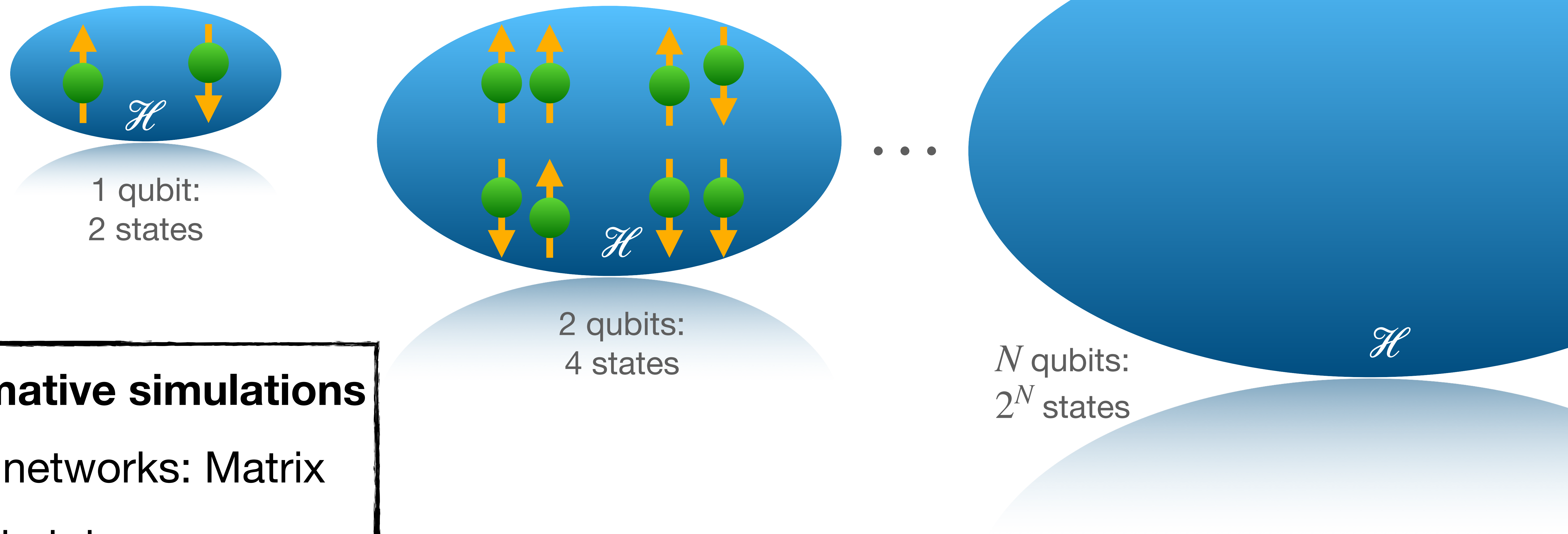
Introduction

Qubit systems

Infeasible to simulate classically

Powerful when well-controlled

Hilbert space: space of all possible states



Control quantum system

- Quantum computation and quantum simulation
- NISQ era: Noisy intermediate-scale quantum era

Approximative simulations

- Tensor networks: Matrix product states
- Quantum Monte Carlo
- Neural network quantum states

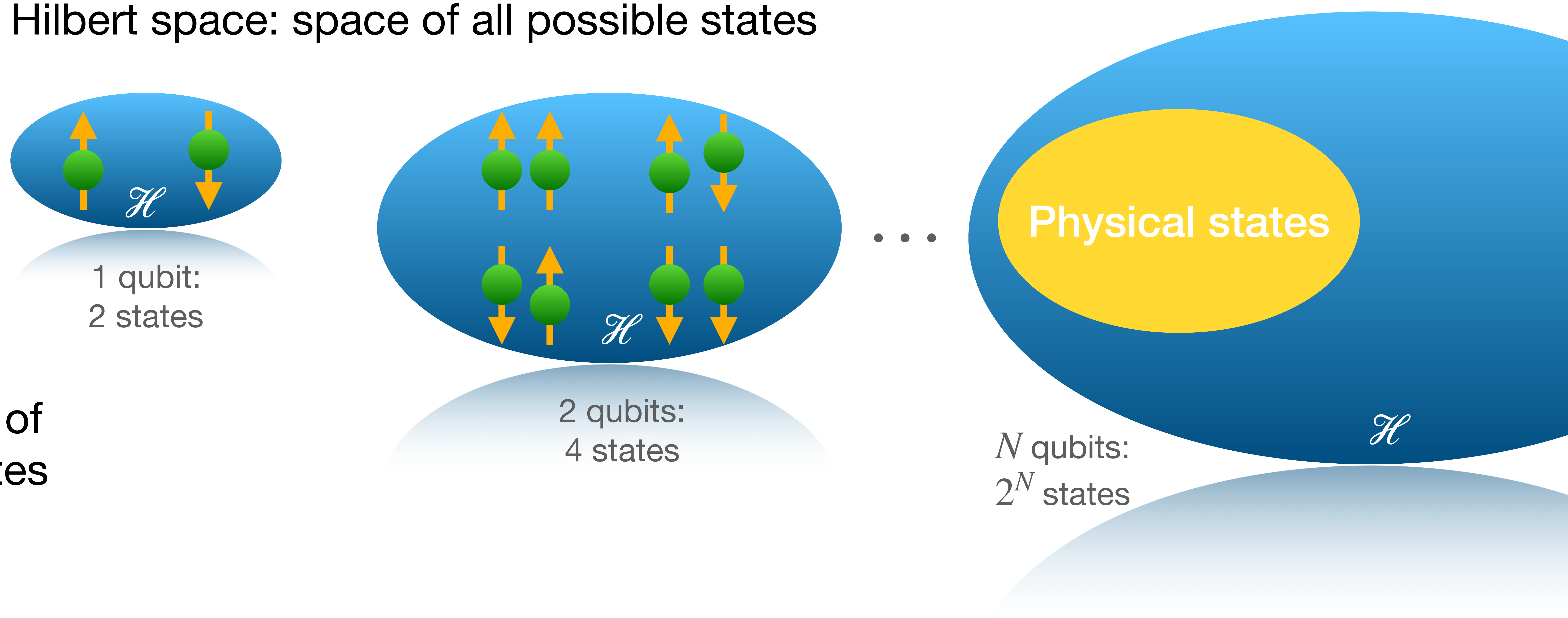
Curse of dimensionality

Stefanie.Czischek@uOttawa.ca

Approximation methods

"Physical states" describe the dominant behaviour of the qubit system.

Hilbert space: space of all possible states



Solve exponentially scaling problem → Find subset of physical states

Squared wave function amplitude $\left|\psi(\boldsymbol{\varphi})\right|^2$:

probability distribution underlying states

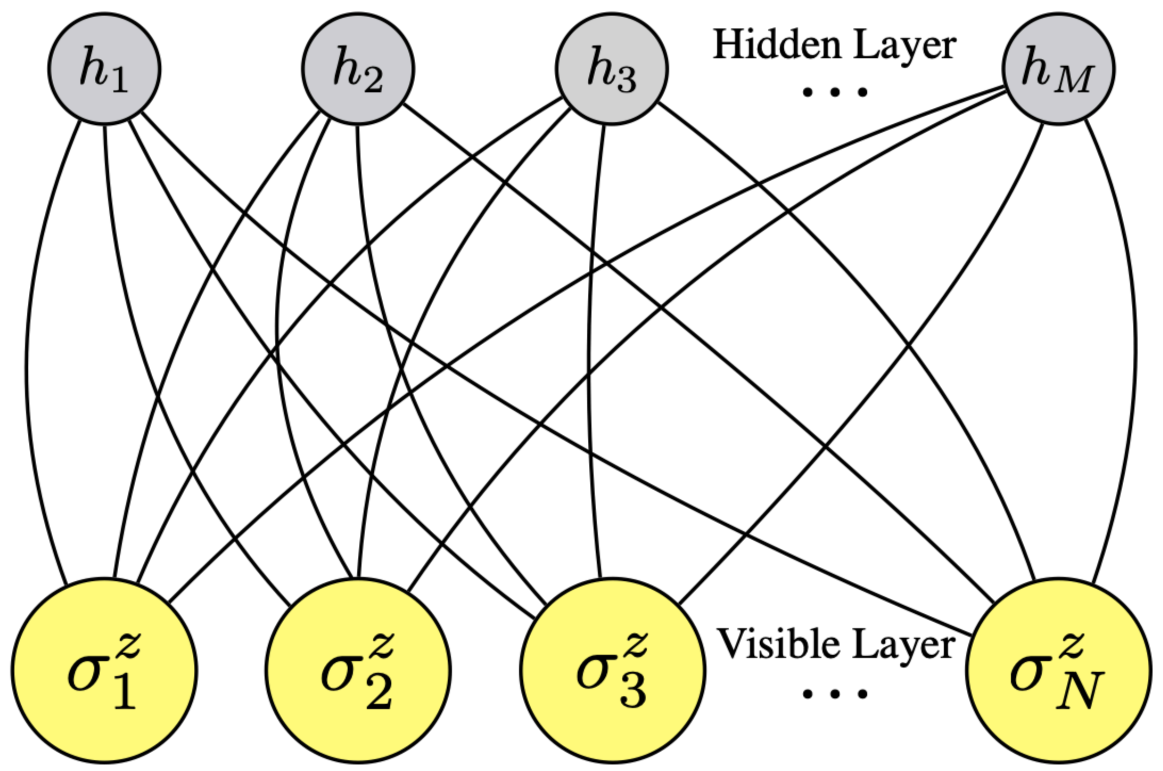
$\boldsymbol{\varphi} \in \{(0,0,\ldots,0), \ldots, (1,1,\ldots,1)\}$

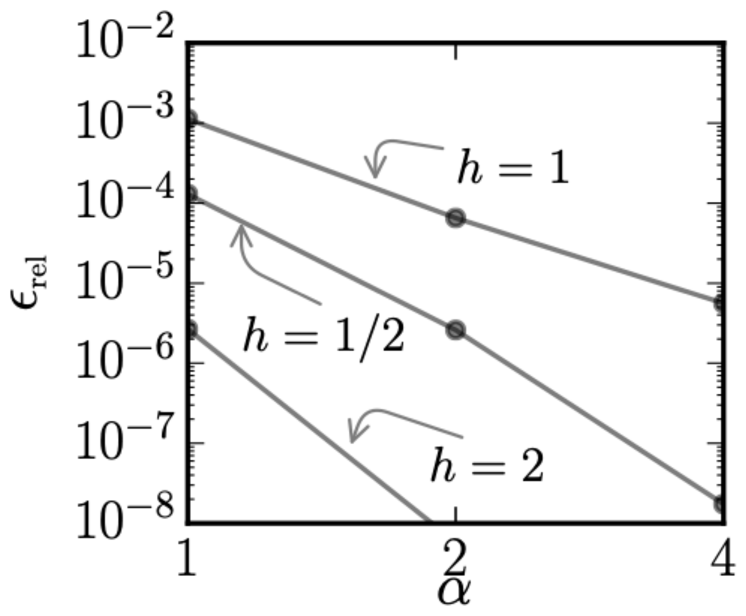
- Find numerical expression for $\left|\psi(\boldsymbol{\varphi})\right|^2$
- Sample physical states

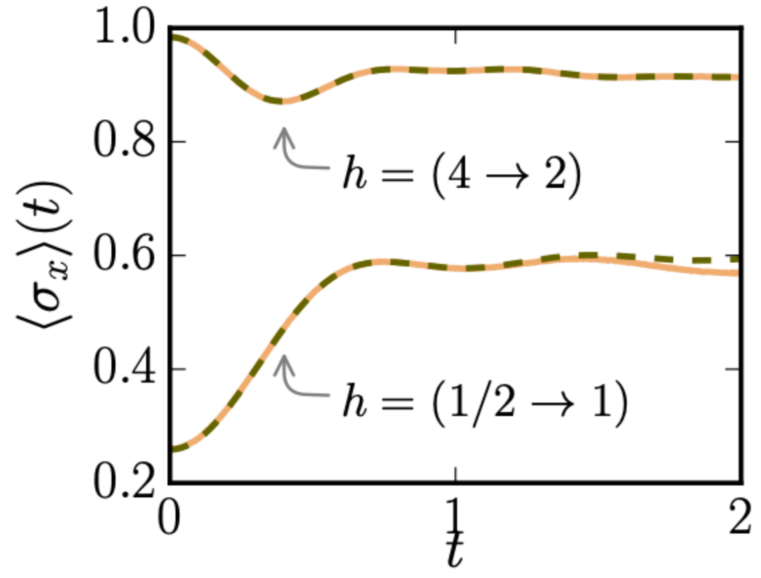
Stefanie.Czischek@uOttawa.ca

Neural-network quantum states

Artificial neural networks to represent ground states and model dynamics

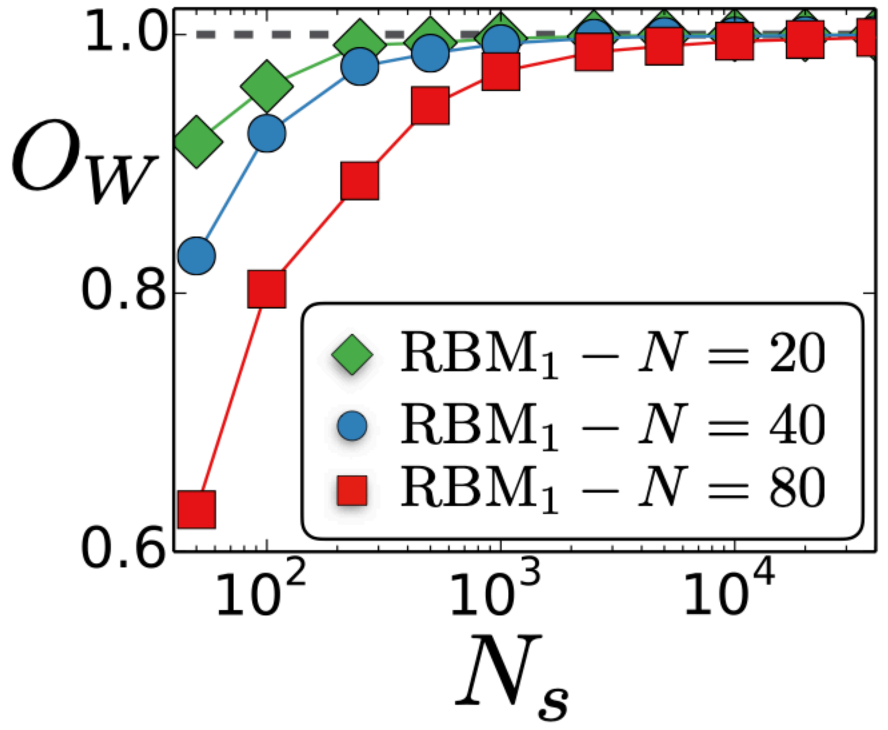


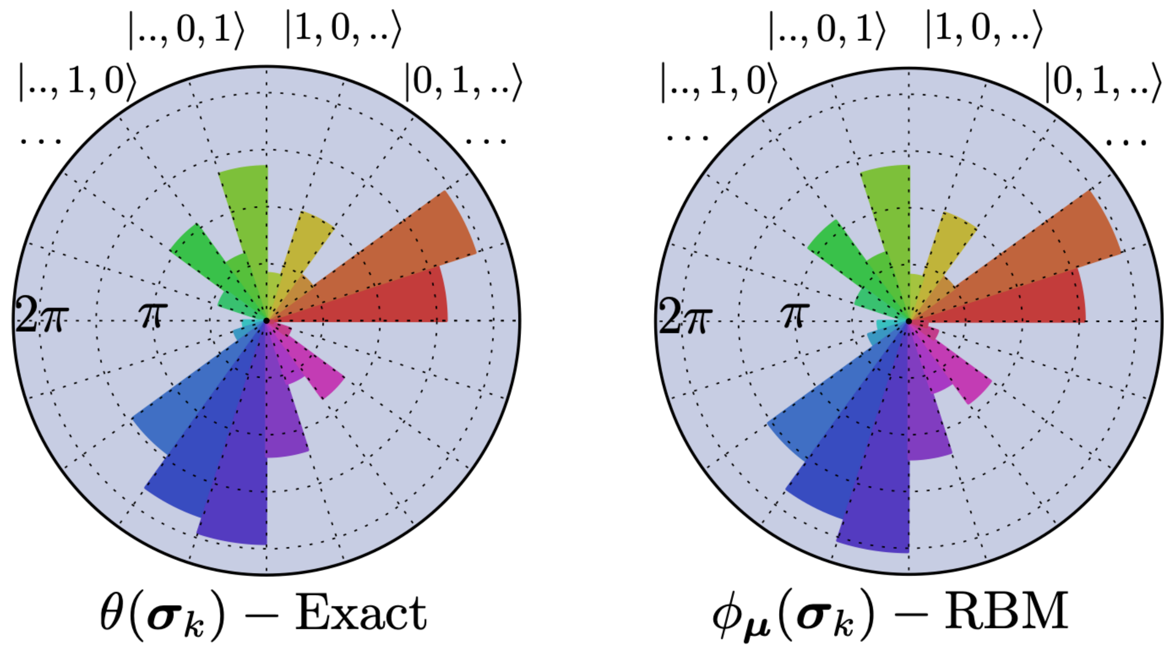




[Carleo & Troyer, Science **355** (2017)]

Artificial neural networks for quantum state tomography



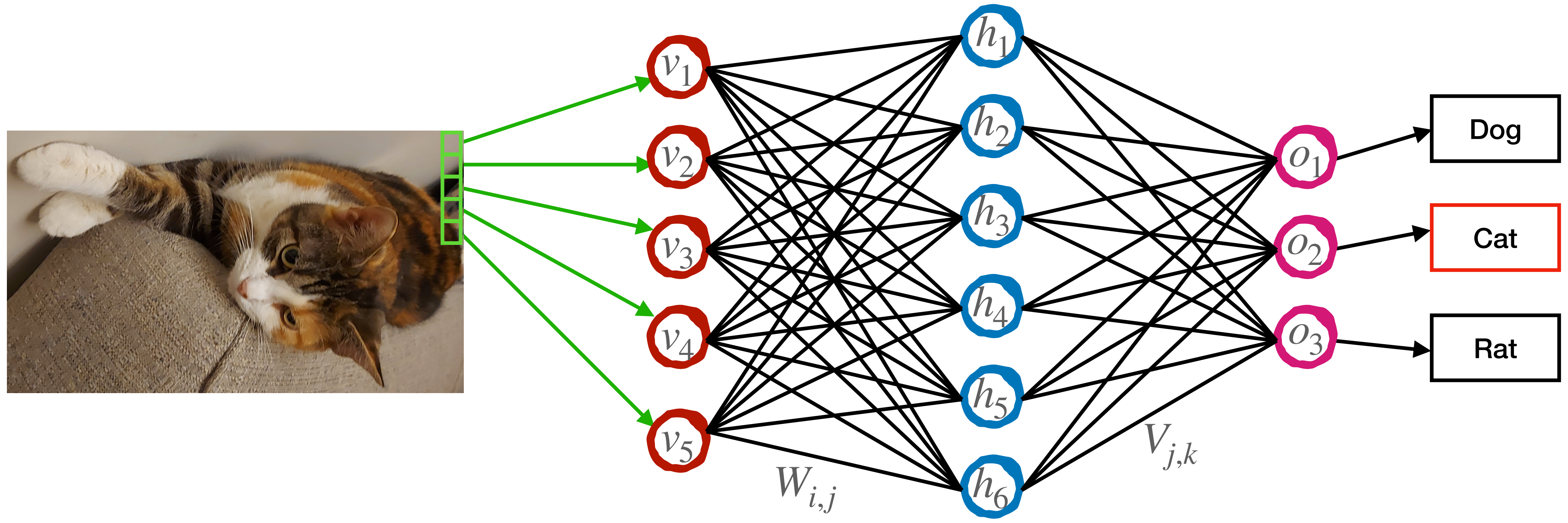


[Torlai, Melko, et al., Nature Physics **14** (2018)]

Stefanie.Czischek@uOttawa.ca

Artificial neural networks: applications

Image classification



Text prediction

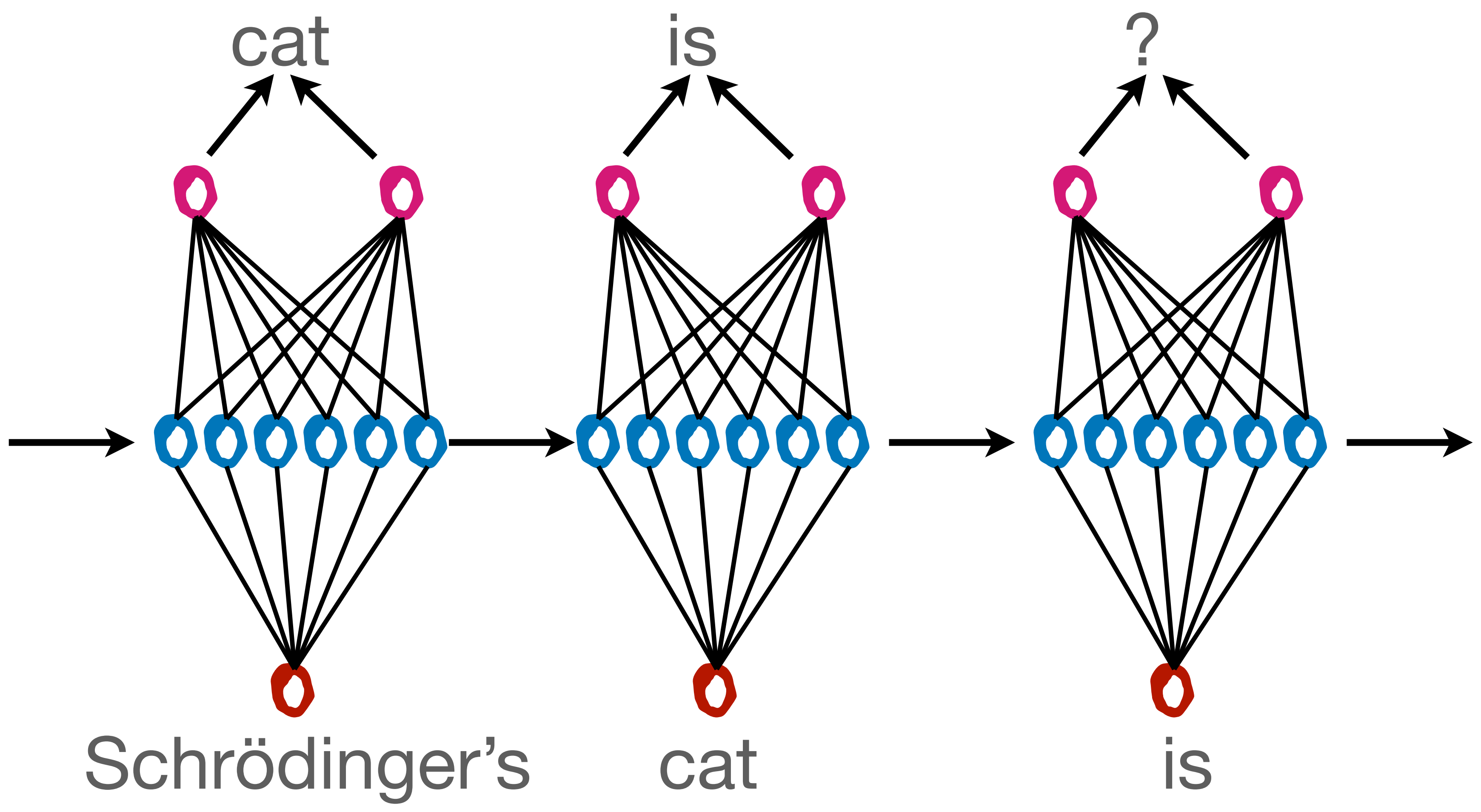


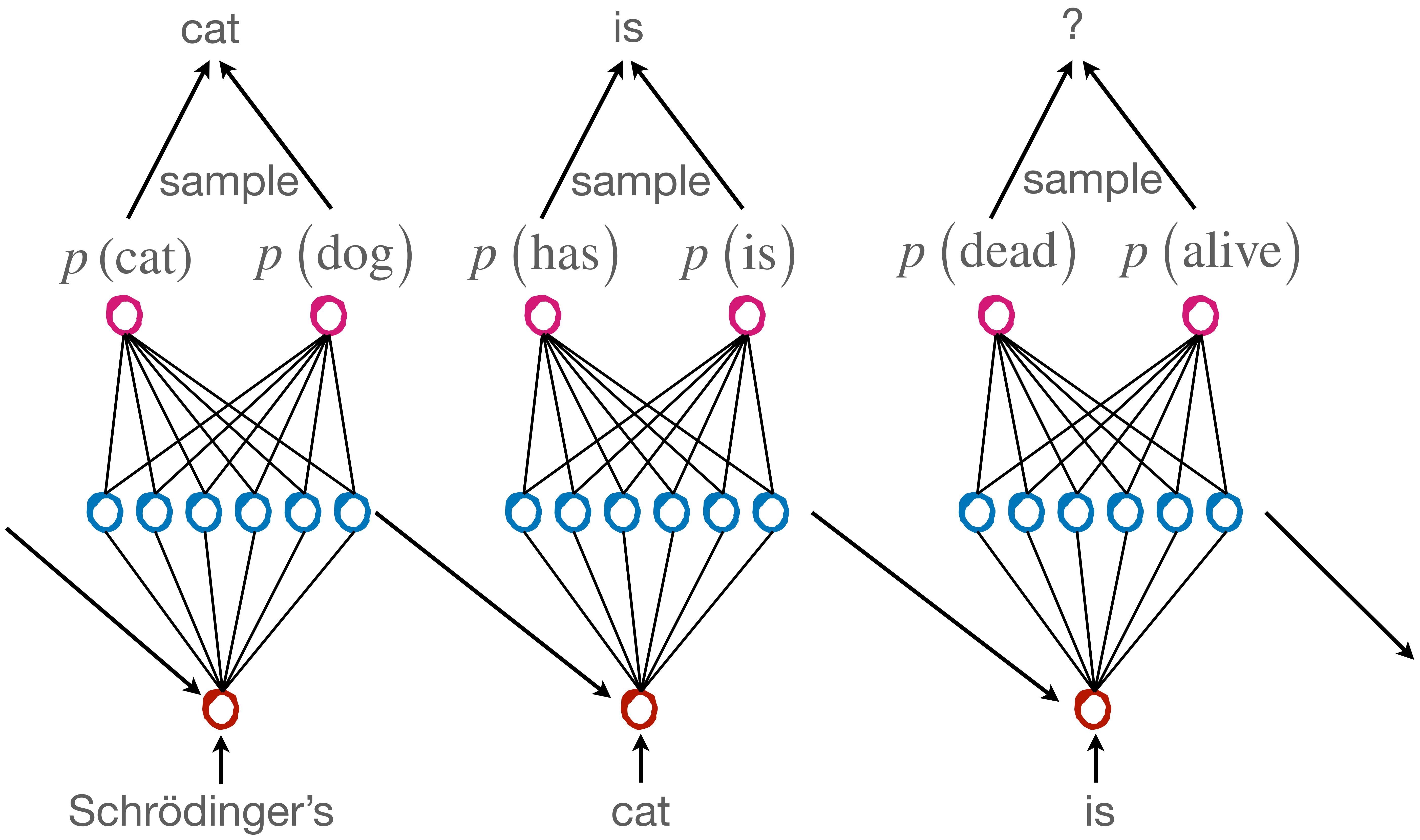
Image generation

thiscatdoesnotexist.com

Stefanie.Czischek@uOttawa.ca

Text prediction: Language models

Text prediction



Memory effect:
Previous hidden state is used as additional input

Output probabilities are conditioned on previous words

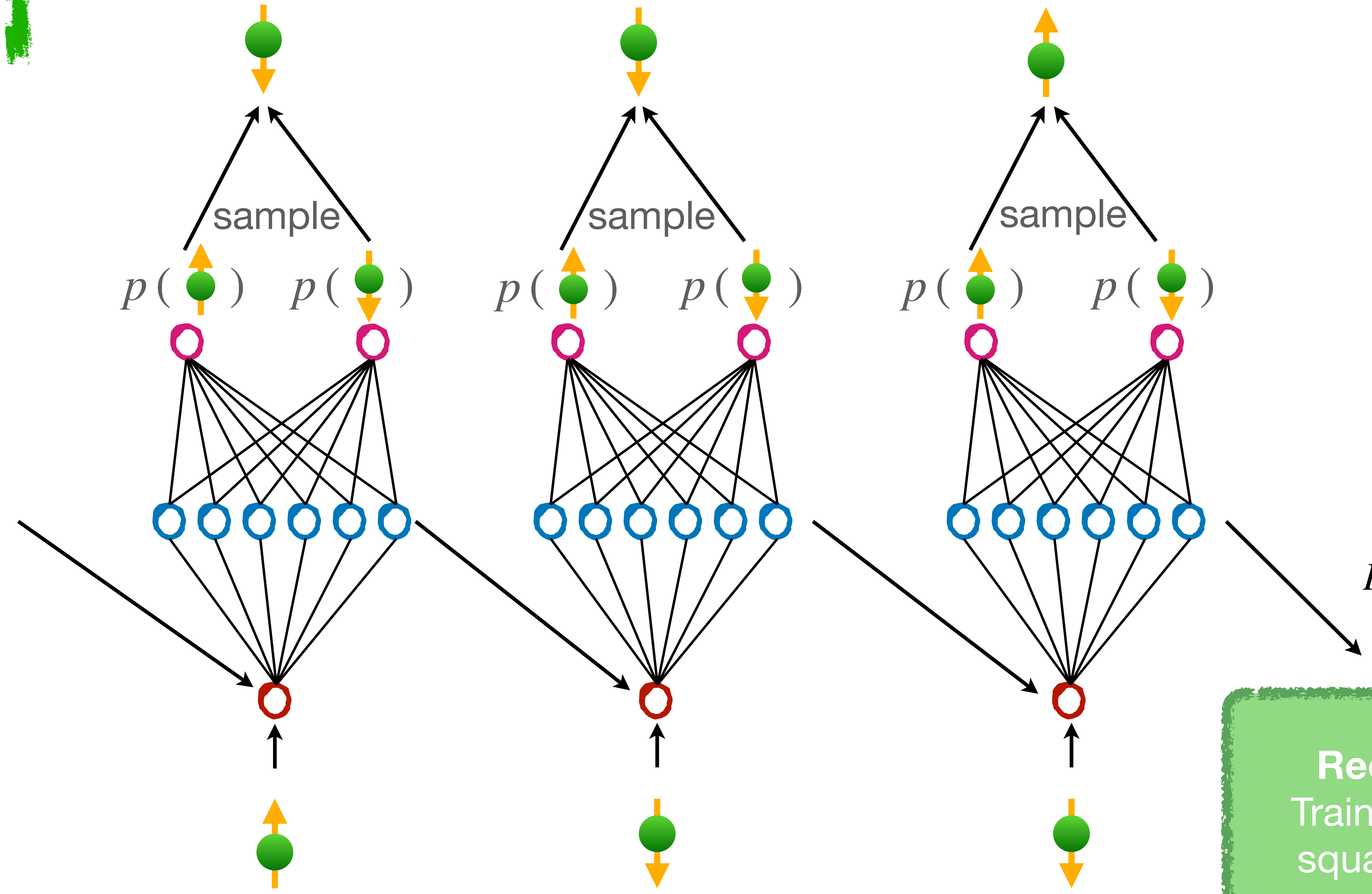
$$p\left(\text{alive}\right) = P\left(\text{alive}\,|\,\text{Schrödinger's, cat, is}\right)$$

$$P\left(\boldsymbol{w}\right) = \prod_{i=1}^{N} P\left(w_i\,|\,w_{<i}\right)$$

Stefanie.Czischek@uOttawa.ca

Language model quantum states

[O Sharir et al., PRL **124** (2020)]
[M Hibat-Allah et al., PRR **2** (2020)]

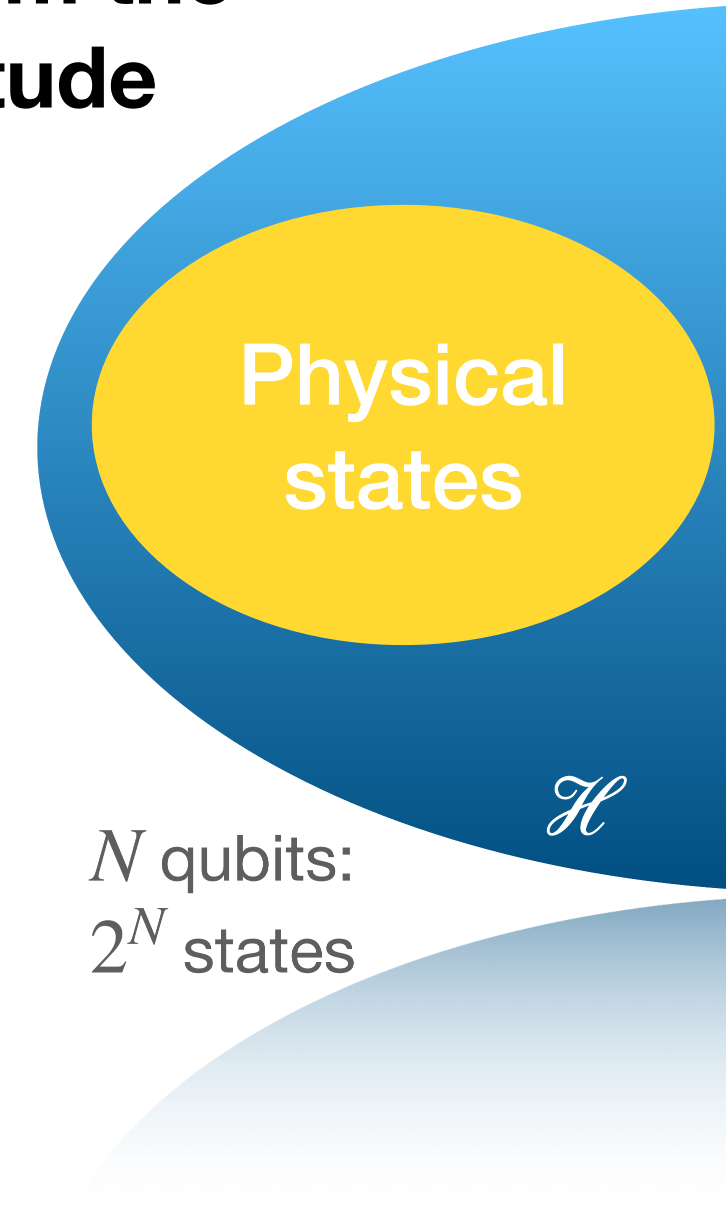


Language models sample from a probability distribution

Physical states are sampled from the squared wave function amplitude

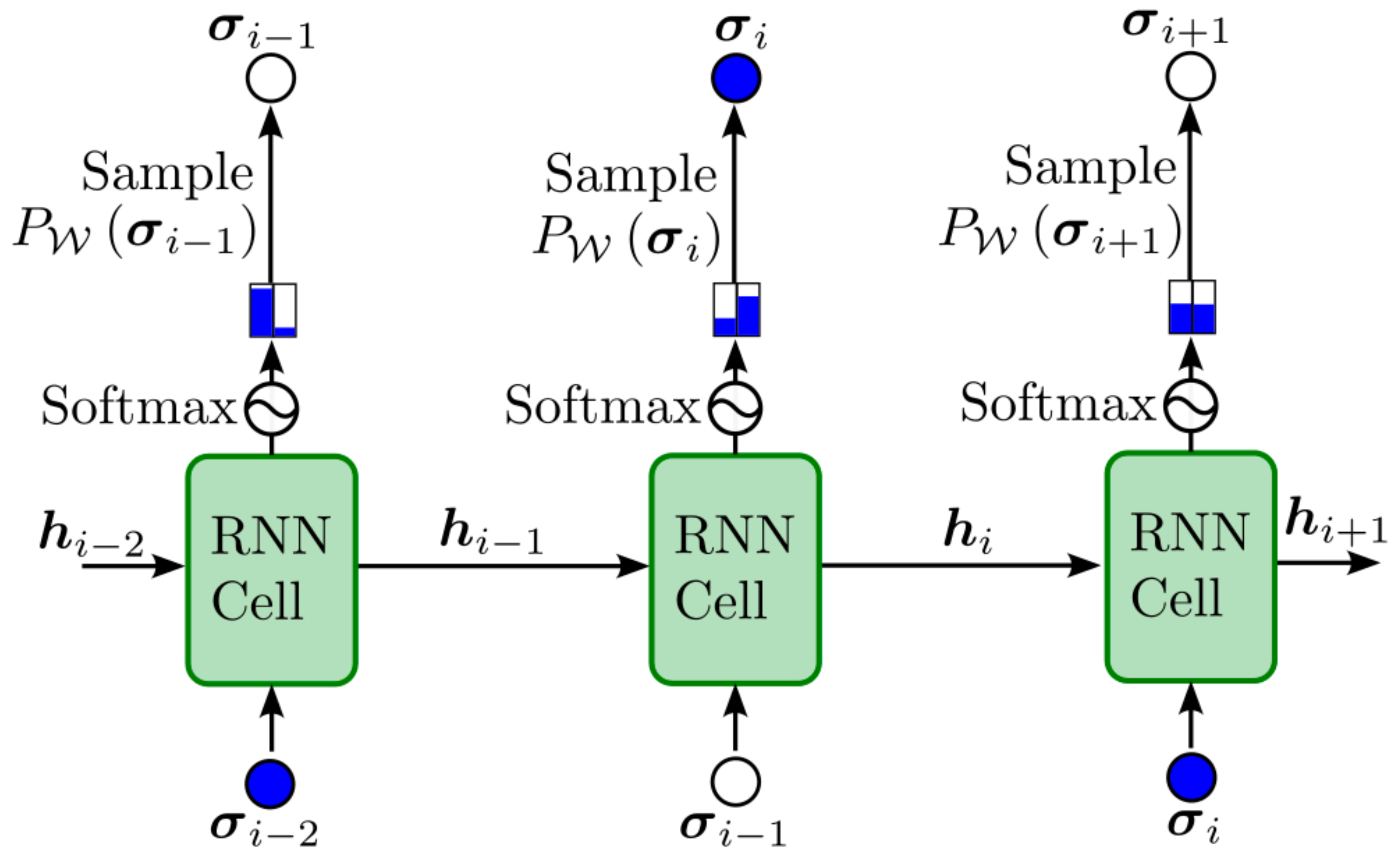
$$P(\boldsymbol{\varphi}) = \prod_{i=1}^{N} P\left(\varphi_i | \varphi_{<i}\right) \stackrel{!}{=} \left|\Psi\left(\boldsymbol{\varphi}\right)\right|^2$$

Reconstruct quantum states:
Train a language model to encode squared wave function amplitude

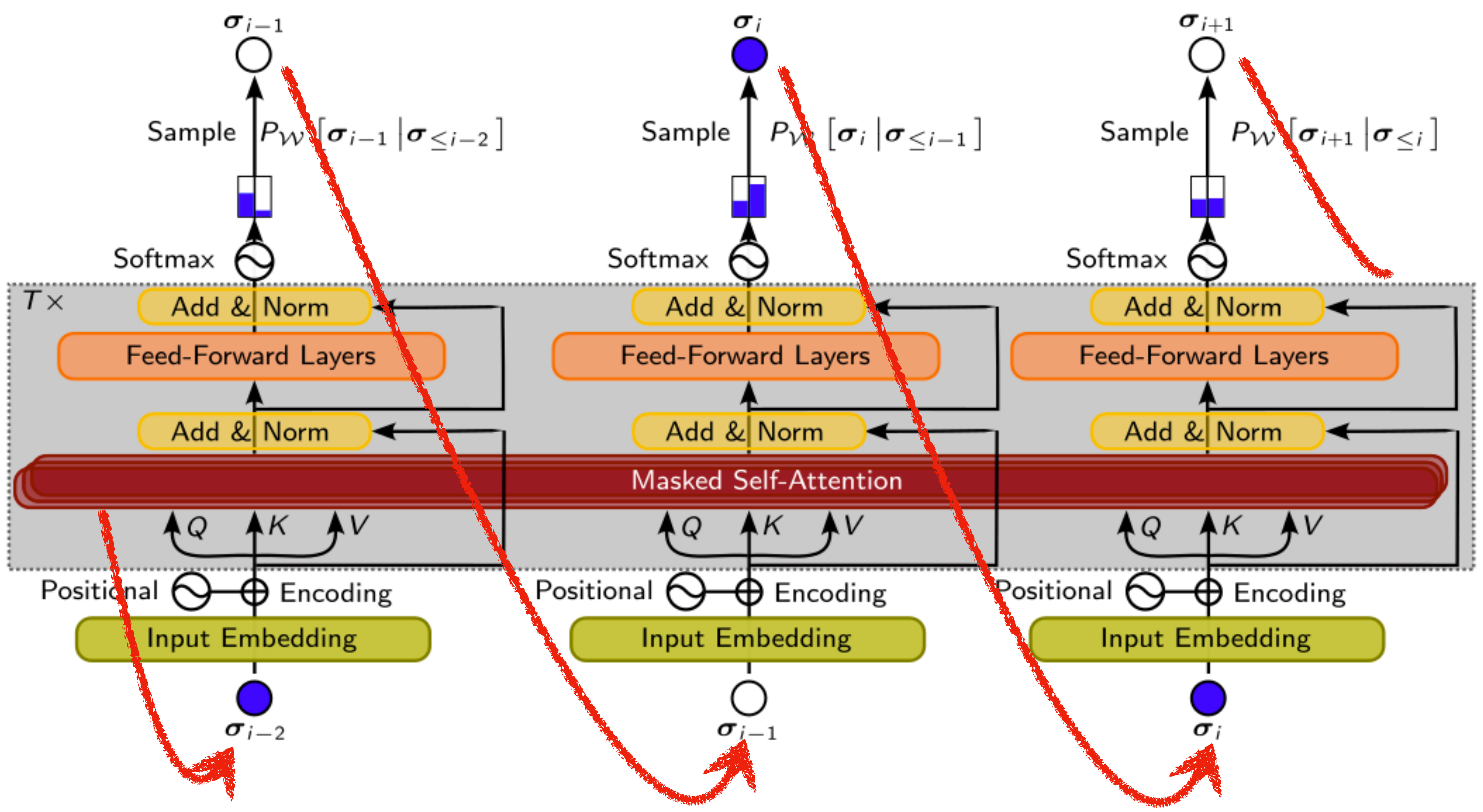


Stefanie.Czischek@uOttawa.ca

Language model quantum states



Recurrent Neural Networks (RNNs)



Transformer Models

[Sharir et al., PRL **124** (2020)]
[Hibat-Allah et al., PRR **2** (2020)]
[SC et al., PRB **105** (2022)]
[M Hibat-Allah et al., Nat Phys **15** (2019)]
[Hibat-Allah et al., PRB **108** (2023)]
⋮

[Melko & Carrasquilla, Nat Comput Sci **4** (2023)]
[Cha et al., ML: Sci Tech **3** (2021)]
[Ma et al., arXiv:2305.05433 (2023)]
[Zhang & Di Ventra, PRB **107** (2023)]
[Viteritti et al., PRL **130** (2023)]
[Sharir et al., arXiv:2212.11296 (2022)]
[Sprague & SC, Comm Phys **7** (2024)]
⋮

Nice review:
From Architectures to Applications: A Review of Neural Quantum States
[Lange et al., arXiv:2402.09402 (2024)]

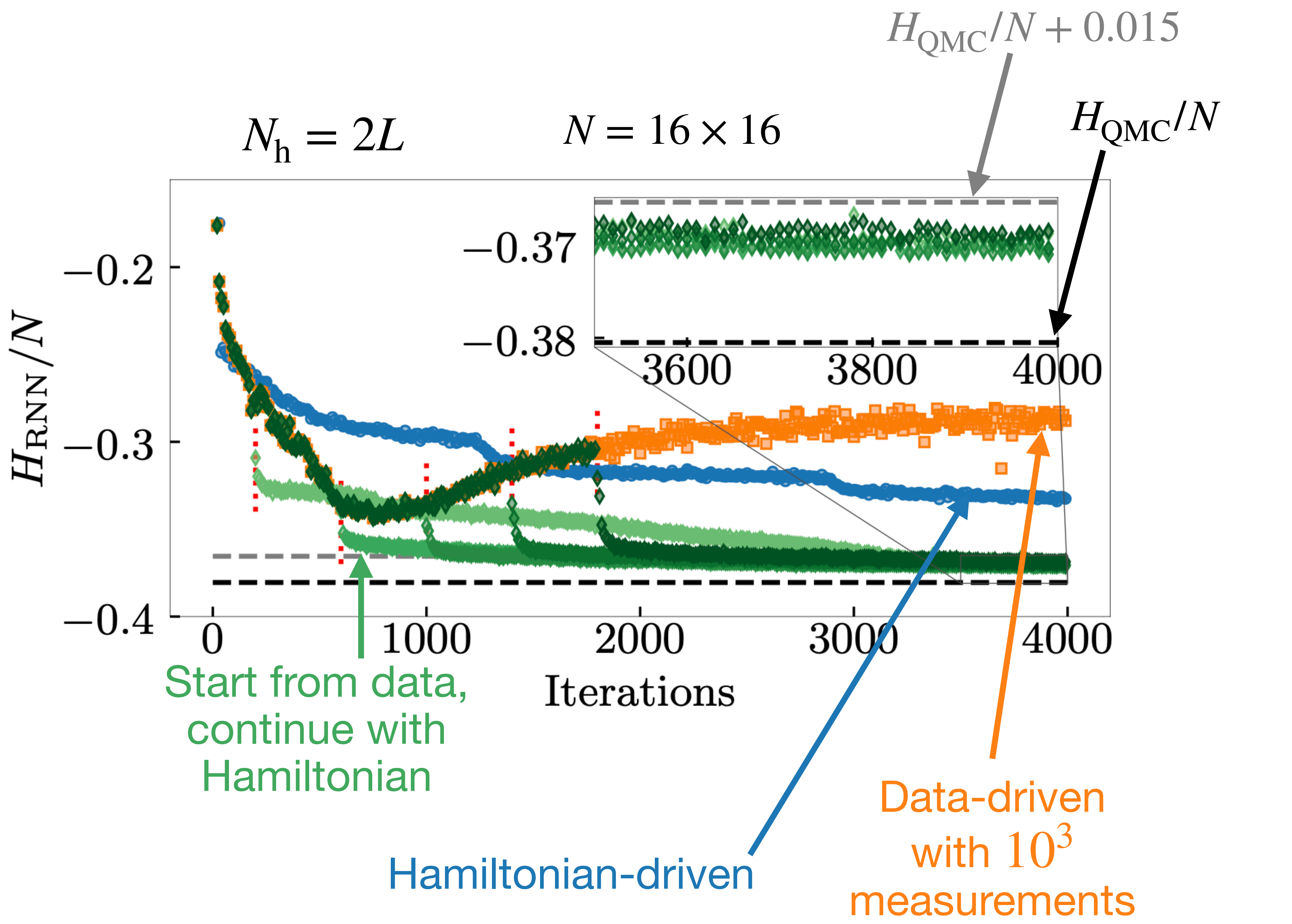
Stefanie.Czischek@uOttawa.ca

APRIQuOt uOttawa

Outlook

Hybrid training

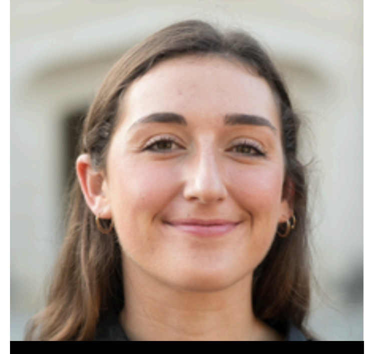
[SC et al., Phys. Rev. B **105** (2022)]



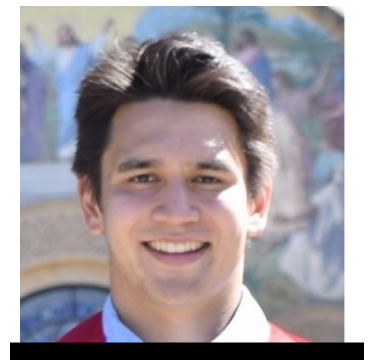
[E. Merali et al., arXiv:2107.00766 (2021)]

Slow convergence…

Overfitting…

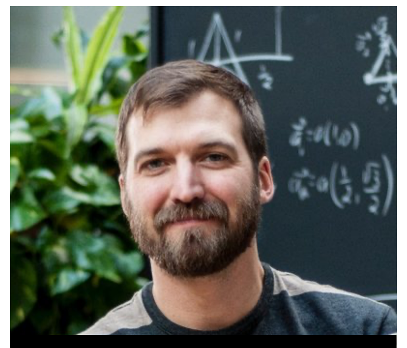


Schuyler Moss



Matthew Radzihovsky

Ejaaz Merali



Roger Melko

$$\hat{H} = -\frac{\Omega}{2}\sum_{i=1}^{N}\hat{\sigma}_i^x - \delta\sum_{i=1}^{N}\hat{n}_i + \sum_{i,j} V_{ij}\hat{n}_i\hat{n}_j \qquad V_{ij} = \frac{\Omega R_{\mathrm{b}}^6}{|\boldsymbol{r}_i - \boldsymbol{r}_j|^6} = \frac{7}{|\boldsymbol{r}_i - \boldsymbol{r}_j|^6}$$

$$\Omega = \delta = 1$$

- Start with data-driven training
- Switch to Hamiltonian-driven training at different points in time
- Independent of when the switch happens, we can significantly reduce the convergence time

A limited amount of measurement data can enhance ground state searches!

[Moss, Melko et al., arXiv:2308.02647 (2023)]

Stefanie.Czischek@uOttawa.ca

Long-Range Correlations: Transformers

Transformer neural networks and quantum simulators: a hybrid approach for simulating strongly correlated systems

Hannah Lange, Guillaume Bornet, Gabriel Emperauger, Cheng Chen, Thierry Lahaye, Stefan Kienle, Antoine Browaeys, Annabelle Bohrdt

Towards Neural Variational Monte Carlo That Scales Linearly with System Size

Or Sharir, Garnet Kin-Lic Chan, Anima Anandkumar

Transformer Wave Function for the Shastry-Sutherland Model: emergence of a Spin-Liquid Phase

Luciano Loris Viteritti, Riccardo Rende, Alberto Parola, Sebastian Goldt, Federico Becca

Transformer variational wave functions for frustrated quantum spin systems

Luciano Loris Viteritti, Riccardo Rende, Federico Becca

Are queries and keys always relevant? A case study on Transformer wave functions

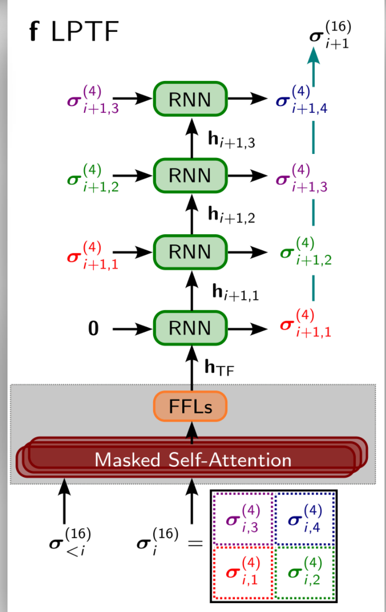
Riccardo Rende, Luciano Loris Viteritti

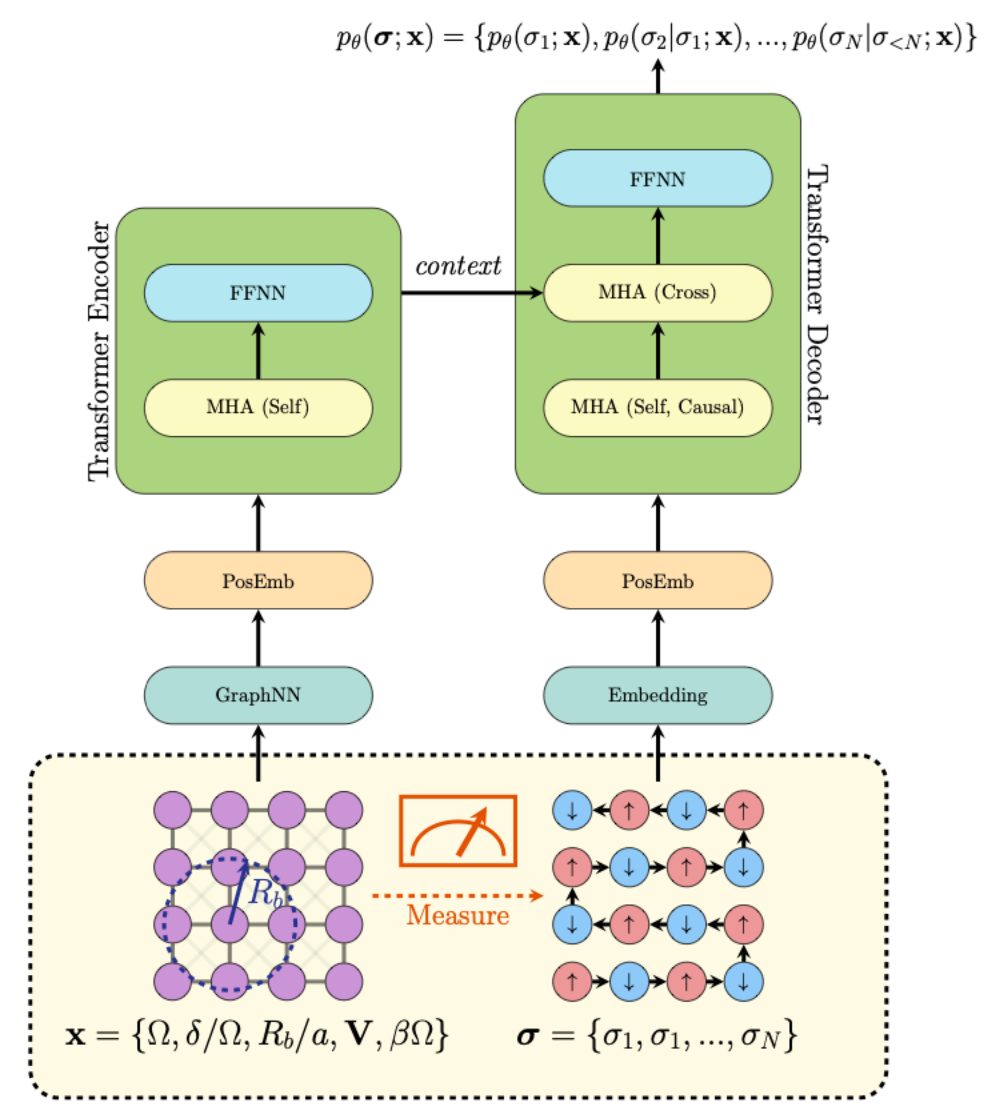
A simple linear algebra identity to optimize Large-Scale Neural Network Quantum States

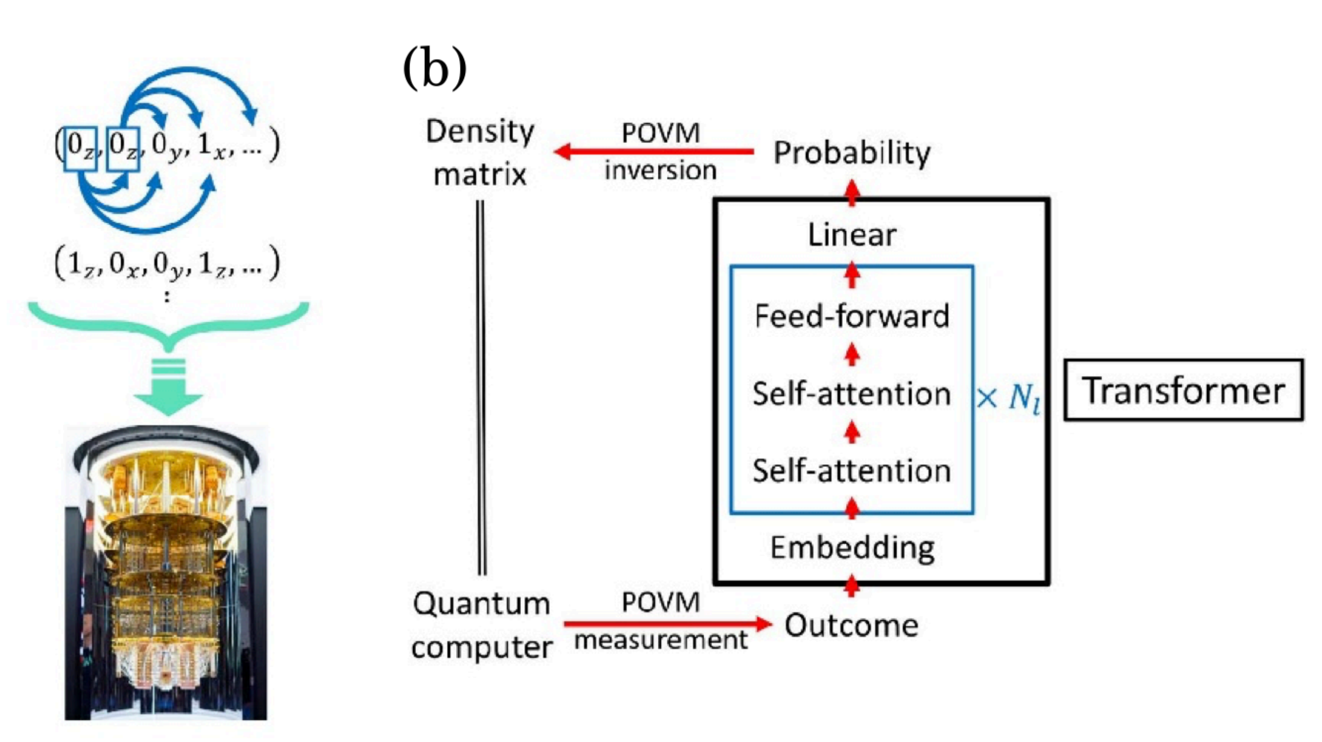
Riccardo Rende, Luciano Loris Viteritti, Lorenzo Bardone, Federico Becca, Sebastian Goldt

Variational Monte Carlo with Large Patched Transformers

Kyle Sprague, Stefanie Czischek







Attention-based quantum tomography

Peter Cha[6,1], Paul Ginsparg[2], Felix Wu[2], Juan Carrasquilla[3,4], Peter L McMahon[5] and Eun-Ah Kim[1]

Published 23 November 2021 • © 2021 The Author(s). Published by IOP Publishing Ltd
Machine Learning: Science and Technology, Volume 3, Number 1
Citation Peter Cha *et al* 2022 *Mach. Learn.: Sci. Technol.* **3** 01LT01
DOI 10.1088/2632-2153/ac362b

Attention-Based Transformer Networks for Quantum State Tomography

Hailan Ma, Zhenhong Sun, Daoyi Dong, Chunlin Chen, Herschel Rabitz

RydbergGPT

David Fitzek, Yi Hong Teoh, Hin Pok Fung, Gebremedhin A. Dagnew, Ejaaz Merali, M. Schuyler Moss, Benjamin MacLellan, Roger G. Melko

Transformer quantum state: A multipurpose model for quantum many-body problems

Yuan-Hang Zhang and Massimiliano Di Ventra
Phys. Rev. B **107**, 075147 – Published 22 February 2023

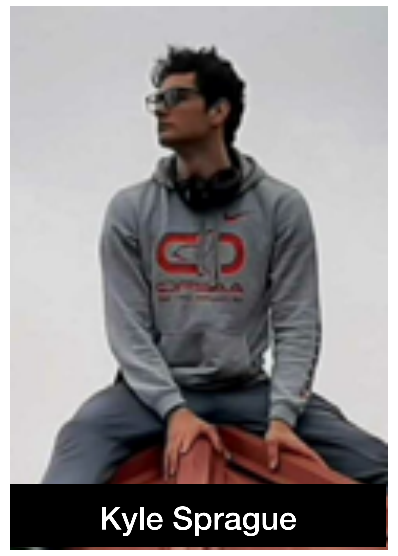
Unified Quantum State Tomography and Hamiltonian Learning Using Transformer Models: A Language-Translation-Like Approach for Quantum Systems

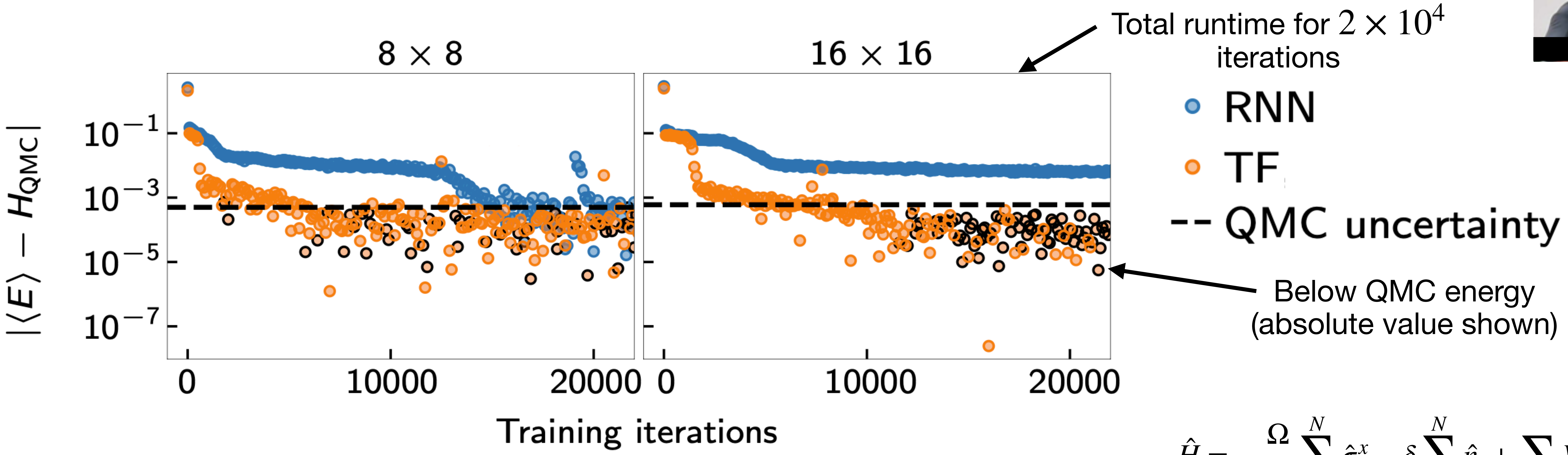
Zheng An, Jiahui Wu, Muchun Yang, D. L. Zhou, Bei Zeng

Stefanie.Czischek@uOttawa.ca

Performance comparison

[Sprague & SC, Comm Phys **7** (2024)]





$$\hat{H} = -\frac{\Omega}{2}\sum_{i=1}^{N}\hat{\sigma}_i^x - \delta\sum_{i=1}^{N}\hat{n}_i + \sum_{i,j} V_{ij}\hat{n}_i\hat{n}_j$$

$$\Omega = \delta = 1 \qquad V_{ij} = \frac{7}{|\boldsymbol{r}_i - \boldsymbol{r}_j|^6}$$

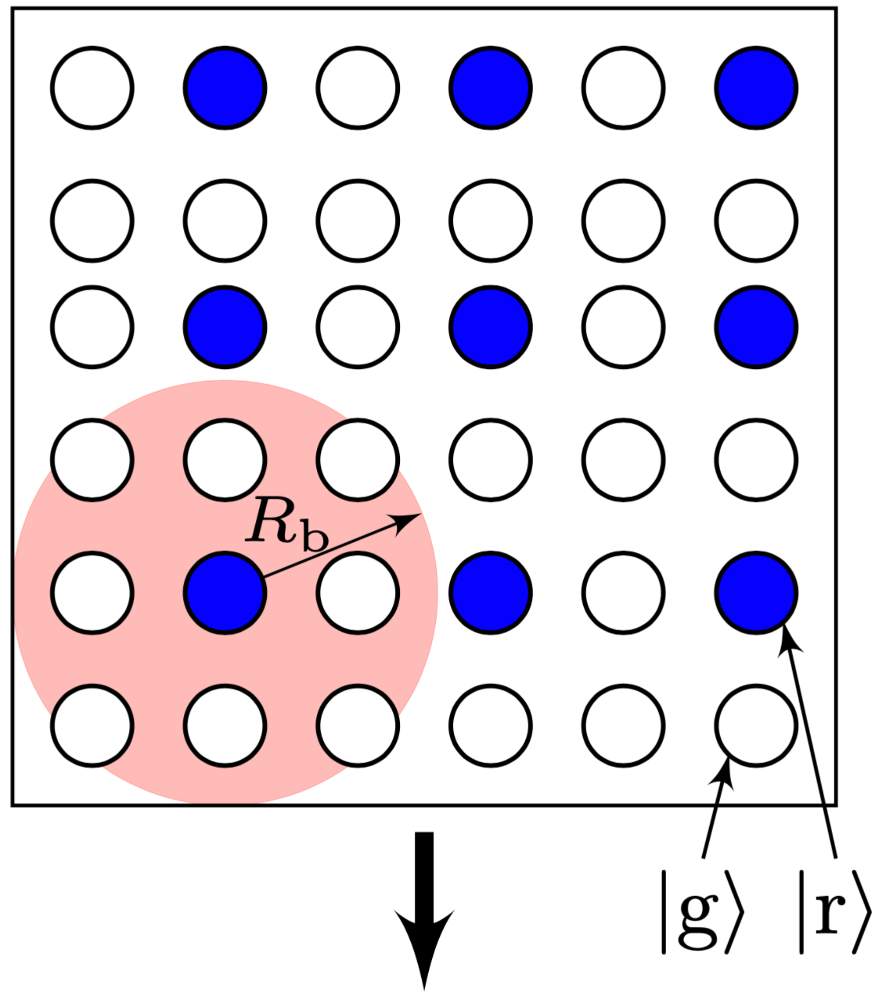
- $\langle E \rangle$: Energy evaluated on 512 neural network samples
- H_{QMC}: Energy evaluated on 7×10^4 quantum Monte Carlo samples
- Transformers outperform RNNs
- But at a high computational cost…

We cannot scale to larger system sizes!

Stefanie.Czischek@uOttawa.ca

Tutorial

Rydberg atom arrays

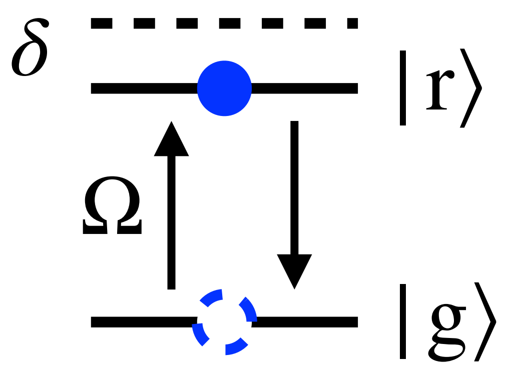


Projective measurement

$|\boldsymbol{\sigma}\rangle = |\mathrm{g\ r\ g} \ldots \mathrm{g\ g}\rangle$

$N = L \times L$

atoms on square lattice



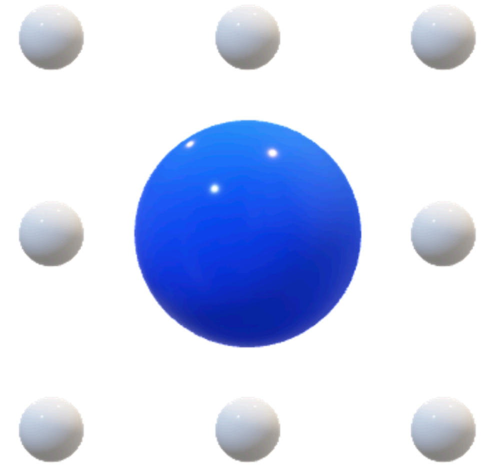
$\hat{\sigma}_i^x = |\mathrm{g}\rangle_i\langle \mathrm{r}|_i$

$\hat{n}_i = |\mathrm{r}\rangle_i\langle \mathrm{r}|_i$

Stoquastic:
Positive, real-valued ground states
All information covered in $\left|\Psi(\boldsymbol{\sigma})\right|^2$

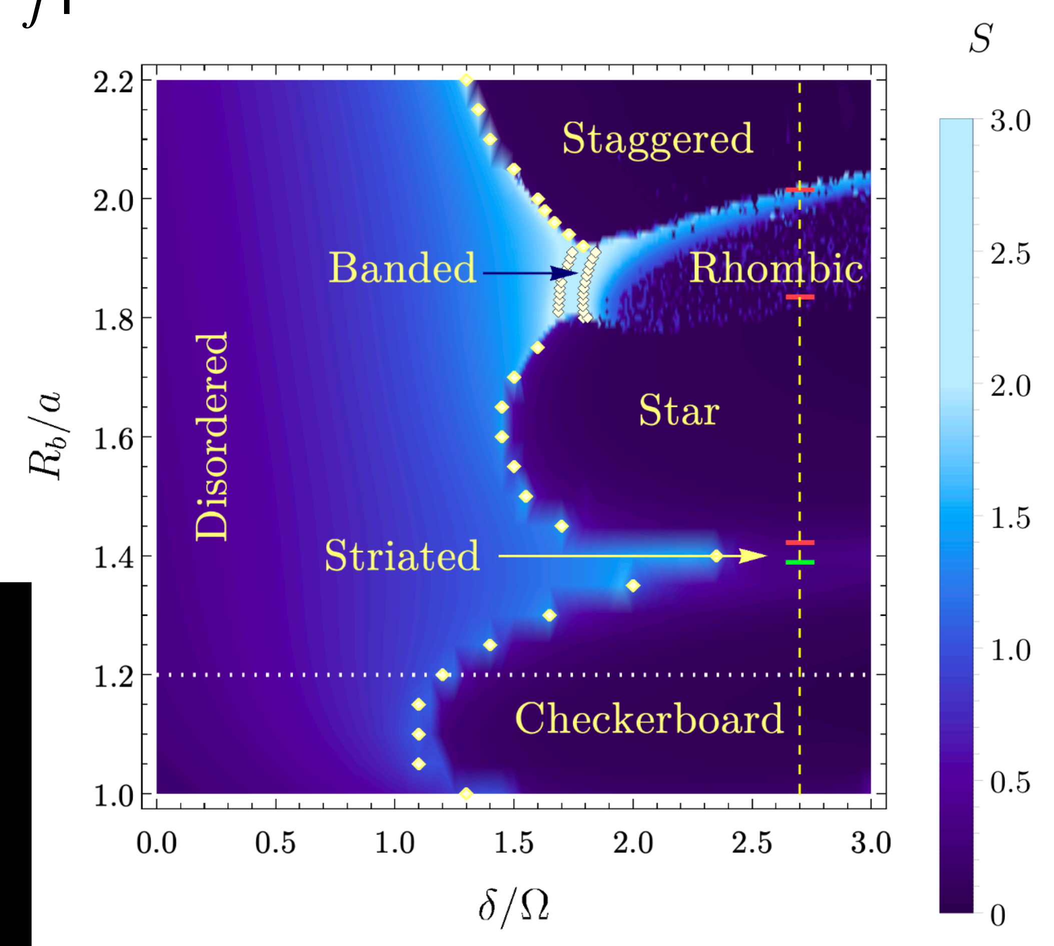
van der Waals interaction:
penalize two excitations within R_b
(Rydberg blockade)

$$V_{ij} = \frac{\Omega R_b^6}{|\boldsymbol{r}_i - \boldsymbol{r}_j|^6}$$

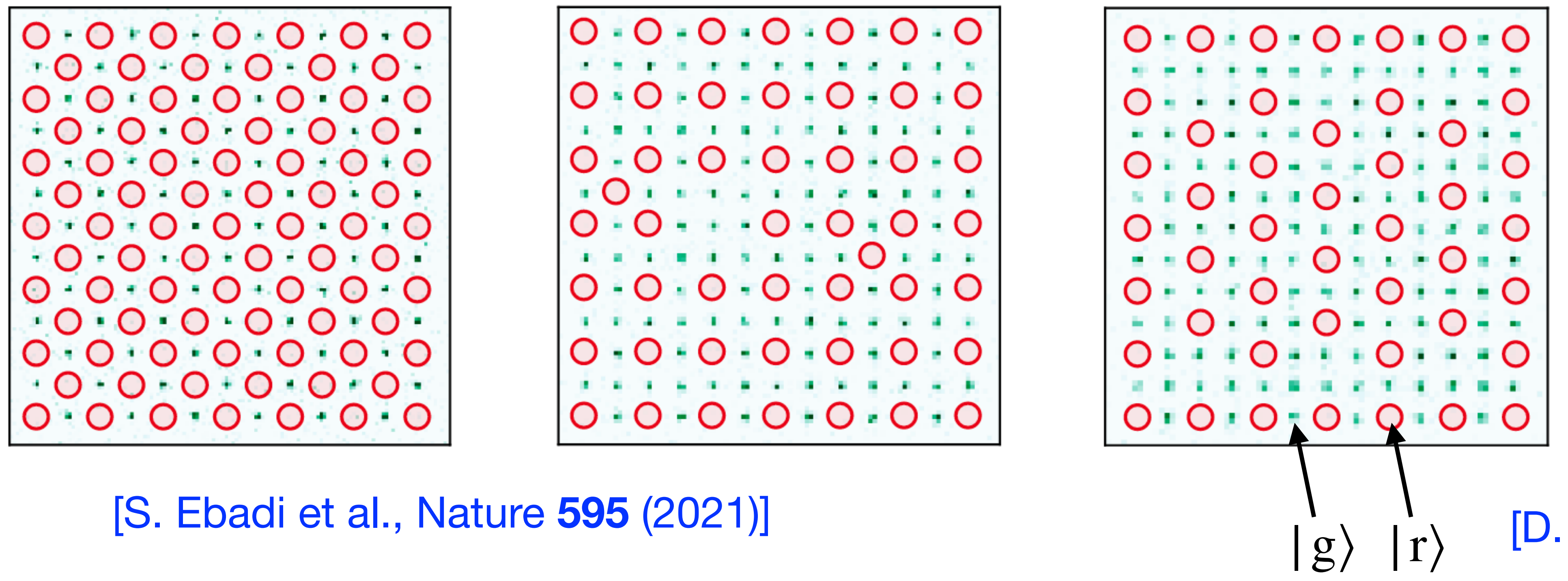


$$\hat{H} = -\frac{\Omega}{2}\sum_{i=1}^{N}\hat{\sigma}_i^x - \delta\sum_{i=1}^{N}\hat{n}_i + \sum_{i,j} V_{ij}\hat{n}_i\hat{n}_j$$

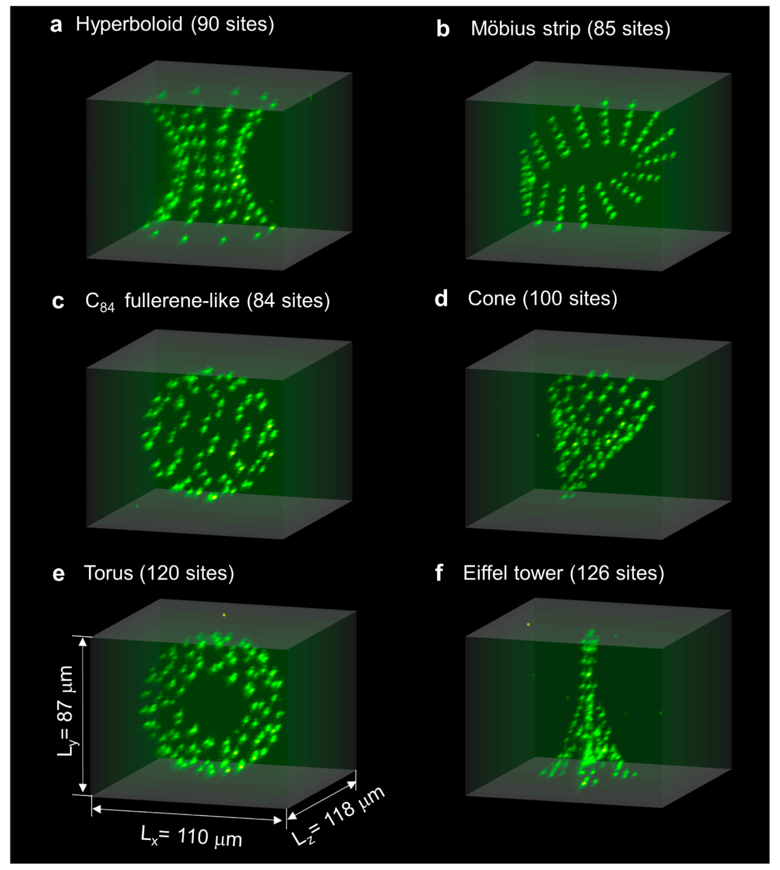
Laser driving:
detuning δ, Rabi frequency Ω



[R. Samajdar et al., PRL **124** (2020)]



[S. Ebadi et al., Nature **595** (2021)]



[D. Barredo et al., Nature **561** (2018)]

Quantum computation & simulation

[D. Jaksch et al., PRL **85** (2000)]
[M. Lukin et al., PRL 87 (2001)]

Stefanie.Czischek@uOttawa.ca

uOttawa