Title: Machine Learning (2021/2022)

Speakers: Lauren Hayward

Collection: Machine Learning (2021/2022)

Date: May 03, 2022 - 11:30 AM

URL: https://pirsa.org/22050010

Abstract: This course is designed to introduce modern machine learning techniques for studying classical and quantum many-body problems encountered in condensed matter, quantum information, and related fields of physics. Lectures will focus on introducing machine learning algorithms and discussing how they can be applied to solve problem in statistical physics. Tutorials and homework assignments will concentrate on developing programming skills to study the problems presented in lecture.

Quantum many-body systems through the lens of autoregressive models

Juan Felipe Carrasquilla Álvarez
Vector Institute
@carrasqu

PSI Machine Learning for Many-Body Physics course
May 3rd, 2022

ML and quantum science research

- Condensed matter, quantum information, statistical physics, and atomic, molecular, and optical physics communities are exploring research at the intersection of ML and quantum physics.
- Interest is shaped in part by the **commonalities** in the structure of the problems that these disciplines address.


Commonalities

Dimensionality of quantum systems vs neural machine translation

$|\Psi\rangle$ vector with 2^N

- Today's best supercomputers can solve Schroedinger's equation **exactly** for systems with a maximum of ~45 spins.

$$2^N \sim 3.5 \times 10^{13}$$

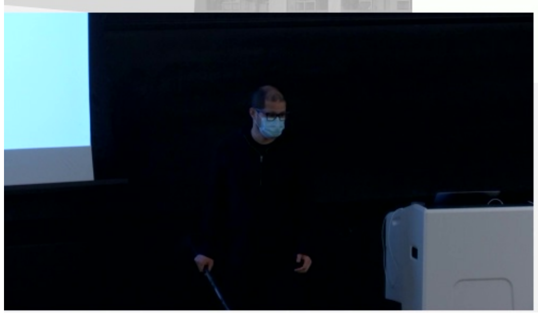
- We want N as large as possible

- Language models live in very high dimensional spaces too (example from "Attention is all you need", 2017)

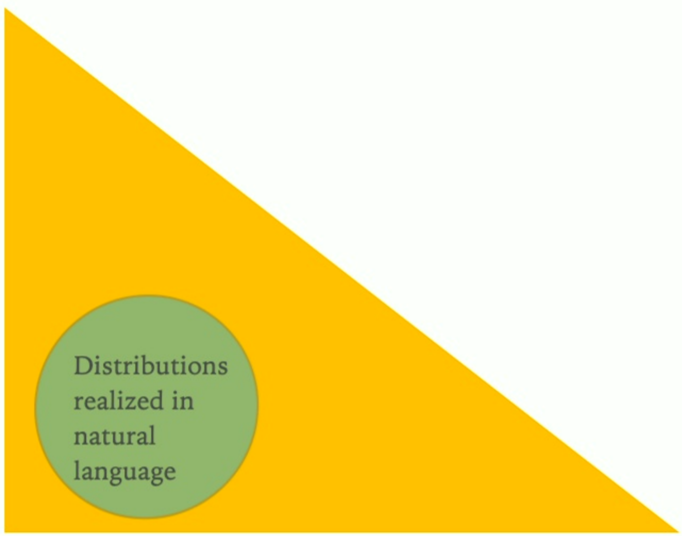
$$\text{Vocab. Size}^{\text{Max length of sentence}}$$

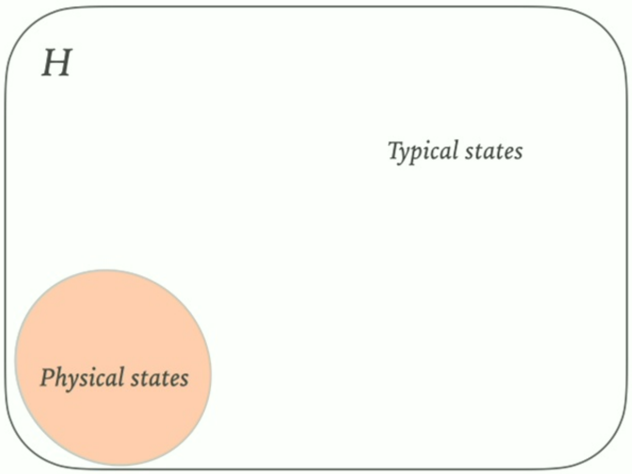
$$8000^{100} \sim 2.03 \times 10^{390}$$

Very large state space



Most paragraphs are noise — Probability distributions over the sentences our brain understands live in low-dimensional subspace. Similarly, physical states realized in nature live in a lower dimensional subspace.

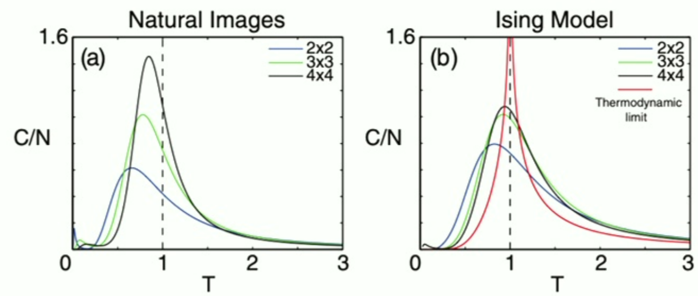




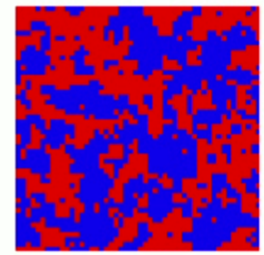
Most importantly is the common structure these problems share

Correlations and symmetries with strikingly similar structure

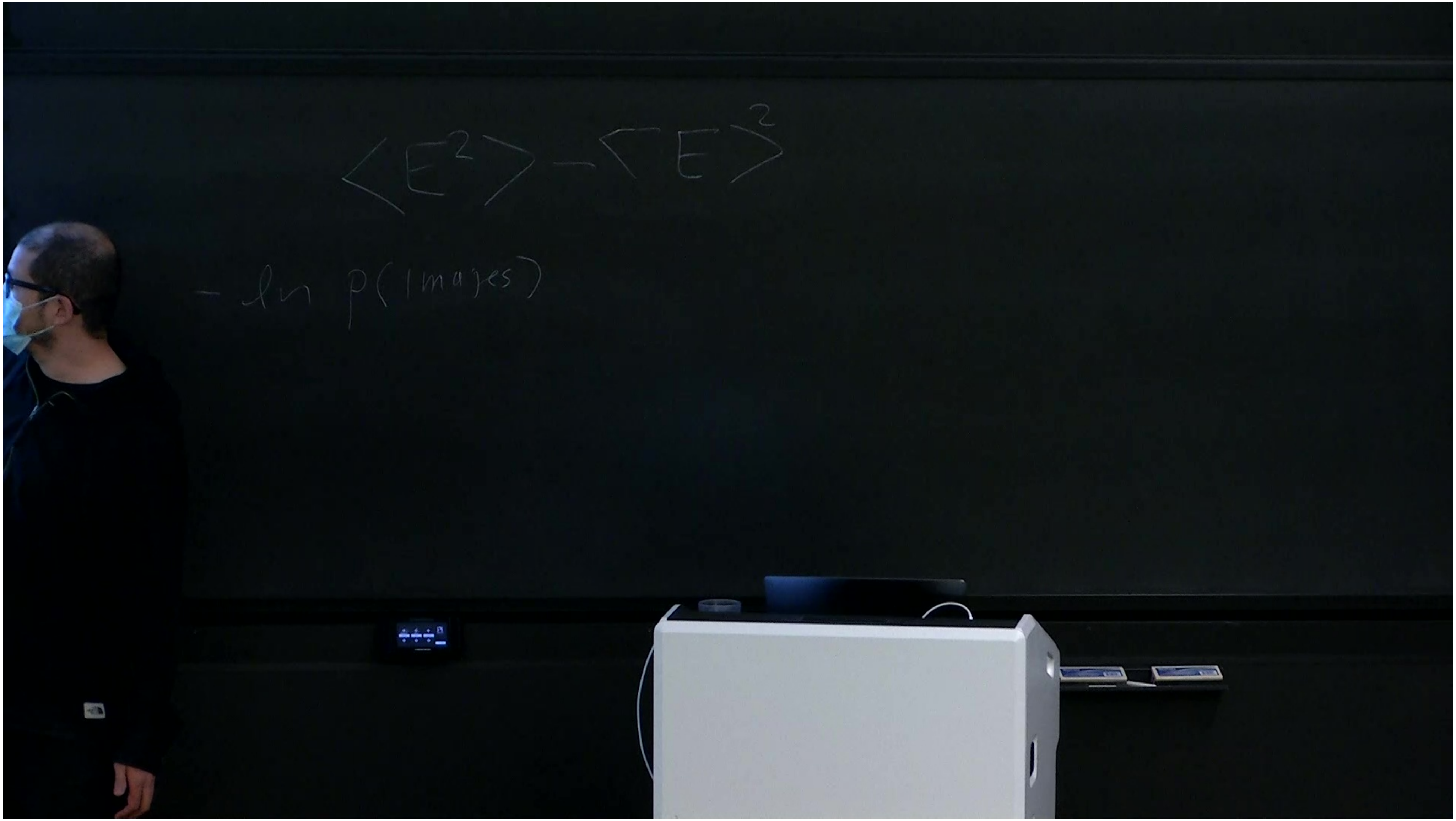
- Critical correlations:
- Natural language and natural images
- Music
- All exhibit power-law decaying correlations identical to a (classical or quantum) at a critical point
- Translational, rotational, reflection, and other symmetries— enrich out understanding and improve sample complexity in ML.

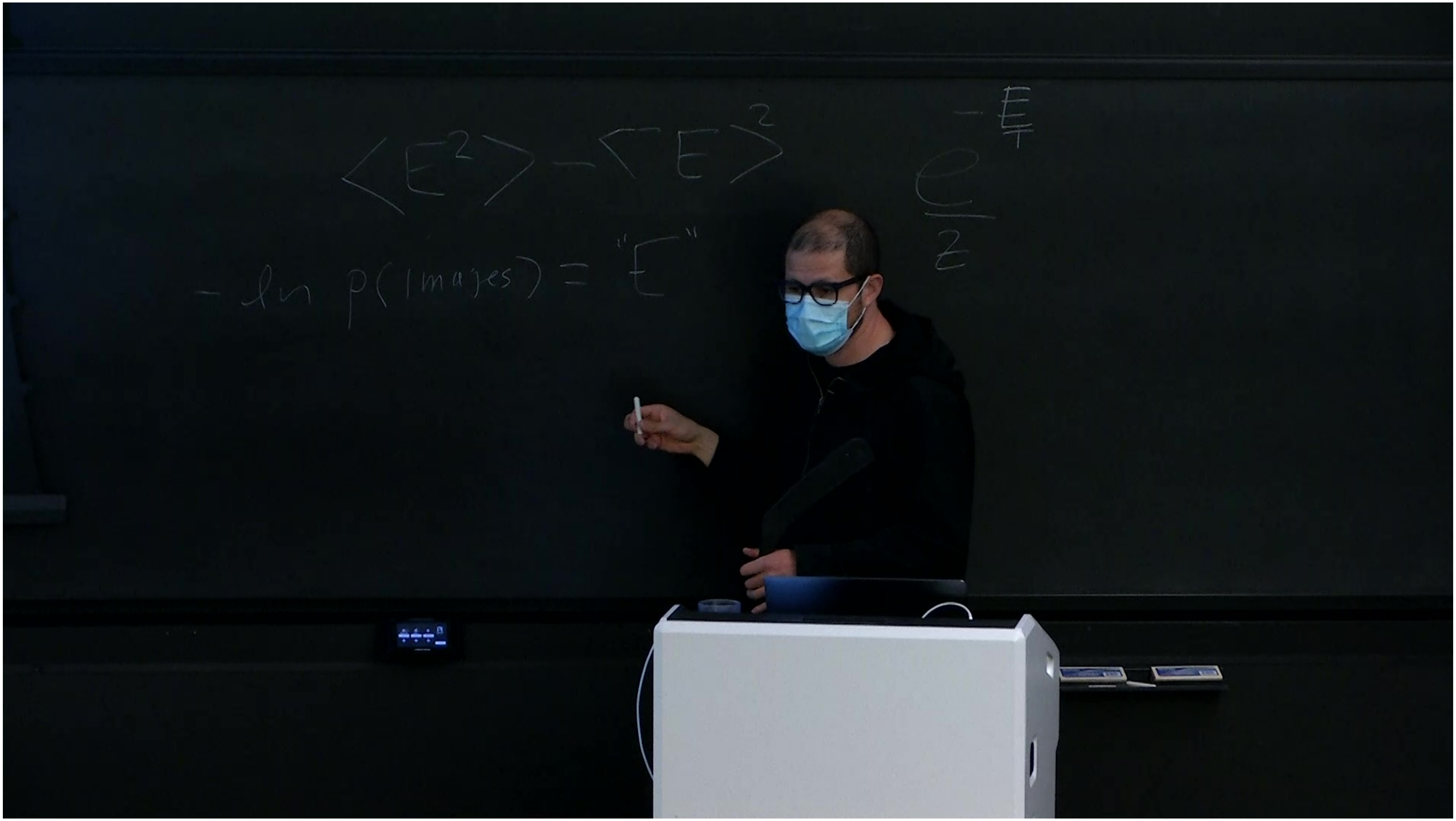


Statistical Thermodynamics of Natural Images
PRL 110, 018701 (2013)

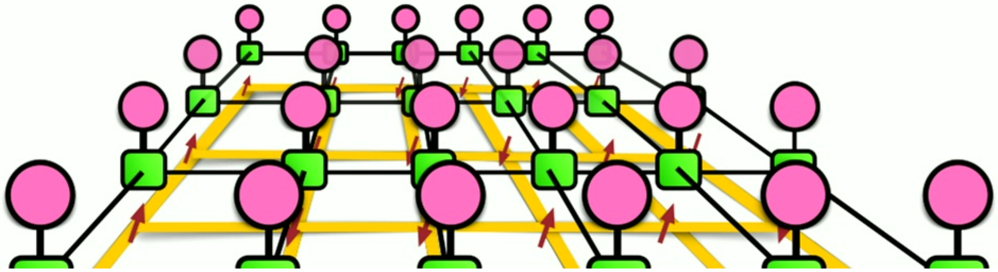


Scale-free correlations in starling flocks.
PNAS 107 (26) 11865-11870


$\langle E^2 \rangle - \langle E \rangle^2$
$- \ln p(\text{images})$


$\langle E^2 \rangle - \langle E \rangle^2$
$\frac{e^{-\frac{E}{T}}}{Z}$
$-\ln p(\text{images}) =$ "E"

WHAT ARE AUTOREGRESSIVE MODELS?



PROBABILISTIC AUTOREGRESSIVE MODELS

- The term *autoregressive* originates from time-series models: observations from the previous time-steps are used to predict the value at the current time step.
- Consider a probability distribution $P(\boldsymbol{\sigma}) = P(\sigma_1, \sigma_2, \ldots, \sigma_N)$,

$$P(\sigma_1, \sigma_2, \ldots, \sigma_N) = P(\sigma_1)P(\sigma_2|\sigma_1)P(\sigma_3|\sigma_1, \sigma_2)\ldots P(\sigma_N|\sigma_1, \sigma_2, \ldots, \sigma_{N-1})$$

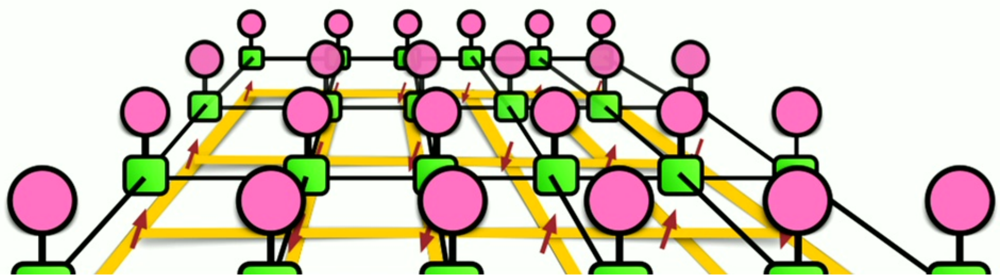
- To specify P in a tabular form requires **exponential** resources
- To alleviate this exponential issue: parametrize the conditionals

$$P(\sigma_i|\sigma_{<i}) = P_\theta(\sigma_i|\sigma_{<i})$$

PROBABILISTIC AUTOREGRESSIVE MODELS AND WAVE FUNCTIONS

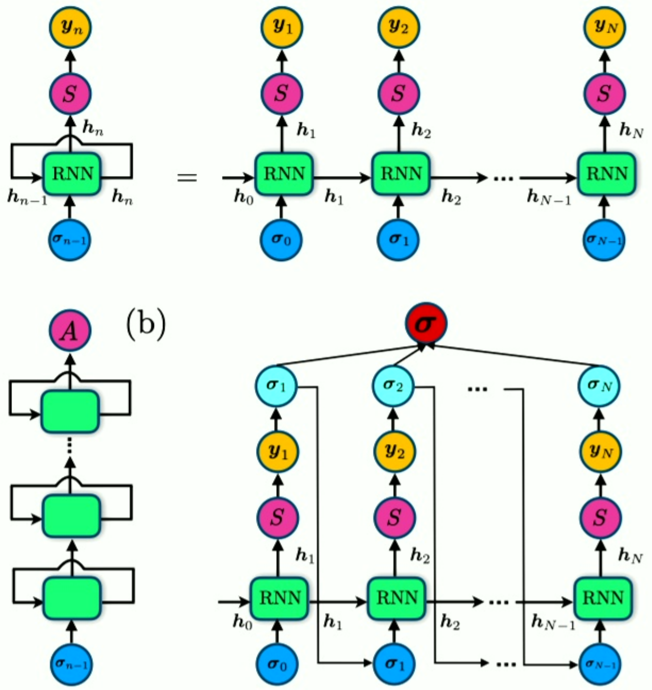
- ✓ Can be exactly sampled easily
- ✓ Computing the probability of a configuration $P(\boldsymbol{\sigma}) = P(\sigma_1, \sigma_2, \ldots, \sigma_N)$ is easy
- ✓ Can be defined in any spatial dimension
- ✓ Easy to encode mean-field theories (e.g. Gutzwiller mean-field theory)
- ✓ We can impose symmetry and other useful physical properties
- ✓ These properties remain true for autoregressive models of the quantum state

A CANONICAL EXAMPLE: THE RECURRENT NEURAL NETWORK



RECURRENT NEURAL NETWORKS (RNN)

➤ The key building block of an RNN is a recurrent cell



$$\boldsymbol{h}_n = f\left(W[\boldsymbol{h}_{n-1}; \boldsymbol{\sigma}_{n-1}] + \boldsymbol{b}\right)$$

$$\boldsymbol{y}_n \equiv S\left(U\boldsymbol{h}_n + \mathbf{c}\right) \quad \text{S = Softmax}$$

$$P\left(\sigma_n | \sigma_{n-1}, \ldots, \sigma_1\right) = \boldsymbol{y}_n \cdot \boldsymbol{\sigma}_n$$

Sampling:

- Sample each conditional
- Input the sample to the next step

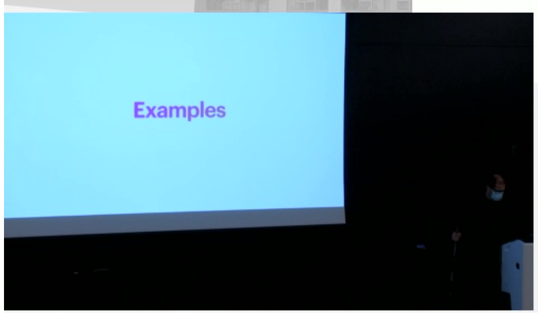
$$P(\sigma_1, \sigma_2, \ldots, \sigma_N) = P(\sigma_1)P(\sigma_2|\sigma_1)P(\sigma_3|\sigma_1, \sigma_2)\ldots P(\sigma_N|\sigma_1, \sigma_2, \ldots, \sigma_{N-1})$$

RNNs are universal function approximators. Schäfer and Zimmermann (2006)

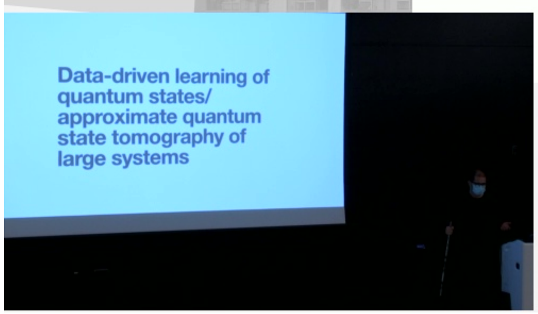
BUT THERE ARE MORE AND LIST IS LONG

- Transformers
- Neural autoregressive density estimators
- Autoregressive flows
- PixelRNN
- PixelCNN
- Wavenet
- ….

Examples



Data-driven learning of quantum states/ approximate quantum state tomography of large systems



Quantum state tomography

Quantum state tomography is the process of reconstructing the quantum state by **measurements** on the system. It "is the gold standard for verification and benchmarking of quantum devices"*

Useful for:

- Characterizing optical signals
- Diagnosing and detecting errors in state preparation, e.g. states produced by quantum computers reliably.
- Entanglement verification

* Efficient quantum state tomography. Marcus Cramer, Martin B. Plenio, Steven T. Flammia, Rolando Somma, David Gross, Stephen D. Bartlett, Olivier Landon-Cardinal, David Poulin & Yi-Kai Liu. Nature Communications volume 1, Article number: 149

Traditionally QST requires exponential resources

Examples

- Maximum likelihood estimation. Requires an explicit "physical" density matrix → representation scales poorly

$$L(\hat{\rho}) = \prod_a P(a)^{f_a}$$

Maximize probability of observed data with respect to a parametrization of $\hat{\rho}$

- Issues: Exponential scaling in the parametrization
- Estimation of errors due to finite statistics in the measurements is difficult

On the Measurement of Qubits. Daniel F. V. James, Paul G. Kwiat, William J. Munro, Andrew G. White. Physical Review A 64, 052312 (2001)

Need to go beyond standard Quantum state tomography

- Progress in controlling large quantum systems.
- Availability of arbitrary measurements performed with great accuracy.
- The bottleneck limiting progress in the estimation of states: curse of dimensionality of traditional techniques.

Synthetic Quantum devices are growing fast

Article | Published: 07 July 2021

Quantum phases of matter on a 256-atom programmable quantum simulator

Sepehr Ebadi, Tout T. Wang, Harry Levine, Alexander Keesling, Giulia Semeghini, Ahmed Omran, Dolev Bluvstein, Rhine Samajdar, Hannes Pichler, Wen Wei Ho, Soonwon Choi, Subir Sachdev, Markus Greiner, Vladan Vuletić & Mikhail D. Lukin

Nature **595**, 227–232 (2021) | Cite this article

Science

Current Issue First release papers Archive About Submit manuscript

HOME > SCIENCE > VOL. 374, NO. 6572 > PROBING TOPOLOGICAL SPIN LIQUIDS ON A PROGRAMMABLE QUANTUM SIMULATOR

RESEARCH ARTICLE TOPOLOGICAL MATTER

Probing topological spin liquids on a programmable quantum simulator

G. SEMEGHINI, H. LEVINE, A. KEESLING, S. EBADI, T. T. WANG, D. BLUVSTEIN, R. VERRESEN, H. PICHLER, M. KALINOWSKI, [...] M. D. LUKIN

+7 authors Authors Info & Affiliations

nature.com > nature > letters > article

nature
International journal of science

Letter | Published: 22 August 2018

Observation of topological phenomena in a programmable lattice of 1,800 qubits

Andrew D. King, Juan Carrasquilla, [...] Mohammad H. Amin

Nature **560**, 456–460 (2018) | Download Citation

PHYSICAL REVIEW X

Highlights Recent Subjects Accepted Collections Authors Referees Search Press About Staff

Open Access

Quantum Chemistry Calculations on a Trapped-Ion Quantum Simulator

Cornelius Hempel, Christine Maier, Jonathan Romero, Jarrod McClean, Thomas Monz, Heng Shen, Petar Jurcevic, Ben P. Lanyon, Peter Love, Ryan Babbush, Alán Aspuru-Guzik, Rainer Blatt, and Christian F. Roos

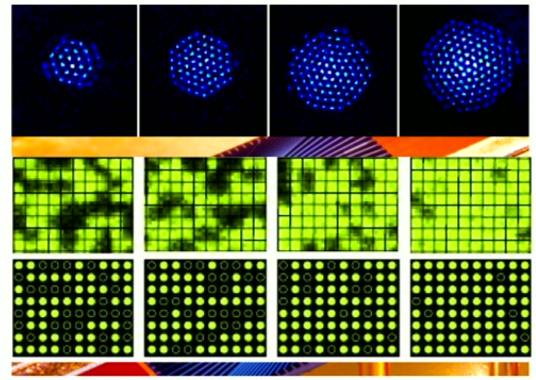
Phys. Rev. X **8**, 031022 – Published 24 July 2018

1. How to collect the datasets

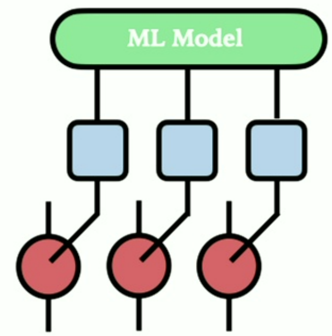
$$\text{dataset} = \begin{bmatrix} 0,0,1,\ldots 1,0 \\ 1,1,0,\ldots 0,1 \\ 0,0,1,\ldots 0,1 \\ 1,1,0,\ldots 0,0 \\ \vdots \\ 0,1,0,\ldots 1,0 \end{bmatrix}$$

Quantum state tomography

- Prepare an unknown quantum state
- Apply a measurement that probes enough information about the quantum state
- Repeat and collect the statistics of the measurement
- Infer a reconstruction of the state consistent with the measurement outcomes



$\rho^* =$

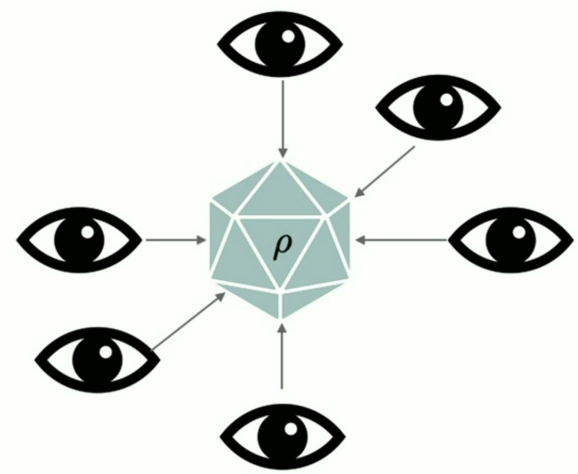


Measurements: Positive operator-valued measure (POVM)

- Born Rule $P\left(\mathbf{a}\right) = \mathrm{Tr}\, \rho M^{\mathbf{a}}$ quantum theory ⇔ experiment
- Measurement operators $\mathbf{M} = \{M^{(a)} \mid a \in \{1, ..., m\}\}$

INFORMATIONALLY COMPLETE MEASUREMENTS

- The measurement statistics $P(\mathbf{a})$ contains all of the information about the state.
- Relation between ρ and distribution $P(\boldsymbol{a})$ can be inverted
- M form a basis for operators



INVERTING BORN RULE

BORN RULE $P(\mathbf{a}) = \mathrm{Tr}\,\rho M^{\mathbf{a}}$

INFORMATIONALLY COMPLETENESS —>THIS RELATION CAN BE INVERTED

$$\rho = \sum_{a,a'} T^{-1}_{a,a'} P(a') M^{(a)}$$

$$T_{a,a'} = \mathrm{Tr}[M^{(a)} M^{(a')}]$$

Insight: parametrize statistics of measurements and invert

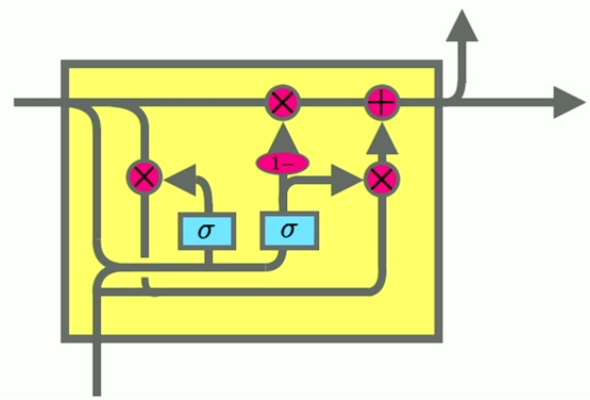
$$P\left(\mathbf{a}\right) = \mathrm{Tr}\,\rho M^{\mathbf{a}}$$

=> Create an autoregressive model of P(a)

Autoregressive models (RNNs and transformer)

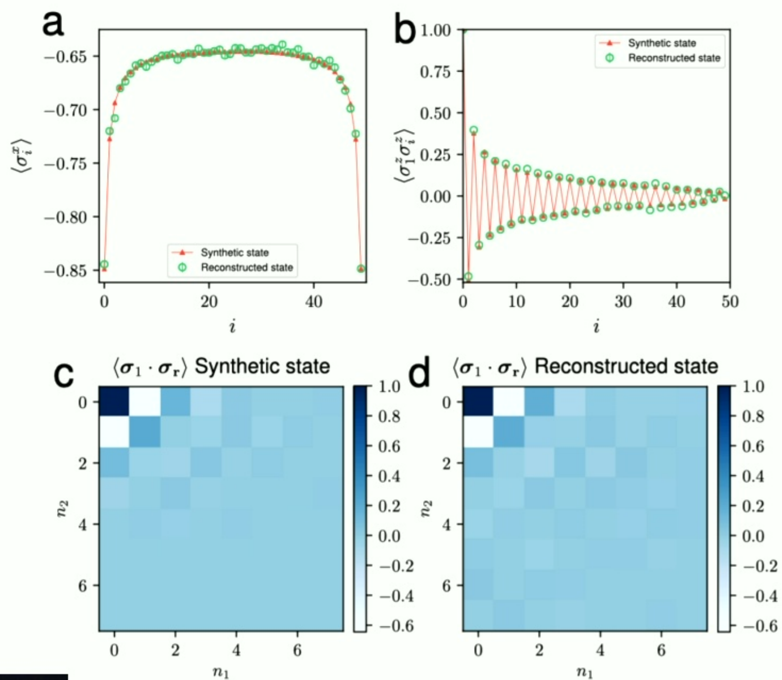
1. Allow for exact sampling
2. Tractable density $P_{\mathrm{model}}\left(\mathbf{a}\right)$
3. Use MLE to learn it.

$P_{\mathrm{model}}\left(\mathbf{a}\right) \longrightarrow$



$$\rho_{\mathrm{model}} = \sum_{a,a'} T^{-1}_{a,a'} P_{\mathrm{model}}(a')\, M^{(a)}$$

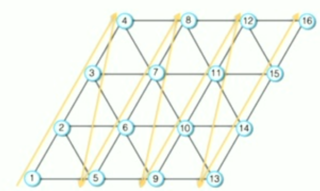
Learning Ground states of local hamiltonians from data



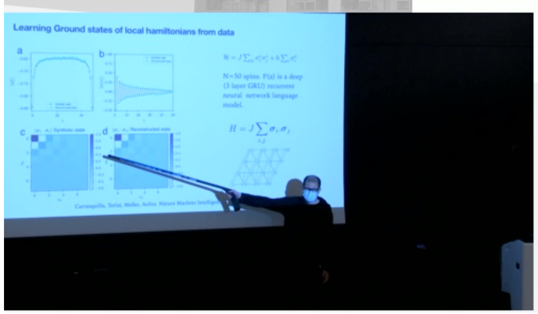
$\mathcal{H} = J \sum_{ij} \sigma_i^z \sigma_j^z + h \sum_i \sigma_i^x$

N=50 spins. P(a) is a deep (3 layer GRU) recurrent neural network language model.

$$H = J \sum_{i,j} \boldsymbol{\sigma}_i . \boldsymbol{\sigma}_j$$



Carrasquilla, Torlai, Melko, Aolita. Nature Machine Intelligence 1, 200 (2019)



EXPERIMENTAL DEMONSTRATION

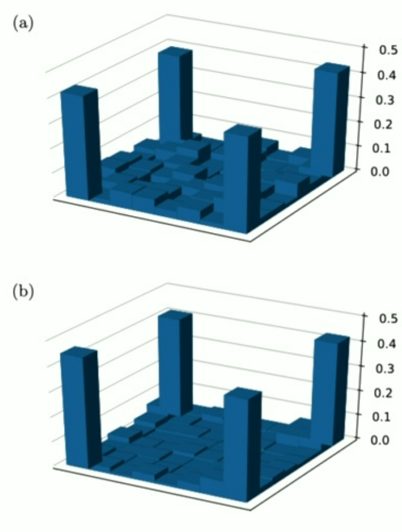
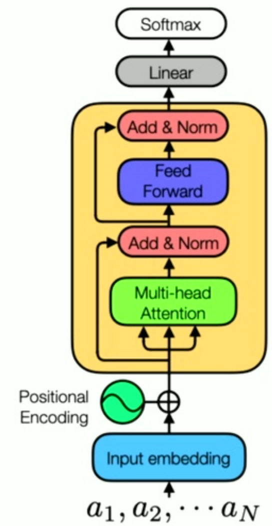


FIG. 5. Benchmarking AQT (a) to MLE tomography offered by IBM's Qiskit library (b) for a noisy 3-qubit QHZ state data generated on the IBMQ_OURENSE quantum computer. Each bar represents the absolute value of a density matrix (DM) element.

GHZ state with 3 qubits



PSI 2015-2016

Peter Cha, Paul Ginsparg, Felix Wu, Juan Carrasquilla, Peter L. McMahon, Eun-Ah Kim. arXiv:2006.12469 (2020)

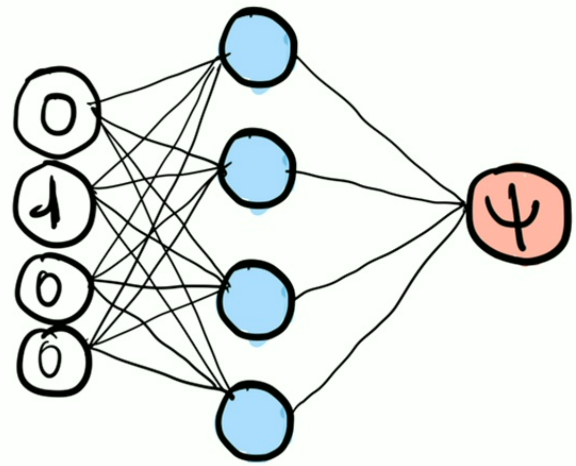
Neural networks as wavefunctions

- Recall that we represent a quantum state as a 2^N-dimensional vector of complex entries
- $$|\psi\rangle = \begin{bmatrix} \psi_{0,0,0,\ldots 0,0} \\ \psi_{1,0,0,\ldots 0,0} \\ \psi_{1,1,0,\ldots 0,0} \\ \vdots \\ \psi_{1,1,1,\ldots 1,1} \end{bmatrix}$$ What does it mean that we represent a quantum state as a neural network?

$$|\psi_\theta\rangle = \begin{bmatrix} \psi_\theta(0,0,0,\ldots 0,0) \\ \psi_\theta(1,0,0,\ldots 0,0) \\ \psi_\theta(1,1,0,\ldots 0,0) \\ \vdots \\ \psi_\theta(1,1,1,\ldots 1,1) \end{bmatrix}$$

Where the complex-valued boolean function

$\psi_\theta(x_1, x_2, \ldots, x_N) =$ **Neural network**$(x_1, x_2, \ldots, x_N)$



As a consequence, we go from an exponential amount of parameters to a neural network with a few parameters at the cost of constraining the type of functions we can represent.

Neural network quantum states

ELSEVIER

Computer Physics Communications 104 (1997) 1–14

Computer Physics Communications

Artificial neural network methods in quantum mechanics

I.E. Lagaris [1], A. Likas, D.I. Fotiadis

Department of Computer Science, University of Ioannina, P.O. Box 1186, GR 45110 Ioannina, Greece

Received 17 March 1997; revised 22 April 1997

3.5. Two-dimensional Schrödinger equation

We consider here the well-studied [2] example of the Henon–Heiles potential. The Hamiltonian is written as

$$H = -\frac{1}{2}\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right) + V(x, y)\,,$$

with $V(x, y) = \frac{1}{2}(x^2 + y^2) + \frac{1}{4\sqrt{5}}(xy^2 - \frac{1}{3}x^3)$.

I.E. Lagaris et al. / Computer Physi

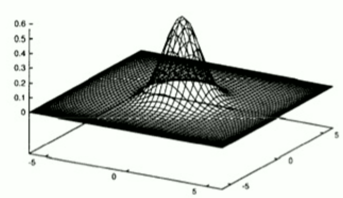


Fig. 4.

Fig. 4. Ground state of the Henon–Heiles problem (ϵ = 0.99866).

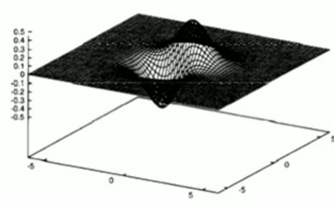
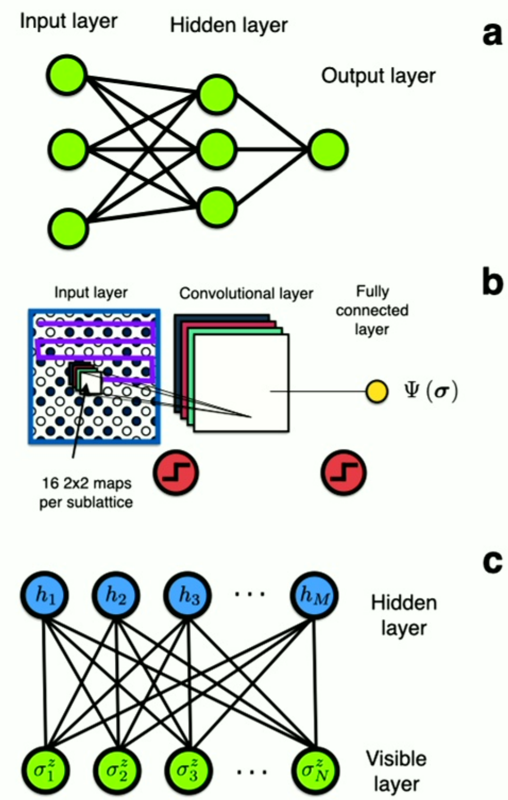


Fig. 5.



Neural network quantum states

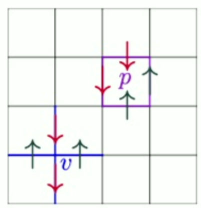
nature physics
LETTERS
PUBLISHED ONLINE: 13 FEBRUARY 2017 | DOI: 10.1038/NPHYS4035

Machine learning phases of matter

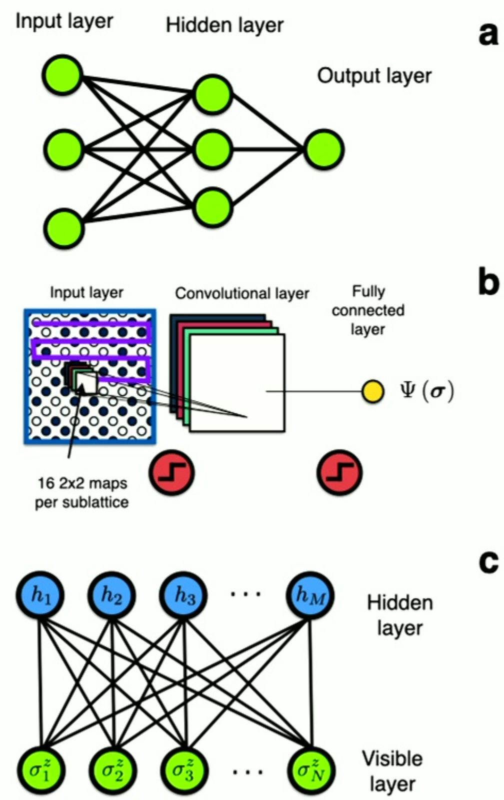
Juan Carrasquilla[1]* and Roger G. Melko[1,2]

KITAEV'S TORIC CODE GROUND STATE

$$H = -J_p \sum_p \prod_{i \in p} \sigma_i^z - J_v \sum_v \prod_{i \in v} \sigma_i^x$$



$$|\Psi_{\mathrm{TC}}\rangle \propto \lim_{\beta \to \infty} \sum_{\sigma_1, \ldots, \sigma_N} e^{\frac{\beta}{2} J \sum_p \prod_{i \in p} \sigma_i^z} |\sigma_1, \ldots, \sigma_N\rangle$$



J. Carrasquilla and R. G. Melko. Nature Physics 13, 431–434 (2017)

Neural network quantum states

RESEARCH

RESEARCH ARTICLE

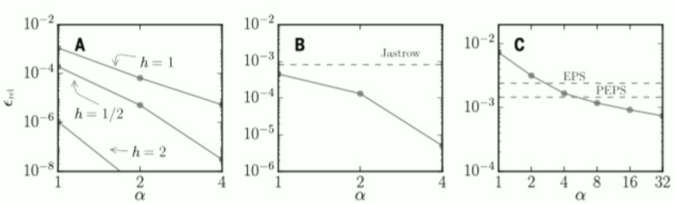
MANY-BODY PHYSICS

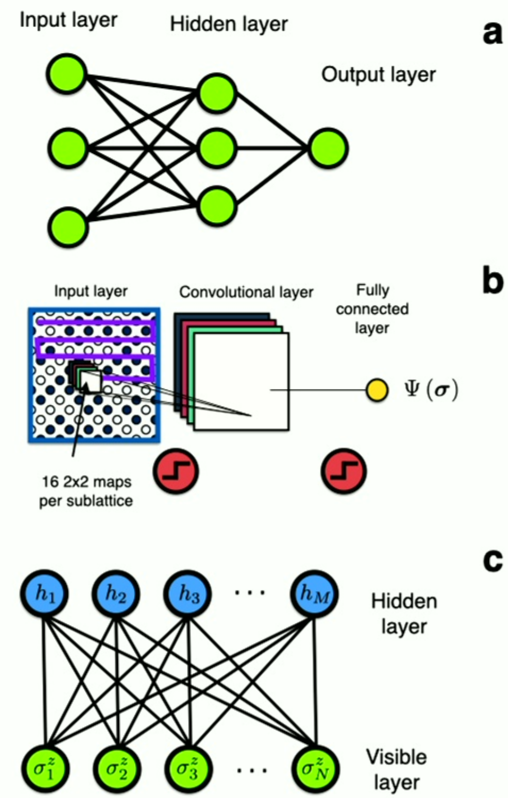
Solving the quantum many-body problem with artificial neural networks

Giuseppe Carleo[1*] and Matthias Troyer[1,2]

The challenge posed by the many-body problem in quantum physics originates from the difficulty of describing the nontrivial correlations encoded in the exponential complexity of the many-body wave function. Here we demonstrate that systematic machine learning of the wave function can reduce this complexity to a tractable computational form for some notable cases of physical interest. We introduce a variational representation of quantum states based on artificial neural networks with a variable number of hidden neurons. A reinforcement-learning scheme we demonstrate is capable of both finding the ground state and describing the unitary time evolution of complex interacting quantum systems. Our approach achieves high accuracy in describing prototypical interacting spins models in one and two dimensions.

$$\mathcal{H}_{\rm TFI} = -h\sum_i \sigma_i^x - \sum_{ij} \sigma_i^z\sigma_j^z \qquad \mathcal{H}_{\rm AFH} = \sum_{ij} \sigma_i^x\sigma_j^x + \sigma_i^y\sigma_j^y + \sigma_i^z\sigma_j^z$$



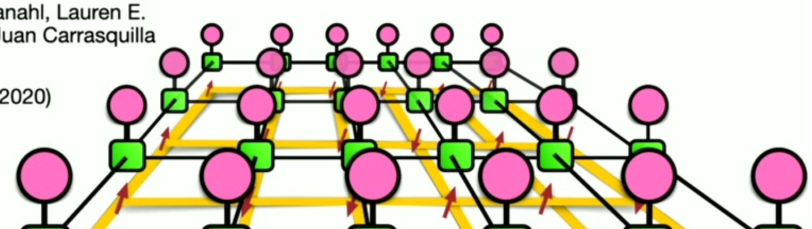


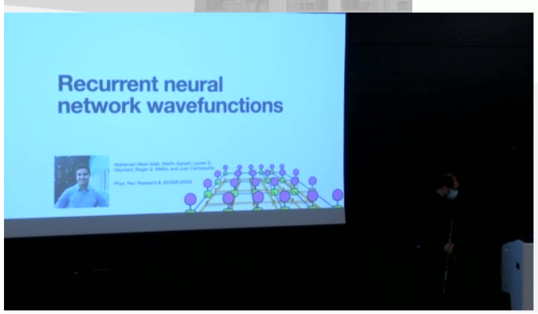
Giuseppe Carleo, Matthias Troyer, Science 355, 602 (2017)

Recurrent neural network wavefunctions

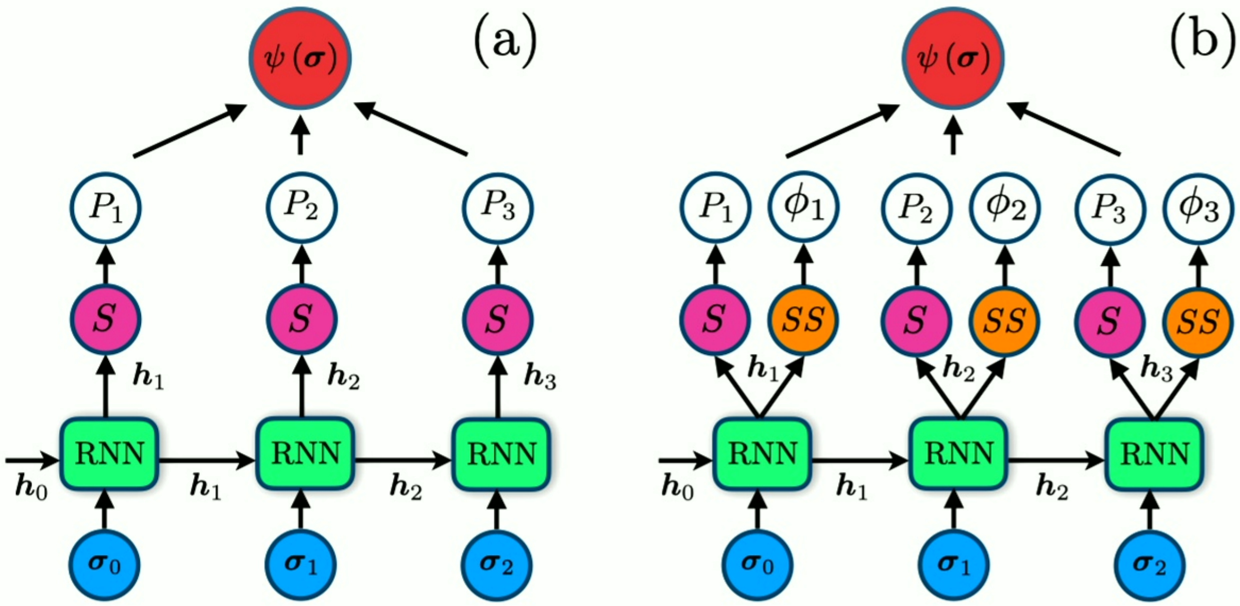
Mohamed Hibat-Allah, Martin Ganahl, Lauren E. Hayward, Roger G. Melko, and Juan Carrasquilla

Phys. Rev. Research **2**, 023358 (2020)





RNNs wave functions

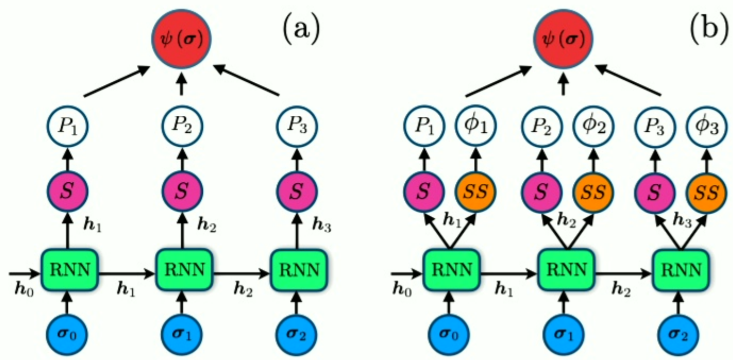


$$|\Psi\rangle = \sum_{\boldsymbol{\sigma}} \psi(\boldsymbol{\sigma})\,|\boldsymbol{\sigma}\rangle = \sum_{\boldsymbol{\sigma}} \sqrt{P(\boldsymbol{\sigma})}\,|\boldsymbol{\sigma}\rangle\,. \qquad |\Psi\rangle = \sum_{\boldsymbol{\sigma}} e^{\mathrm{i}\phi(\boldsymbol{\sigma})}\sqrt{P(\boldsymbol{\sigma})}\,|\boldsymbol{\sigma}\rangle\,.$$

$\boldsymbol{\sigma}$ corresponds to a bit string, eg = (0101010111110100101010001)

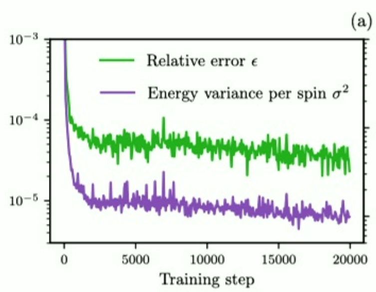
$$\phi(\boldsymbol{\sigma}) = \sum_{n=1}^{N} \phi_n$$

Recurrent neural network wavefunctions

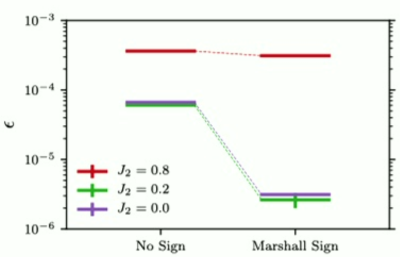


$$|\Psi\rangle = \sum_{\sigma} \psi(\sigma)\,|\sigma\rangle = \sum_{\sigma} \sqrt{P(\sigma)}\,|\sigma\rangle$$

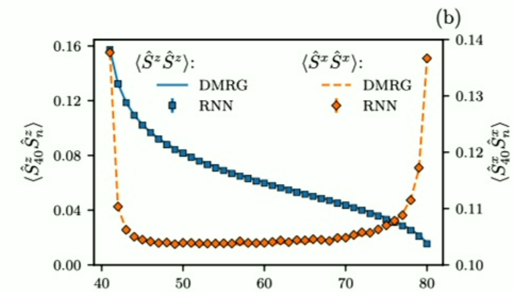
$$|\Psi\rangle = \sum_{\sigma} \exp(\mathrm{i}\phi(\sigma))\sqrt{P(\sigma)}\,|\sigma\rangle$$



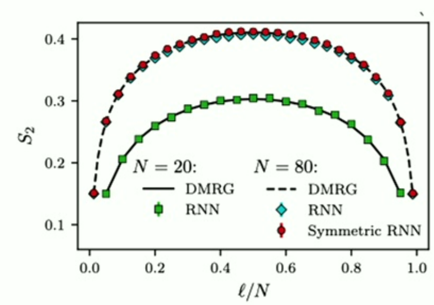
TFIM 1-D N=1000



J1-J2 model in 1-D



TFIM 1-D N=80



TFIM 1-D N=80

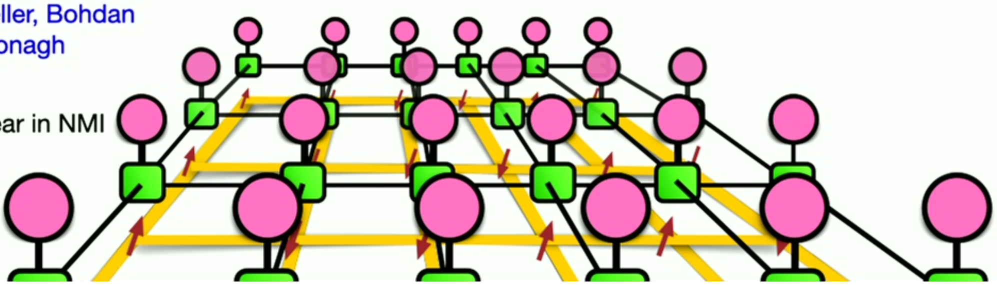
Symmetries: Spin inversion, mirror reflection, Sz. Sign: different Marshall signs for the J1-J2 model

Mohamed Hibat-Allah, Martin Ganahl, Lauren E. Hayward, Roger G. Melko, and Juan Carrasquilla Phys. Rev. Research **2**, 023358 (2020)

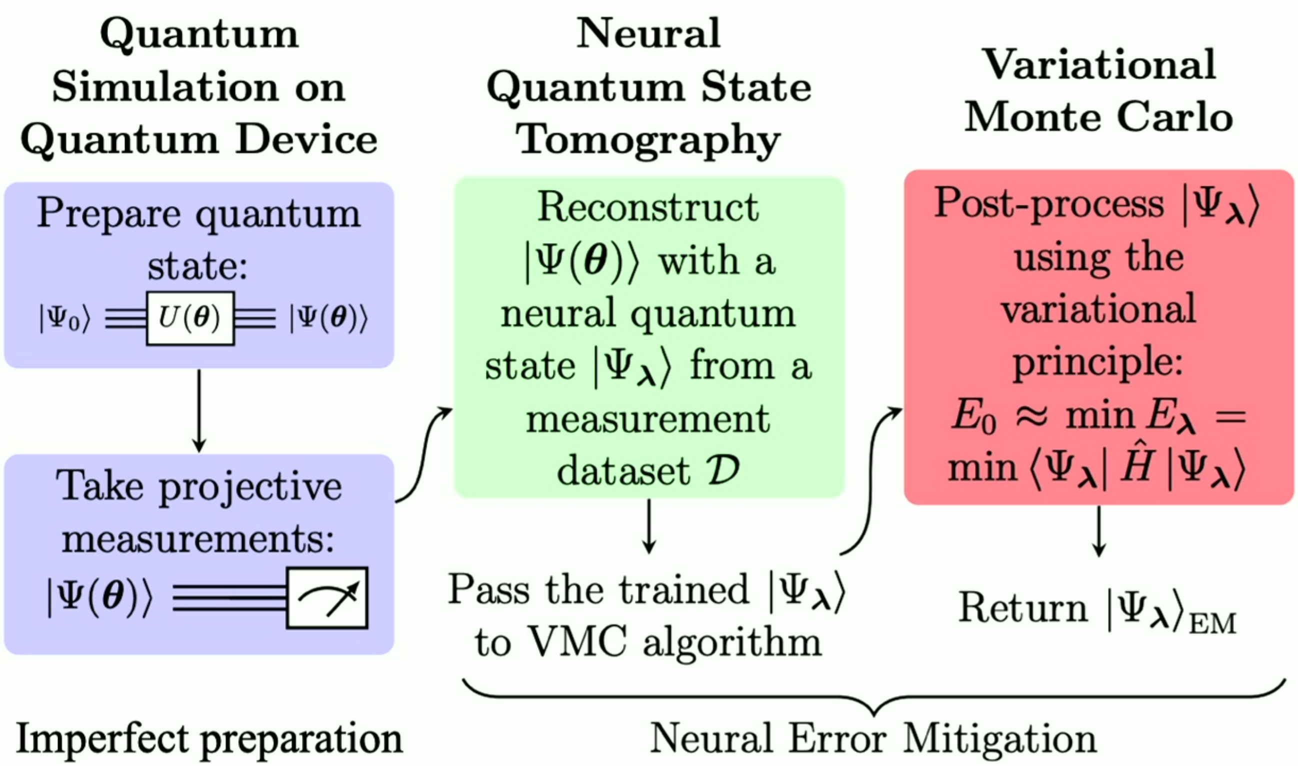
Neural error mitigation of near term quantum simulations

Elizabeth R. Bennewitz, Florian Hopfmueller, Bohdan Kulchytskyy, Juan Carrasquilla, Pooya Ronagh

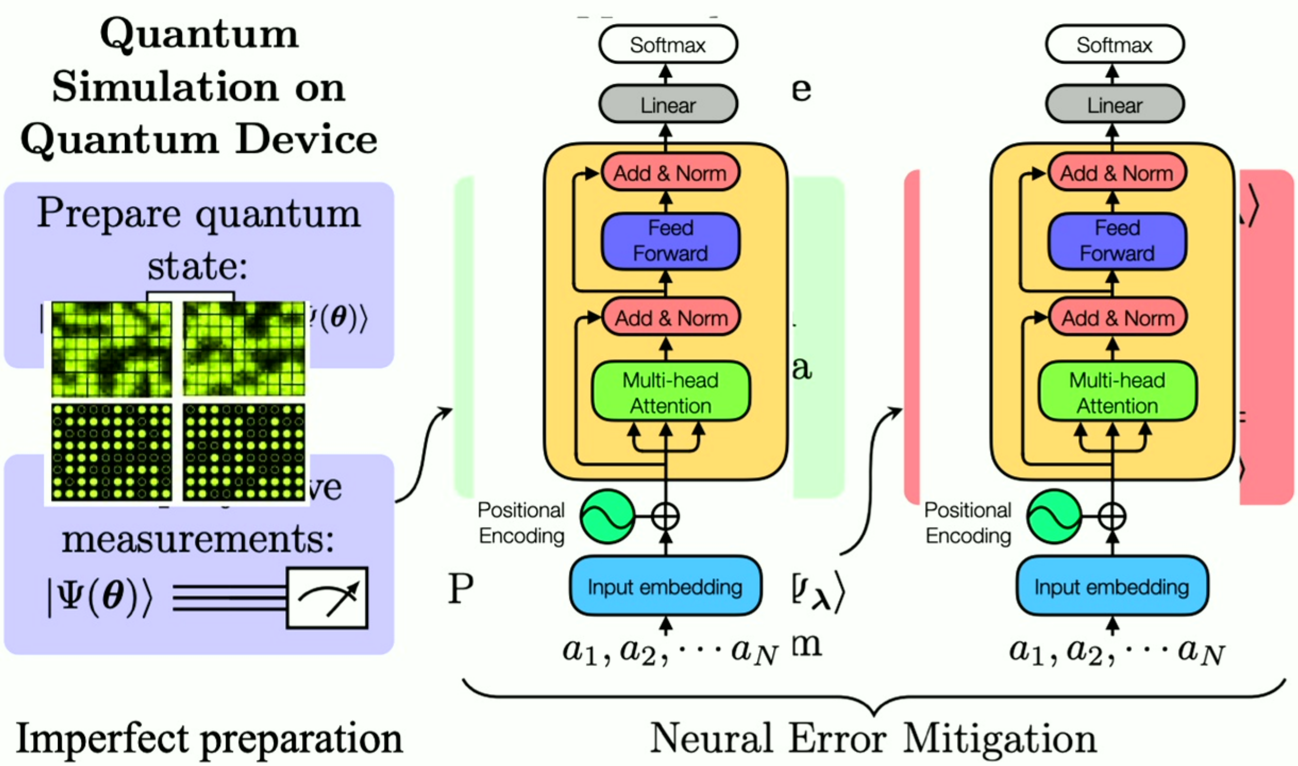
https://arxiv.org/abs/2105.08086 to appear in NMI



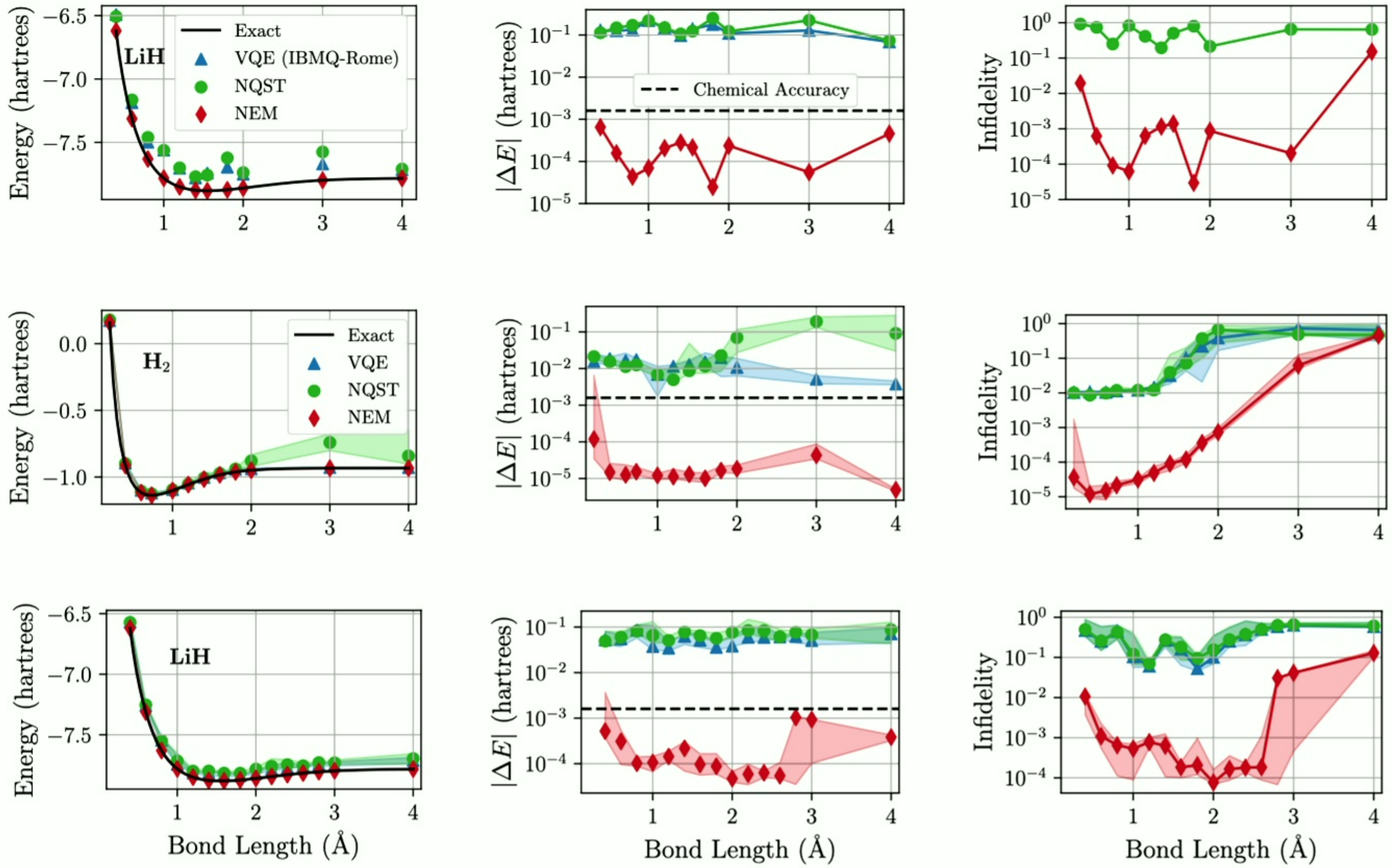
Neural error mitigation



Neural error mitigation



Neural error mitigation



Neural error mitigation

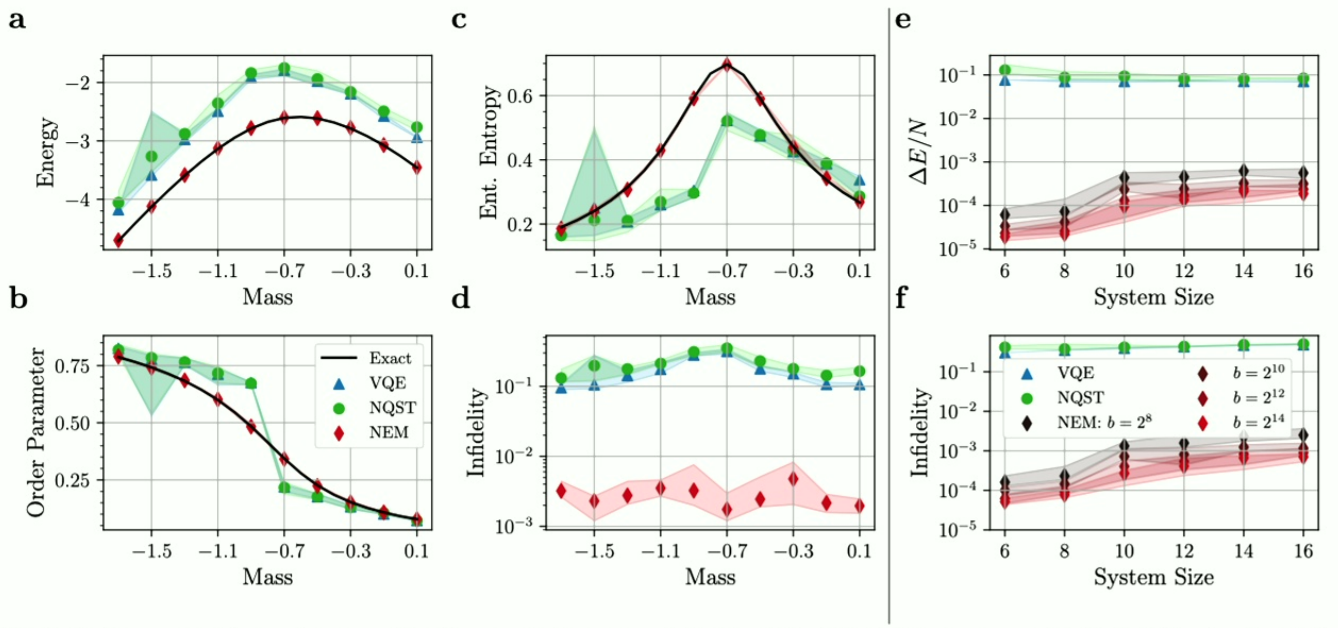
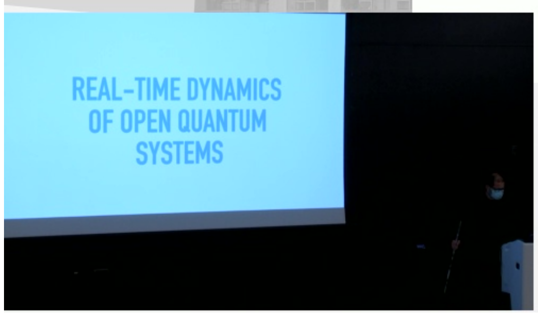


FIG. 3: **Performance of neural error mitigation applied to ground states of the lattice Schwinger model** | ***Left***: Estimates for **(a)** the ground state energy, **(b)** order parameter, **(c)** entanglement entropy between the first three and remaining five sites, and **(d)** infidelity to the exact ground state, for the $N = 8$ site model. Each panel contains results for the quantum states prepared using VQE simulated with a depolarizing noise channel (blue triangles), neural quantum states trained using neural quantum state tomography (green circles), the final neural error mitigated neural quantum states (red diamonds), and, where applicable, exact results (solid black lines). While the qualitative behaviour of the entanglement entropy and the order parameter across the phase transition are not modelled well by VQE or NQST, applying NEM consistently improves the estimates of all observables to low errors and low infidelities. ***Right***: Results of the scaling study of NEM at the phase transition ($m = -0.7$) are shown for **(e)** the energy error per site and **(f)** infidelity, as a function of system size. The performance of VQE without noise (blue triangles) shows an approximately constant energy error per site with infidelities that become slightly worse as system size increases. Across all sizes, applying NEM (red diamonds) improves VQE performance by two to four orders of magnitude, even when using a small VMC batch size of $b = 2^8$, which is the number of samples used to estimate the energy's gradient in one iteration of VMC. In all panels (left and right), median values over 10 runs are shown with the shaded region encompassing three values on either side of the median.

REAL-TIME DYNAMICS OF OPEN QUANTUM SYSTEMS



NEURAL AUTOREGRESSIVE MODELS FOR MANY-BODY PHYSICS

- Quantum state reconstruction with RNNs (Nature Machine Intelligence, vol. 1, 155-161 (2019)) and Transformers (arXiv:2006.12469)
- Simulation of quantum circuits with transformers (arXiv:1912.11052)
- Recurrent neural network wavefunctions — accurate ground states, very compact representation in 1d and 2d (Phys. Rev. Research 2, 023358 (2020))
- Variational neural annealing: produces very accurate solutions to challenging spin glass problems beyond SA and SQA (arXiv:2101.10154)
- Neural Error Mitigation of Near-Term Quantum Simulations (arXiv:2105.08086) (transformer)
- Simulation of open system dynamics (arXiv:2009.05580) (transformer)
- Transfer learning based on physical principles (arXiv:2003.02647) (RNN)
- Data-Enhanced Variational Monte Carlo for Rydberg Atom Arrays (RNN) (arXiv:2203.04988)
- U(1) symmetric recurrent neural networks for quantum state reconstruction (RNN) (arXiv:2010.14514)

Conclusions

- Now is a privileged time for quantum research — enormous opportunities arising from artificial intelligence and quantum computing, two of today's most promising computational paradigms of the century.
- Body of recent showcases the opportunities that machine learning techniques, ideas, and **research culture** can spark in the field of quantum physics.
- What's a plausible goal for the near term?
- Can ML lead to scientific breakthroughs in quantum physics?

Can ML produce scientific breakthroughs in quantum physics?