Self-learning projective quantum Monte Carlo simulations guided by neural network states

Sebastiano Pilati, University of Camerino

Workshop "*Many-Body quantum physics with machine learning*"
ECT*, Trento, 7 Sept. 2023

COLLABORATORS:
L. Brodoloni, S. Cantori, G. Scriva, E. Costa (UniCam)
B. McNaughton, D. Vitali, A . Perali (UniCam), P. Pieri (UniBo)
B. Juliá Díaz, P. Mujal (IFISC, UIB-CSIC)
R. Fazio (ICTP, UniNA)
E. M. Inack (Perimeter Institute)

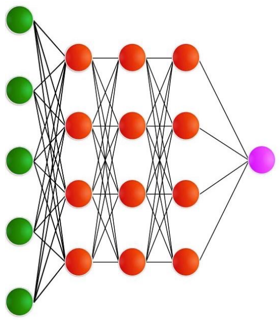
UNICAM Università di Camerino 1336

INFN Istituto Nazionale di Fisica Nucleare

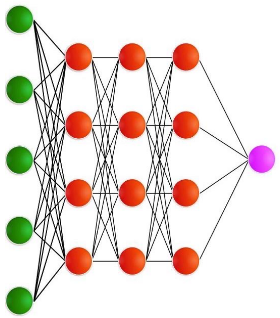
NQSTI QST

ISCRA CINECA

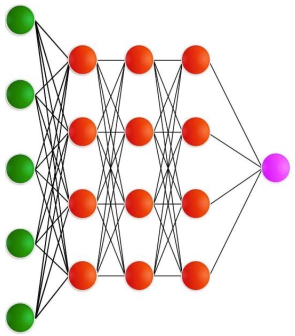
Deep learning for the quantum many body problem

Neural network states: $\boldsymbol{R}$ →  $\psi(\boldsymbol{R})$

$$E_0 = \min\left\{\frac{\langle\psi|H|\psi\rangle}{\langle\psi|\psi\rangle}\right\}$$

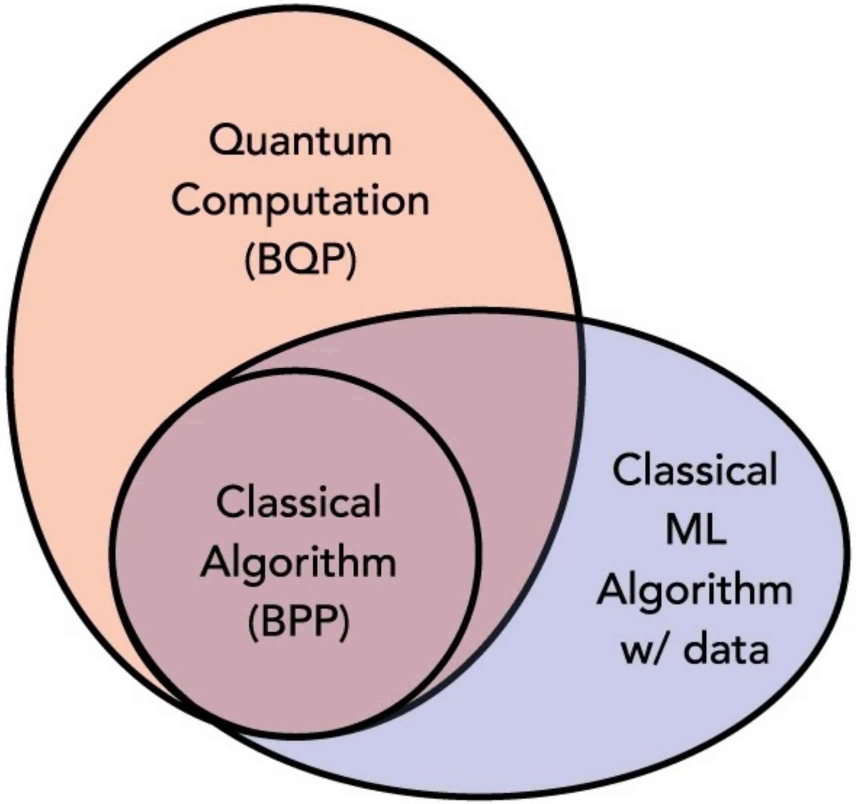
Density functional theory (needs data!) $n(\boldsymbol{R})$ →  $E[n(\boldsymbol{R})]$

$$E_0 \leq E[n(\boldsymbol{R})]$$
$$E_0 = \min\{E[n(\boldsymbol{R})]\}$$

Supervised learning (needs data!) $\hat{H}$ →  E_0

Data might come from:

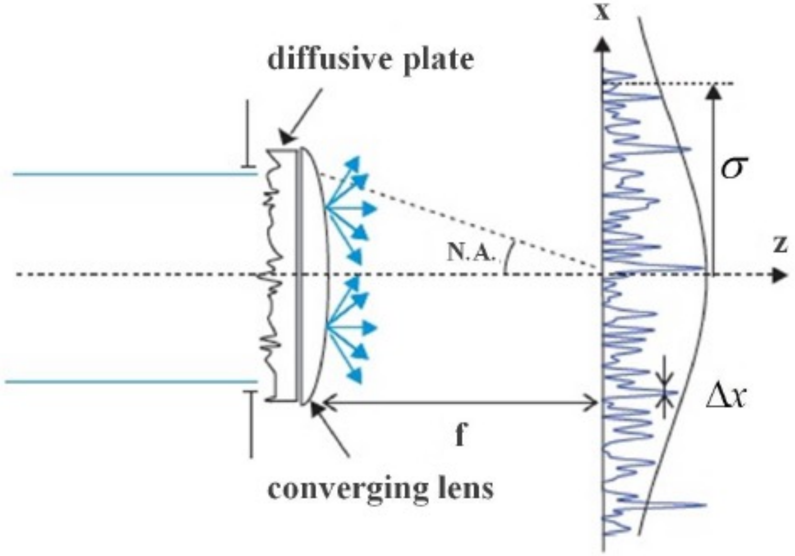
- More ab-initio method (NQS): computational cost
- Smaller systems: scalability
- Quantum device: noise & errors



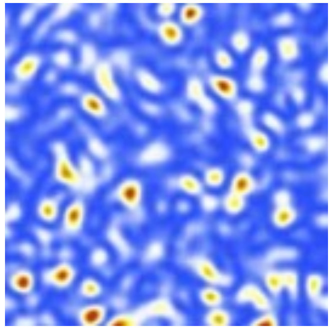
Classical machine learning algorithm trained on (quantum) data can solve otherwise classically intractable problems.

H.-Y. Huang et al, *Power of data in quantum machine learning,* Nat. Commun. **12**, 2631 (2021).
H.-Y. Huang et al., *Provably efficient machine learning for quantum many-body problems,* Science (2022)

Supervised learning: cold-atoms in optical speckle patterns

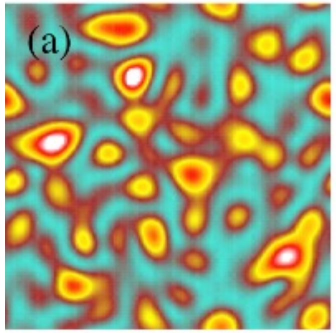


experiment @ LENS



Semeghini et al. Nat. Phys. 2015

num. simulation

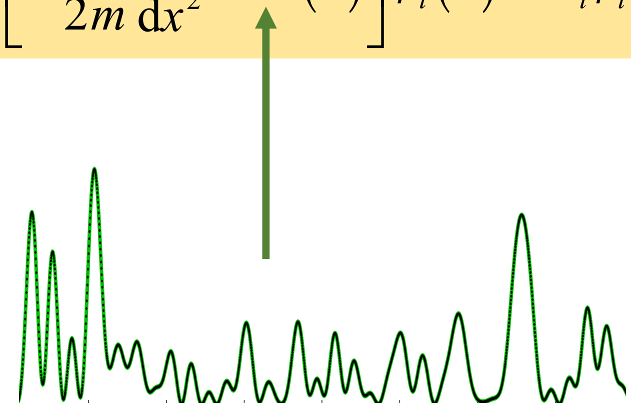


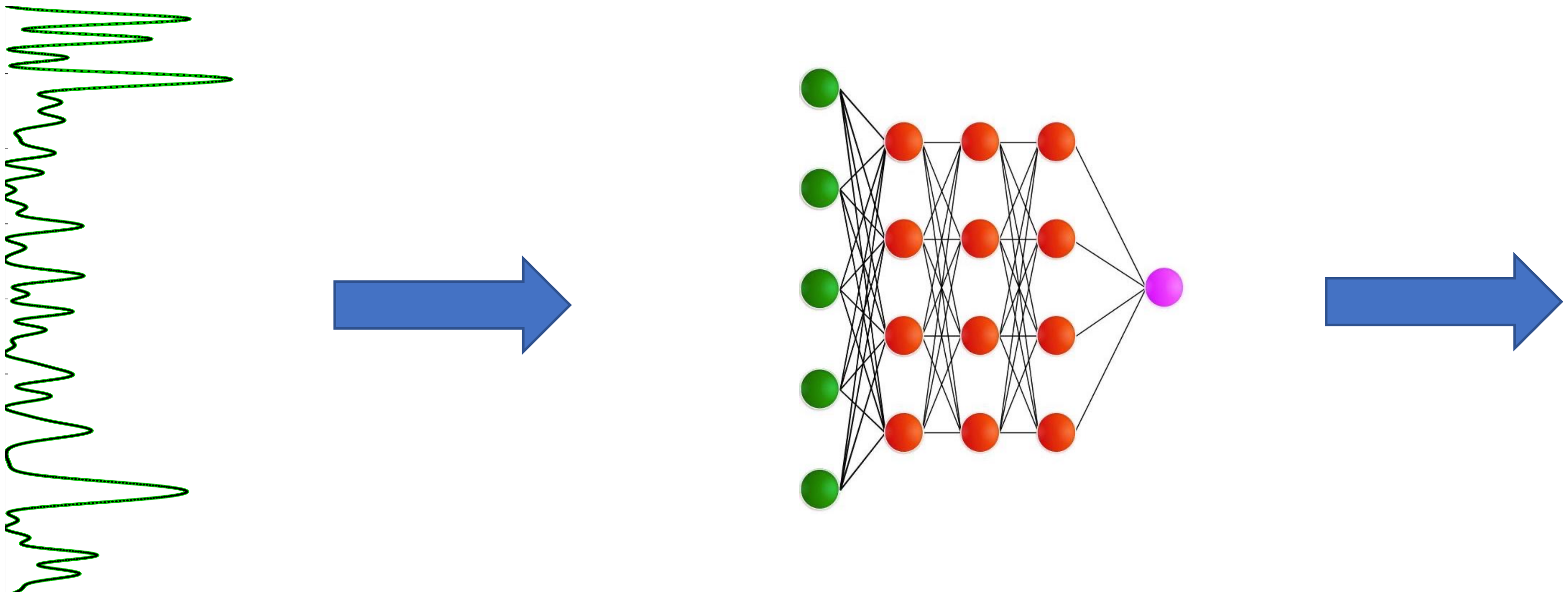
SP, Fratini PRA (2015)

EXP @ LENS, Palaiseau

1D single-particle SE:

$$\left[-\frac{\hbar^2}{2m}\frac{\mathrm{d}^2}{\mathrm{d}x^2}+V\left(x\right)\right]\varphi_i\left(x\right)=E_i\varphi_i\left(x\right)$$

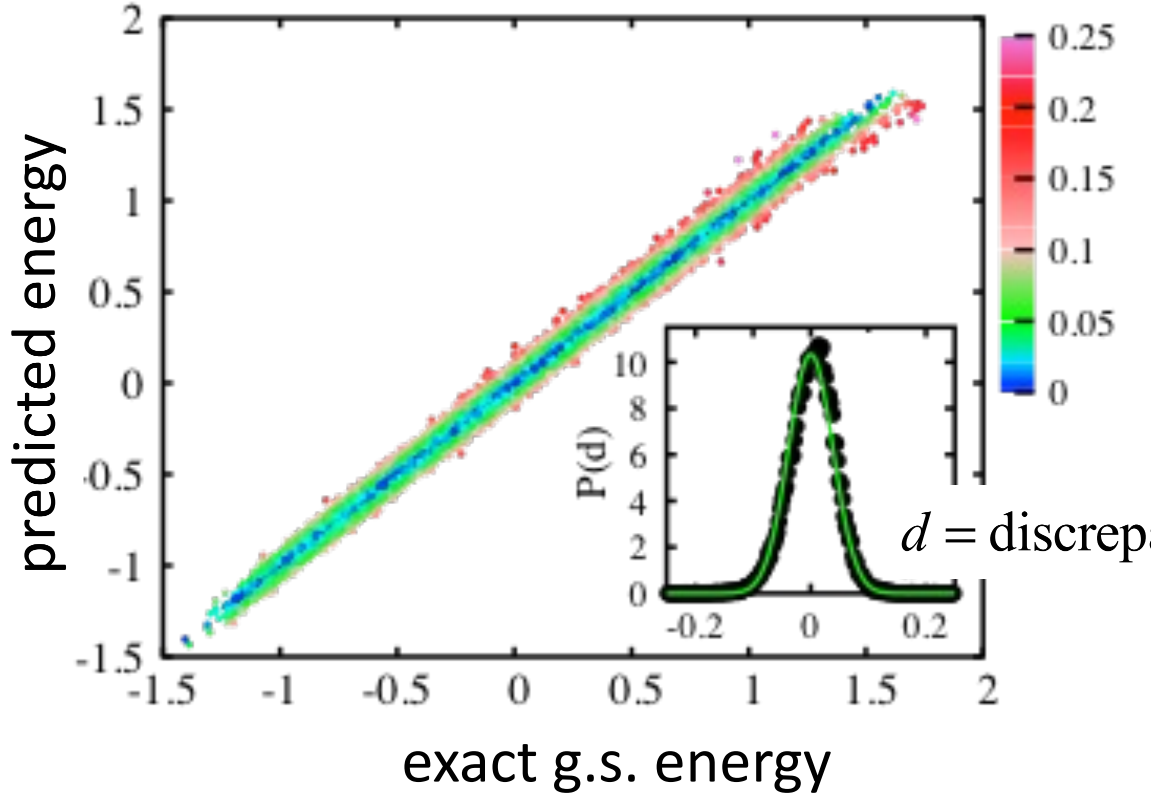




$$E_0 = \langle \psi_0 | \hat{H} | \psi_0 \rangle$$

Neural net. predictions vs. exact g.s. energy

SP, Pieri, Sci. Rep. **9**, 5613 (2019)



$N_{\text{train}} = 8\times10^4$

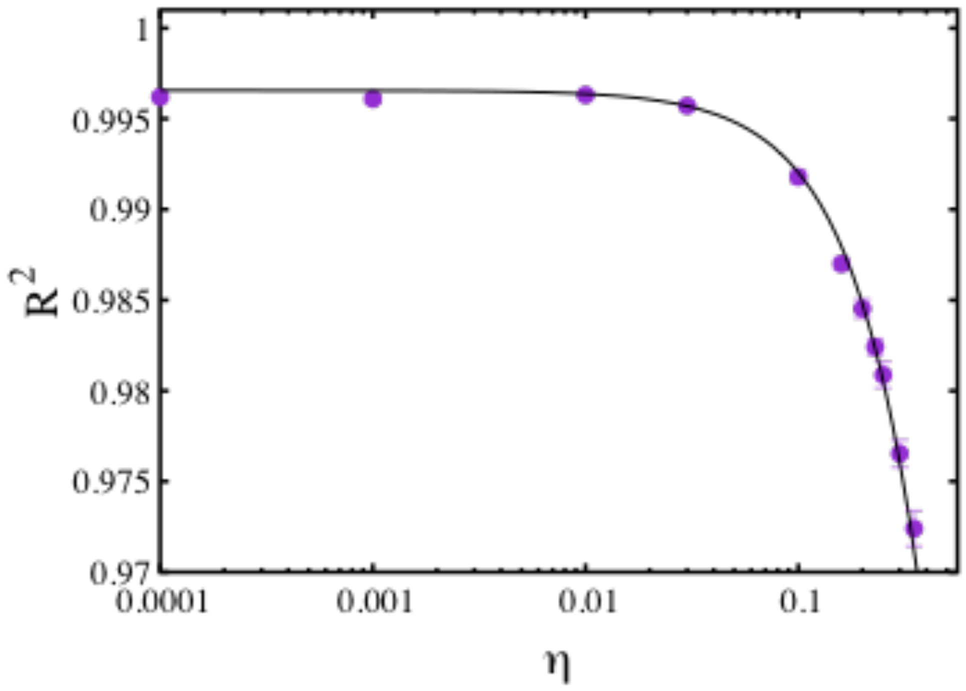
$$d = \text{discrepancy} \equiv \varepsilon^{\text{pred}} - \varepsilon_0$$

Resilinece to noise: cold-atom quantum simulators?

SP, Pieri, Sci. Rep. **9**, 5613 (2019)

Q: could we use QS to train NN to solve computationally intractable problem?

ANALYSIS OF NOISE SENSITIVITY: Training on synthetic data with added noise



added noise

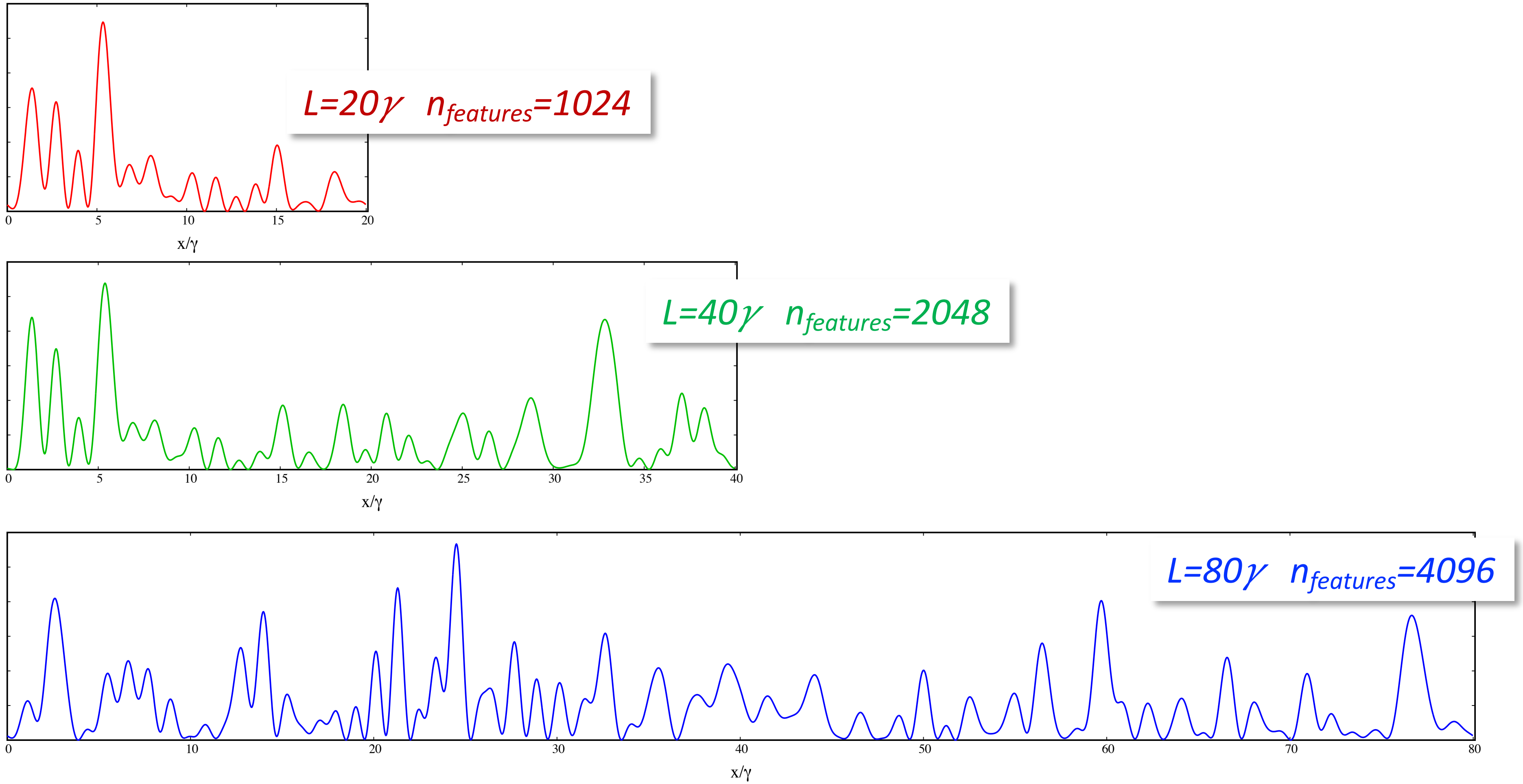
The NN is remarkably resilient, it can filter signal from noise

$$\sigma_{\text{noise}} = \eta\sigma_{data}$$

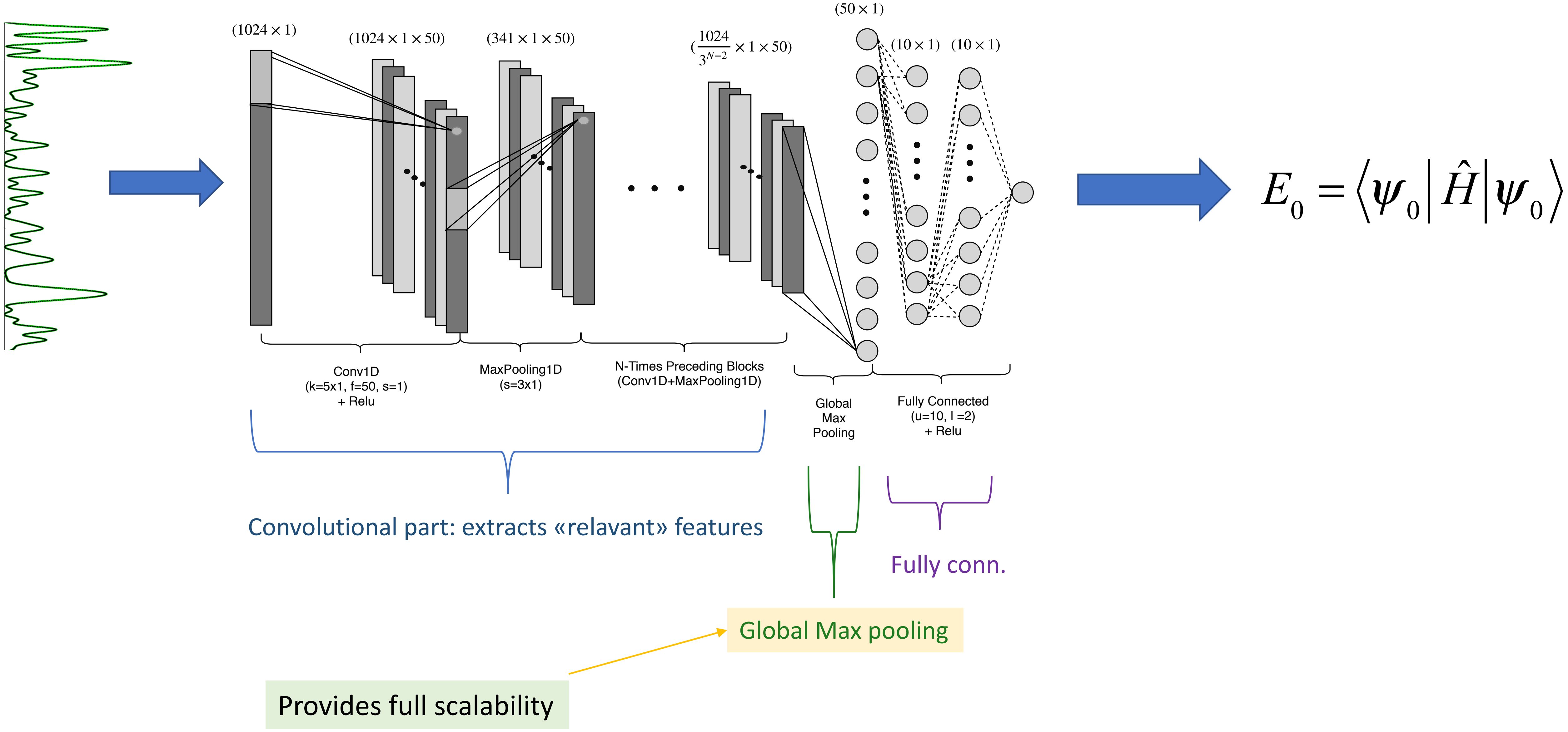
Gaussian noise with stand. dev. proportional to intrinsic data stand. dev.

Scalabale neural networks

- Scalable NN that can address different system sizes
- Heterogeneous training
- Extrapolation: making prediction for larger system sizes

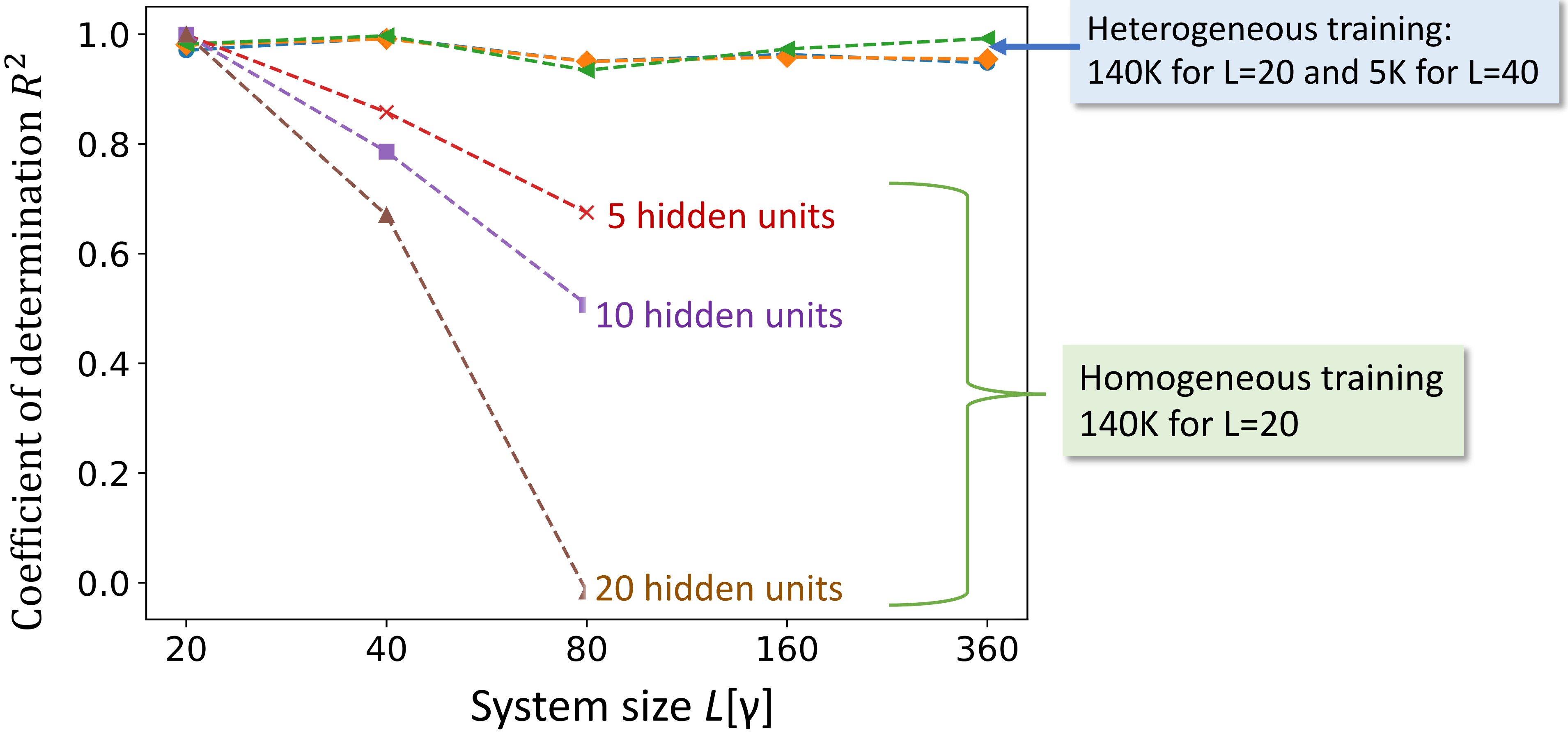


A fully-scalable convolutional neural network



Extrapolation: making predictions for larger systems

Coefficient of determination: $R^2 = 1 - \frac{\text{MSE}}{\text{VAR}}$



Saraceni, Cantori, SP, PRE **102**, 033301 (2020)

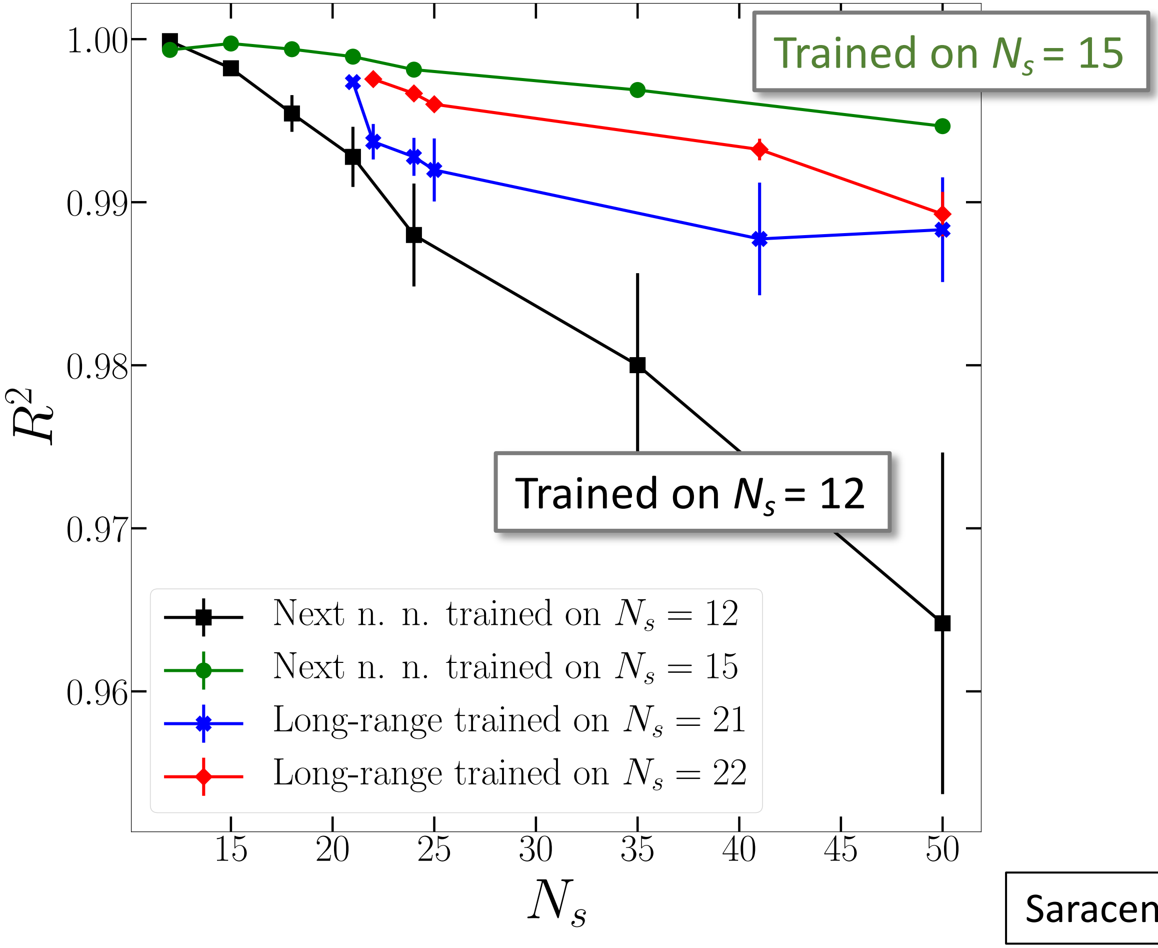
Disordered quantum Ising model in 1D

$$H = -\sum_i \left[J_{i,i+1}\sigma_i^x\sigma_{i+1}^x + J_{i,i+d}\sigma_i^x\sigma_{i+d}^x\right] - \Gamma\sum_i \sigma_i^z$$

$$J_{i,i+1}, J_{i,i+d} \sim \mathrm{Unif}[0,1]$$

Descriptors: $J_{i,i+1}, J_{i,i+d}$

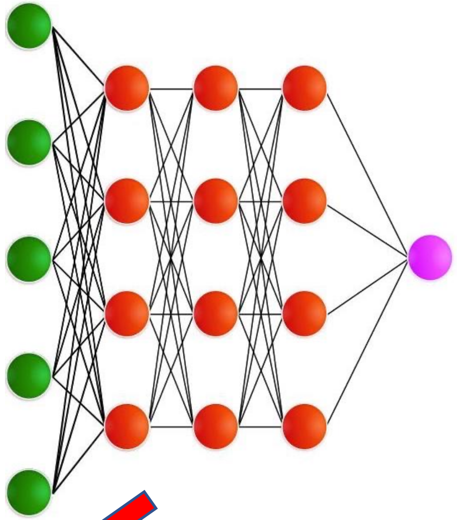
N_{train}: 3×10^4



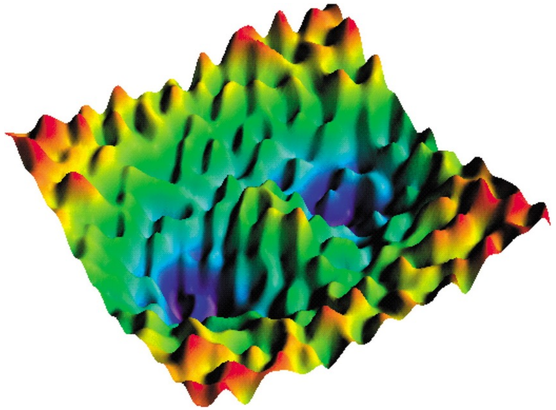
Saraceni, Cantori, SP, PRE **102**, 033301 (2020)

Quantum annealers

Classical deep learning

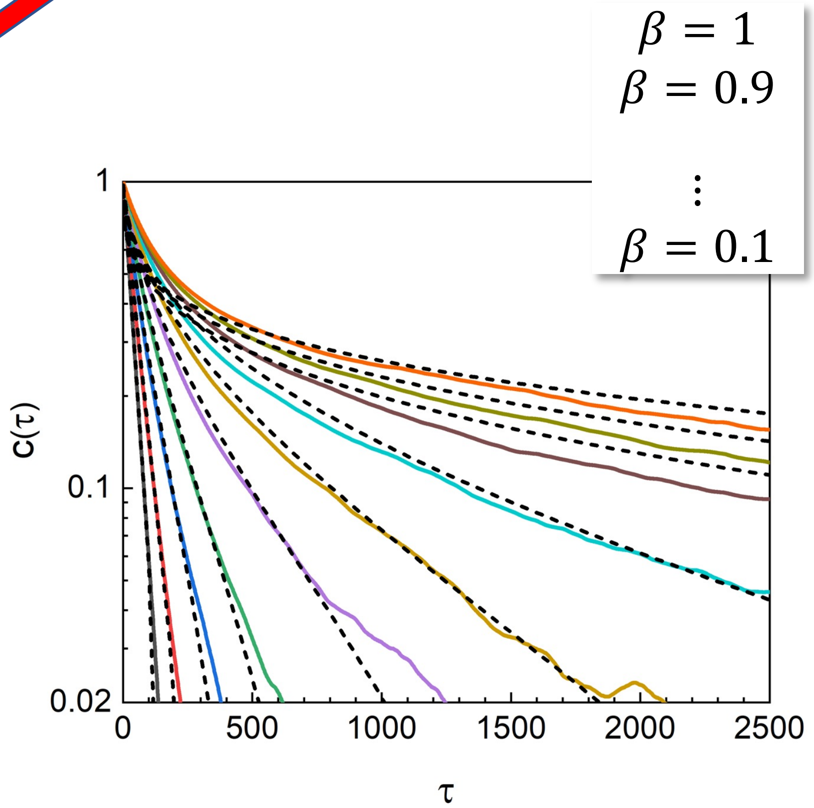


Energy landscape



Low-T equilibrium properties of classical spin glasses

$$H = \sum_{ij} J_{ij} x_i x_j$$



NEURAL CLASSICAL MONTE CARLO SIMULATIONS

McNaughton, Milošević, Perali, SP, PRE (2020)

Proposal matrix in Metropolis-Hastings algorithm: $w_{\mathbf{x}'\mathbf{x}} = p_{\mathrm{NETWORK}}(\mathbf{x}')$ (use ancestral sampling)

Acceptance probability: $A_{\mathbf{x}'\mathbf{x}} = \min\left(1, \frac{p_{\mathrm{BOLTZ}}(\mathbf{x}')w_{\mathbf{x}\mathbf{x}'}}{p_{\mathrm{BOLTZ}}(\mathbf{x})w_{\mathbf{x}'\mathbf{x}}}\right)$

We need: $p_{\mathrm{BOLTZ}}(\mathbf{x}) > 0 \implies p_{\mathrm{NETWORK}}(\mathbf{x}) > 0$

if: $p_{\mathrm{NETWORK}}(\mathbf{x}) \cong p_{\mathrm{BOLTZ}}(\mathbf{x}) \longrightarrow A_{\mathbf{x}'\mathbf{x}} \cong 1$ → efficient simulation!

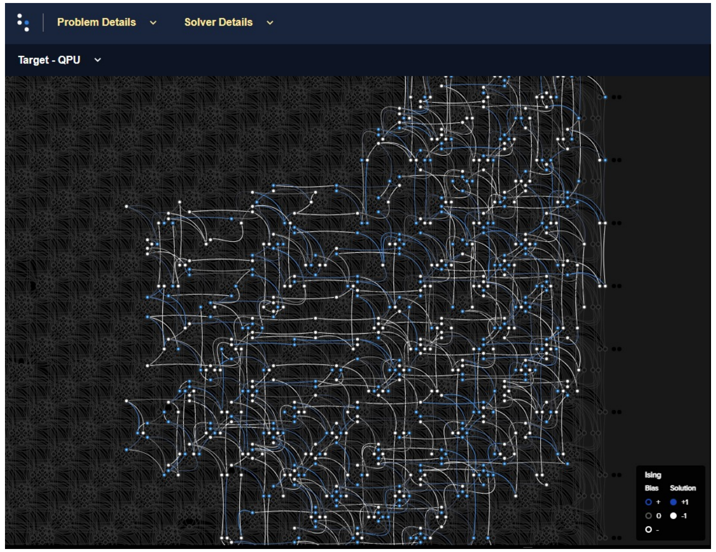
if: $p_{\mathrm{NETWORK}}(\mathbf{x}) \cong 0 \quad \text{where} \quad p_{\mathrm{BOLTZ}}(\mathbf{x}) \approx 1$ → ergodicity problem!

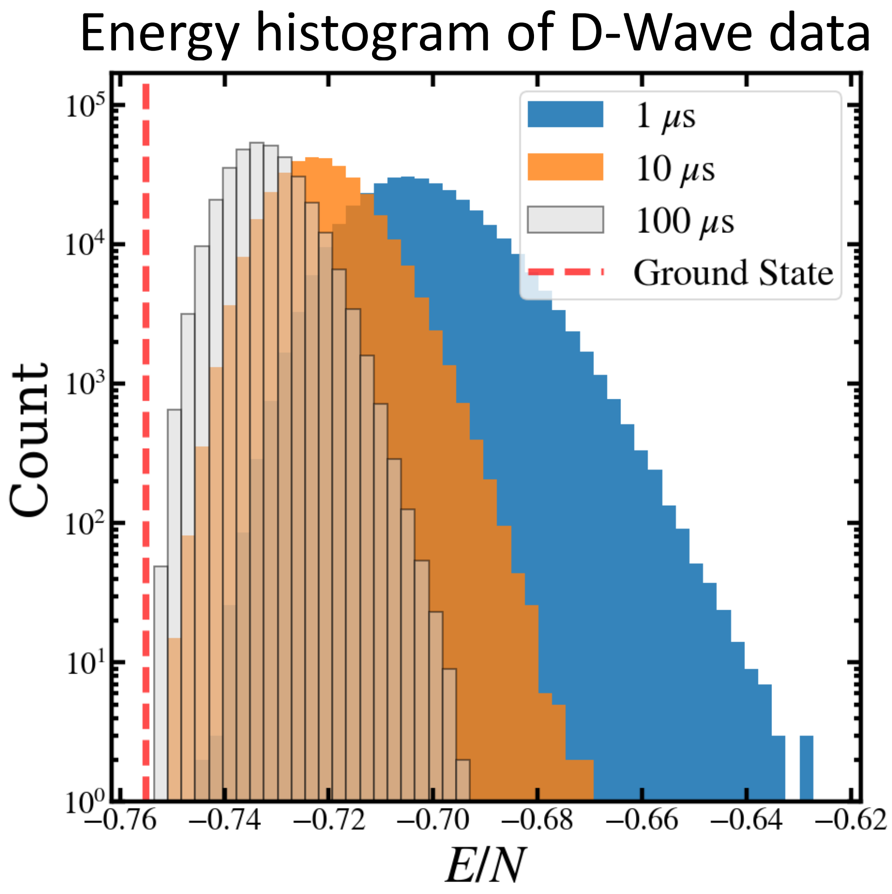
Related work:

- K. A. Nicoli et al., PRE 101, 023304 (2020): MCMC+autoregressive n. for ferromagnetic models.
- F. Noè et al., Science 365, 1147 (2019): normalizing flows for complex-molecule simulations.
- X. Ding et al., J. Phys. Chem. B, 124, 10166 (2020): normalizing flows for free-energy computations.
- M. Gabrié et al., arXiv:2105.12603 (2021): adaptive MCMC via normalizing flows.
- G. S. Hartnett, M. Mohseni,arXiv:2001.00585v2 (2020): spin-glass simulations via normalizing flows.

Neural MC with NN trained on D-WAVE data

Access to D-Wave QPU time via CINECA ISCRA Project





2D square lattice N=484
Nearest and next-nearest neighbor interaction
Uniform random couplings

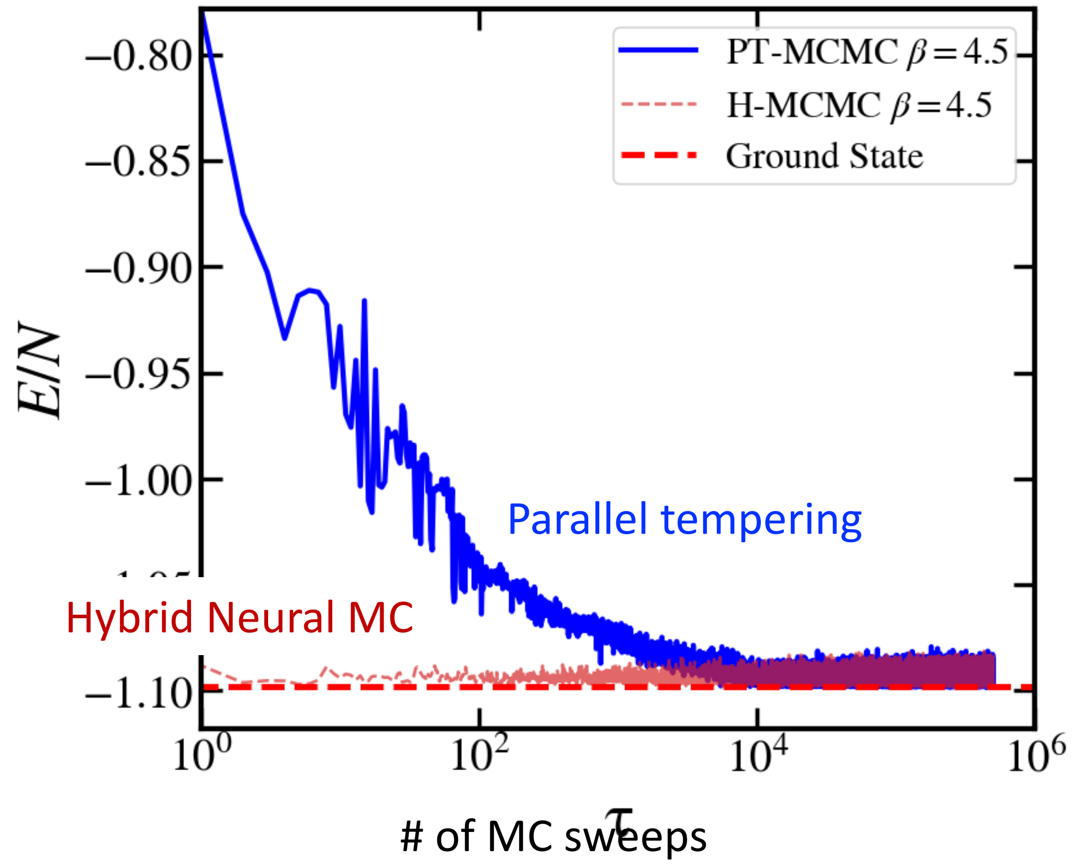
Hybrid neural simulation of a spin glass at low T

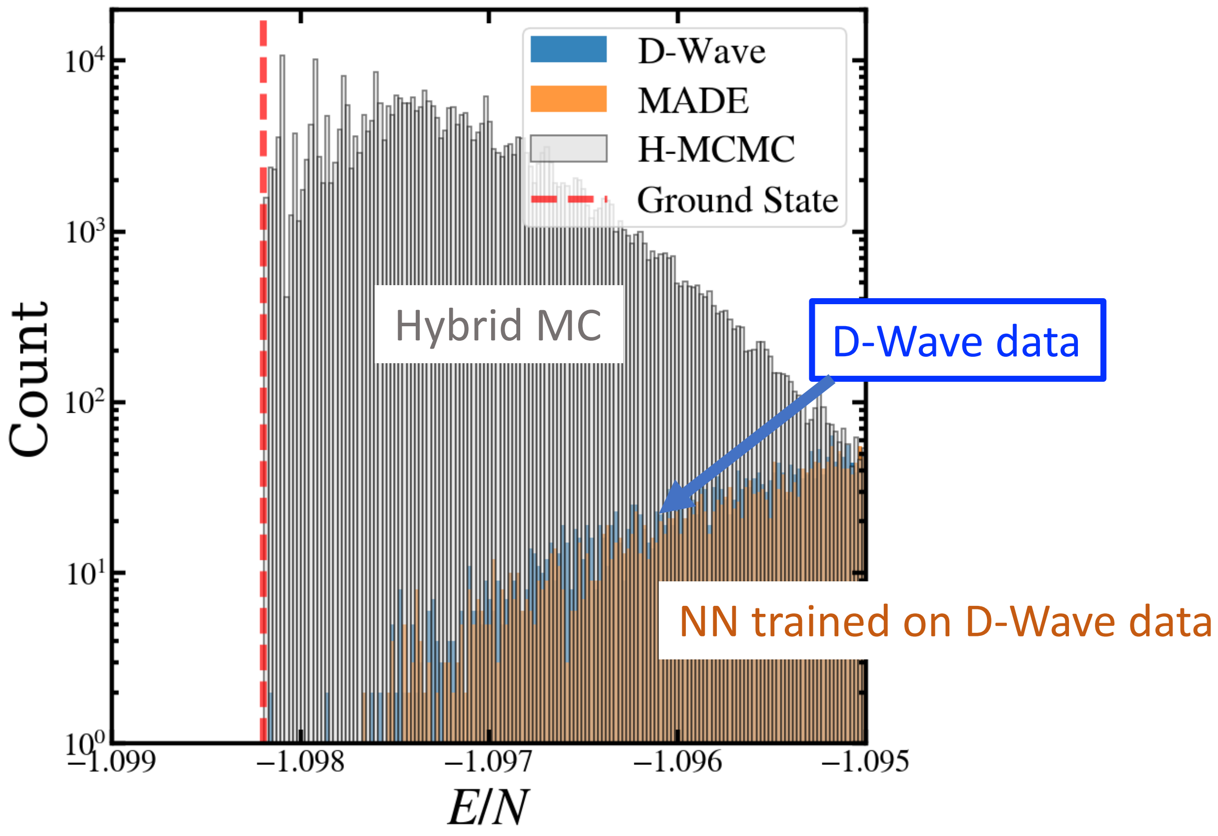
G. Scriva, E. Costa, B. McNaughton, SP, SciPost Physics 15, 018 (2023)

NN trained on D-WAVE data

Annealing time: $100\mu s$

β/J = 4.5
2D square lattice N=484
Nearest and next-nearest neighbor interaction
Uniform random couplings in (-1,1)





DFT for random quantum Ising models via scalable neural networks

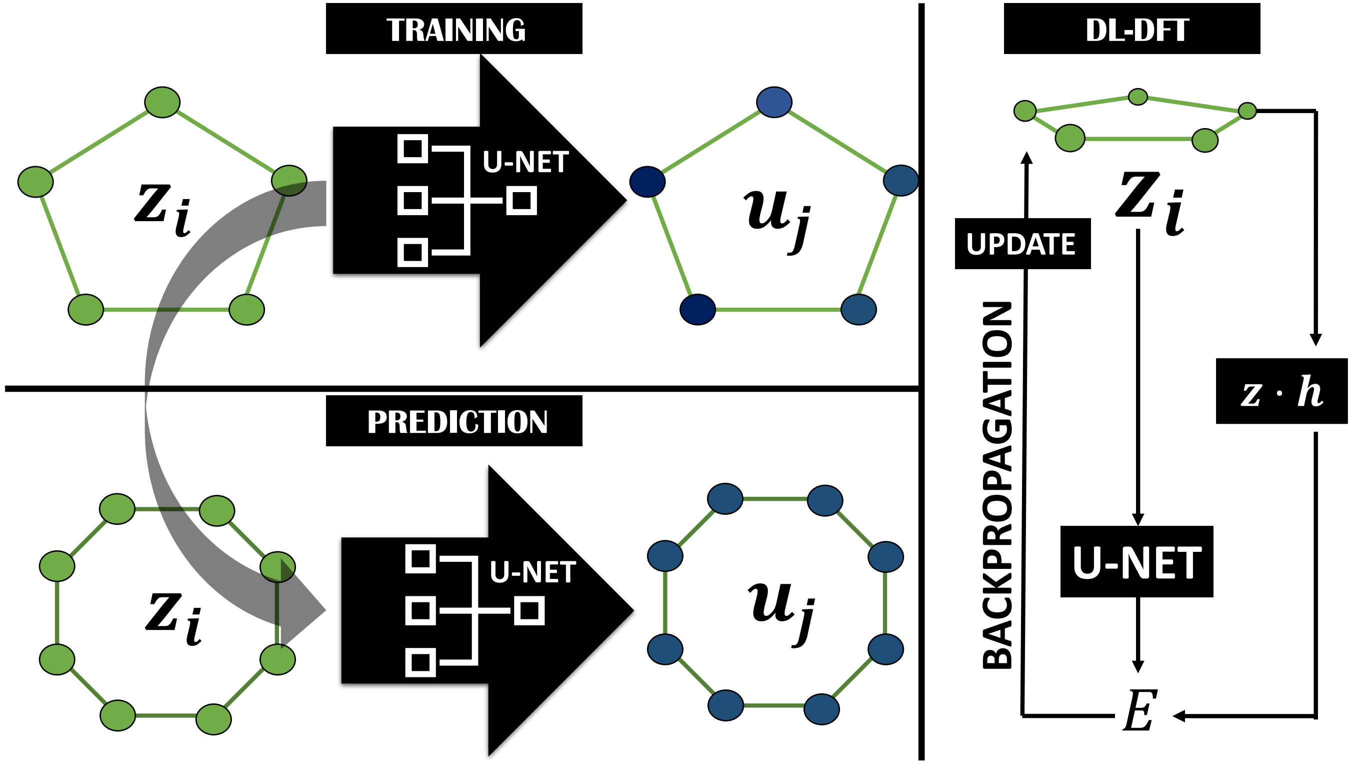
Hamiltonian: $H = -J[\sum_i \sigma_i^x \sigma_{i+1}^x + \sum_i \sigma_i^x \sigma_{i+2}^x] + \sum_i h_i \sigma_i^z$

Universal　　　Random

Functional: $E = E[\boldsymbol{z}] = \sum_j u_j + \boldsymbol{h} \cdot \boldsymbol{z}$

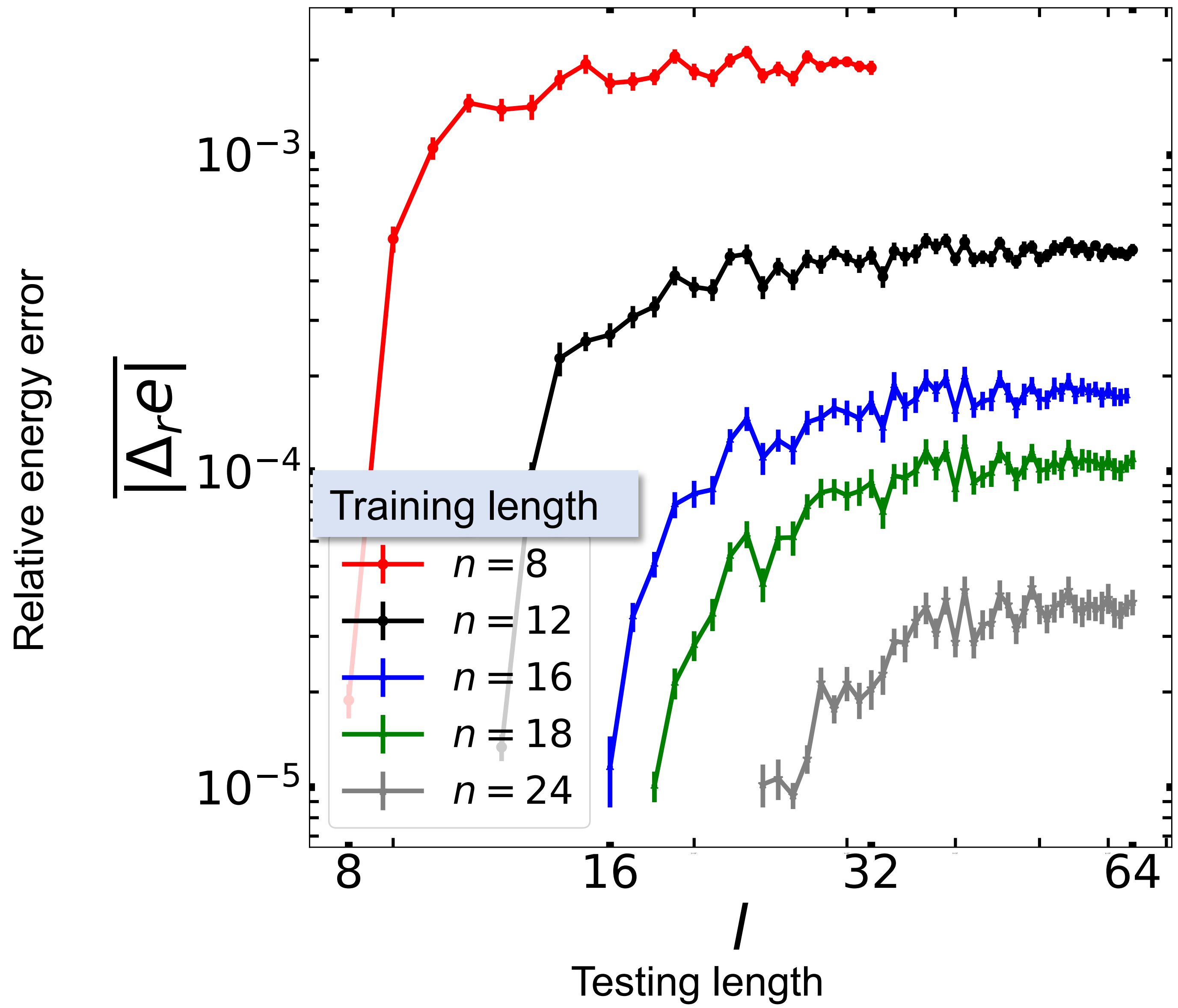
$z_i = \langle \sigma_j^z \rangle$

$u_j = -J \left\langle \left[\sum_i \sigma_i^x \sigma_{i+1}^x + \sum_i \sigma_i^x \sigma_{i+2}^x \right] \right\rangle$



Predicting energies for larger systems than those used for training

E. Costa, R. Fazio, SP, PRB (2023)



Future goal:
dynamics via time-dependent DFT

Projective QMC for Quantum Ising models

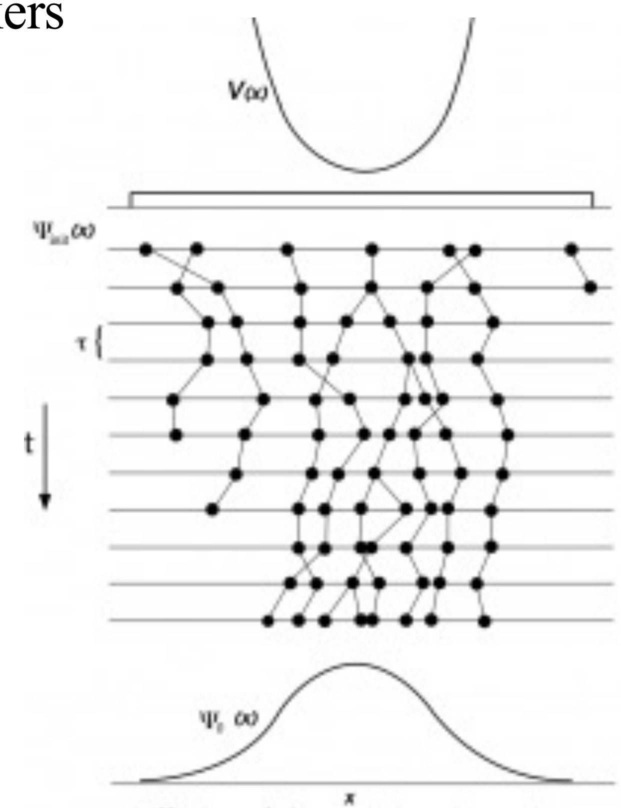
$$H = -\sum_{ij} J_{ij}\sigma_i^z\sigma_j^z - \Gamma\sum_i \sigma_i^x$$

$$\psi(\mathbf{S},\tau) = \exp(-\tau H)\psi(\mathbf{S},0) \underset{\tau\to\infty}{\approx} \psi_0(\mathbf{S},0)$$ Schrödinger eq. in imaginary time

$$\psi(\mathbf{S},\tau+\Delta\tau) = \sum_{\mathbf{S}'} G(\mathbf{S}',\mathbf{S},\Delta\tau)\psi(\mathbf{S}',\tau)$$ defines a Markov process

$G(\mathbf{S}',\mathbf{S},\Delta\tau) \geq 0 \;\Rightarrow$ no sign problem (stoquastic Hamiltonian)

$\sum_{\mathbf{S}'} G(\mathbf{S}',\mathbf{S},\Delta\tau) \neq 1 \Rightarrow$ not a standard Markov process $\Rightarrow$ kill or clone random walkers



Imaginary-time Green's function
$G(\mathbf{S}',\mathbf{S},\Delta\tau) = \langle \mathbf{S}' | \exp(-\Delta\tau(H - E_{ref})) | \mathbf{S} \rangle$

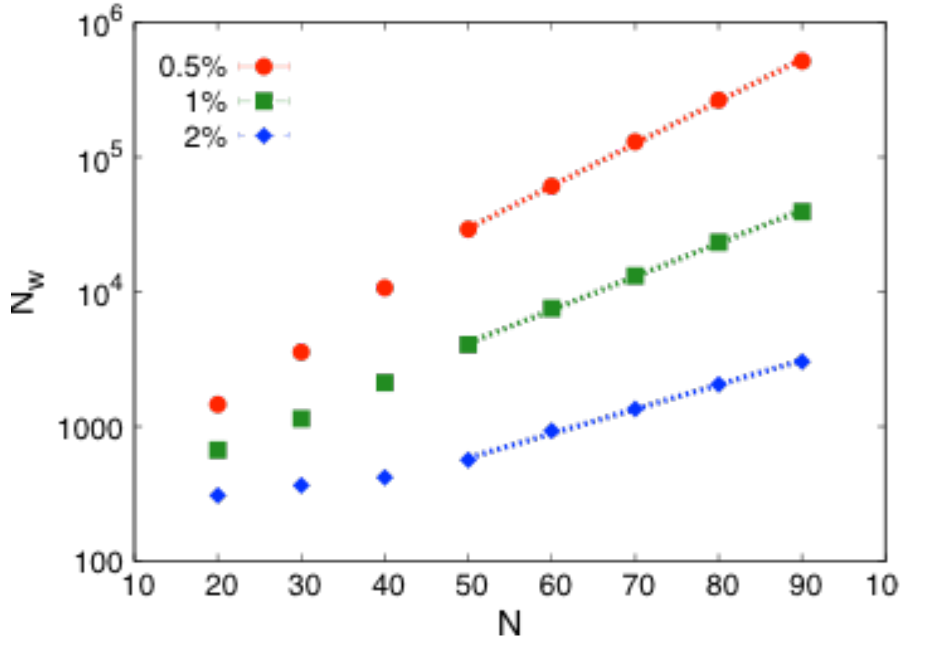
Image from: W. M. C. Foulkes, L. Mitas, R. J. Needs, and G. Rajagopal Rev. Mod. Phys. 73, 33 (2001)

Computational cost of projective QMC simulations

Inack, Giudici, Parolini, Santoro, SP, PRA (2018)

Notice: any diagonal Hamiltonian is stoquastic (sign-problem free).
Finding its ground state encompasses hard classical optimization problems such as k-SAT or MAX-CUT.
Bravyi, Quant. Inf. Comp., Vol. 15, No. 13/14, pp. 1122-1140 (2015)

# of walkers required to keep relative err. fixed



System size

SEE ALSO:
N. Nemec, Phys. Rev. B 81, 035119 (2010).
M. Boninsegni and S. Moroni, Phys. Rev. E 86, 056712 (2012).
K. Ghanem, N. Liebermann, A. Alavi, Phys. Rev. B 103, 155135 (2021)
J. Brand, M. Yang, E, Pahl, Phys. Rev. B 105, 235144 (2022)

Exponentially growing computational cost, even without sign problem

Note: here we use "simple" PQMC algorithm: no guiding wave function.

IMPORTANCE SAMPLING WITH VARIATIONAL ANSATZ

Introduce guiding wave function $\equiv \psi_G(\mathbf{x})$

Modified master eq.: $\Psi(\mathbf{x}, \tau + \Delta\tau)\psi_G(\mathbf{x}) = \sum_{\mathbf{x}'} \tilde{G}(\mathbf{x}, \mathbf{x}', \Delta\tau)\Psi(\mathbf{x}', \tau)\psi_G(\mathbf{x}')$

Modified Green's function: $\tilde{G}(\mathbf{x}, \mathbf{x}', \Delta\tau) = \langle \mathbf{x} | \exp(-\Delta\tau \hat{H} - E_{\mathrm{REF}}) | \mathbf{x}' \rangle \frac{\psi_G(\mathbf{x})}{\psi_G(\mathbf{x}')}$

The guiding wf reduces statistical fluctuations and bias.
Here, we adopt neural network states.

Self-learning projective QMC simulation

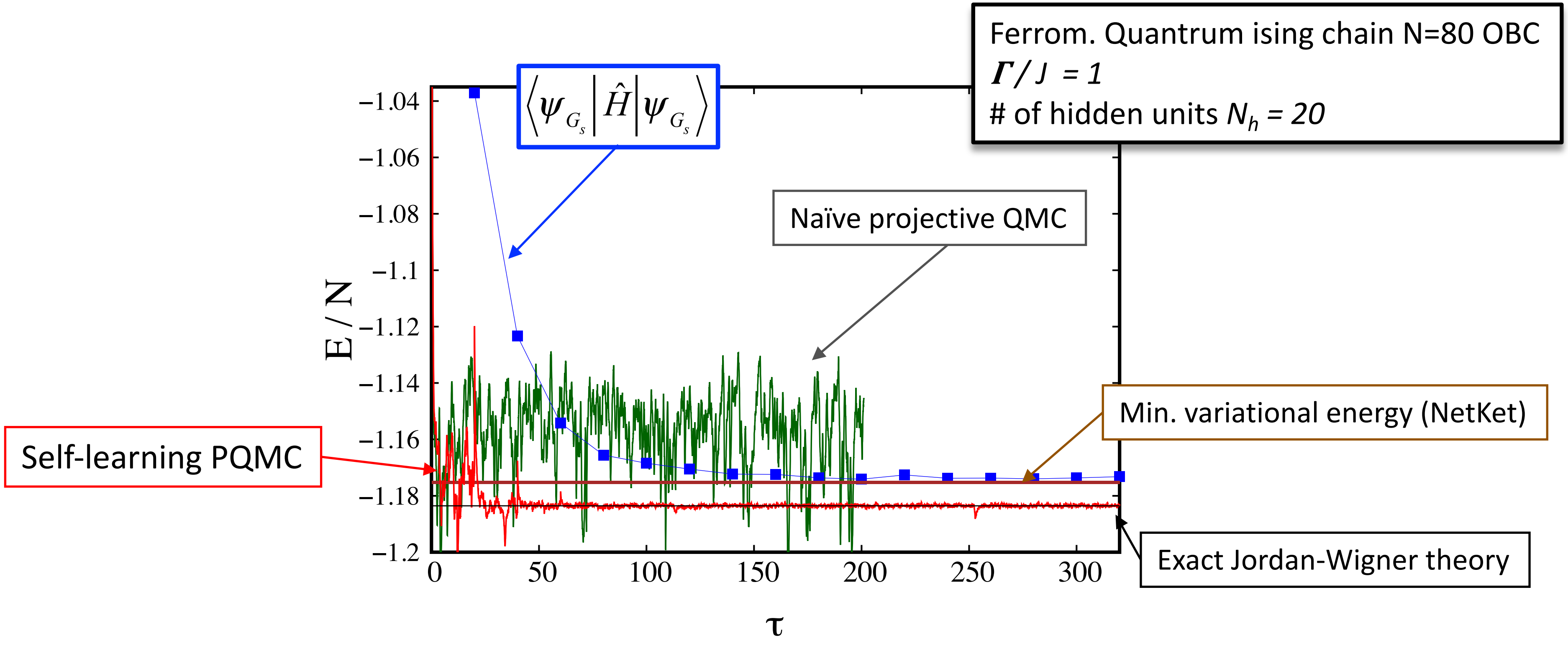
SP, Inack, Pieri, Phys. Rev. E 100, 043301 (2019)

- RBM learns the random-walker distribution: $P(\mathbf{x}) \propto \psi_G(\mathbf{x})\psi_0(\mathbf{x})$
- Guiding wf for the next stint: $\psi_G(\mathbf{x}) = \sqrt{P(\mathbf{x})}$

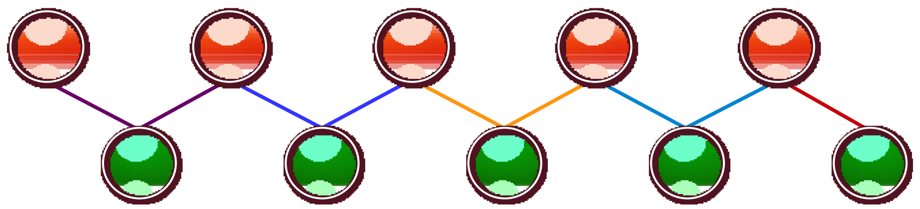
stoquastic model $\Rightarrow \psi_0(\mathbf{x}) \geq 0$

$\psi_{G_0}(\mathbf{x})\psi_0(\mathbf{x})$ $\quad \sqrt{\psi_{G_0}(\mathbf{x})\psi_0(\mathbf{x})}\psi_0(\mathbf{x})$ $\quad$ … $\quad$ … $\quad \psi_{G_0}^{1/2^2}(\mathbf{x})\psi_0^{2-1/2^n}(\mathbf{x})$ $\quad$ …

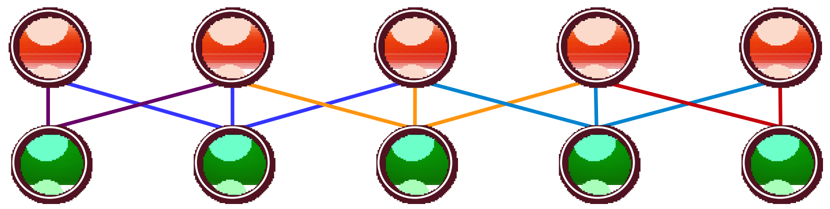
imaginary time



Sparse RBM, connectivity 2

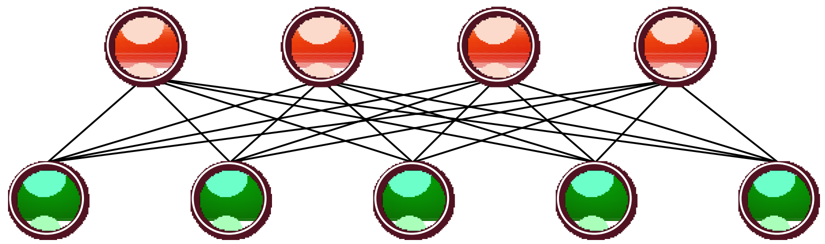


Sparse RBM, connectivity 3



Dense restricted Boltzmann machine

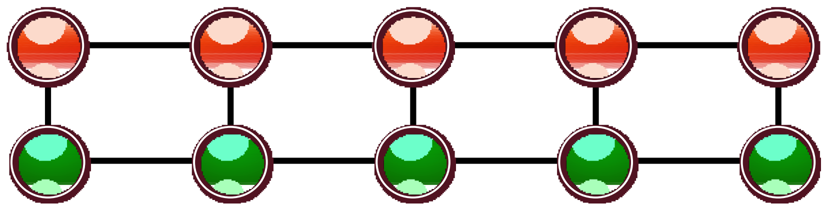
Carleo, Troyer, Science (2017)



Shadow-wave function

Reatto, Masserini, PRB (1988)

Vitiello, Runge, Kalos, PRL (1988)



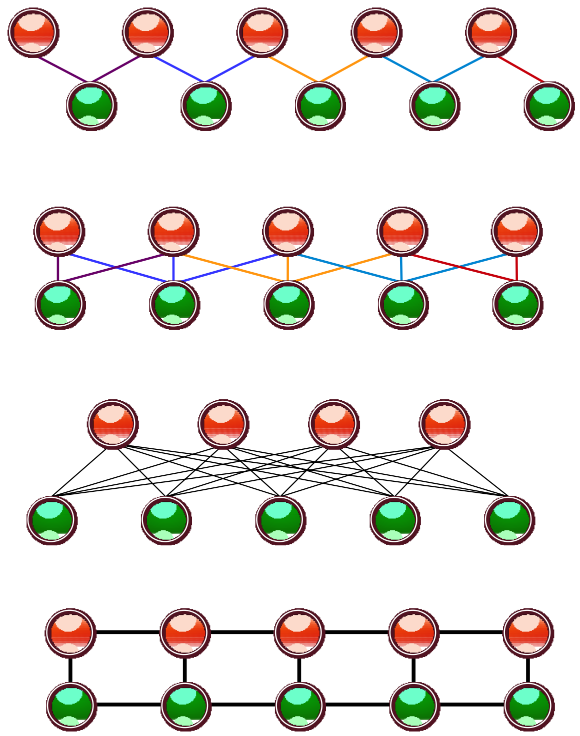
Random 1D quantum Ising chain: sparse neural networks

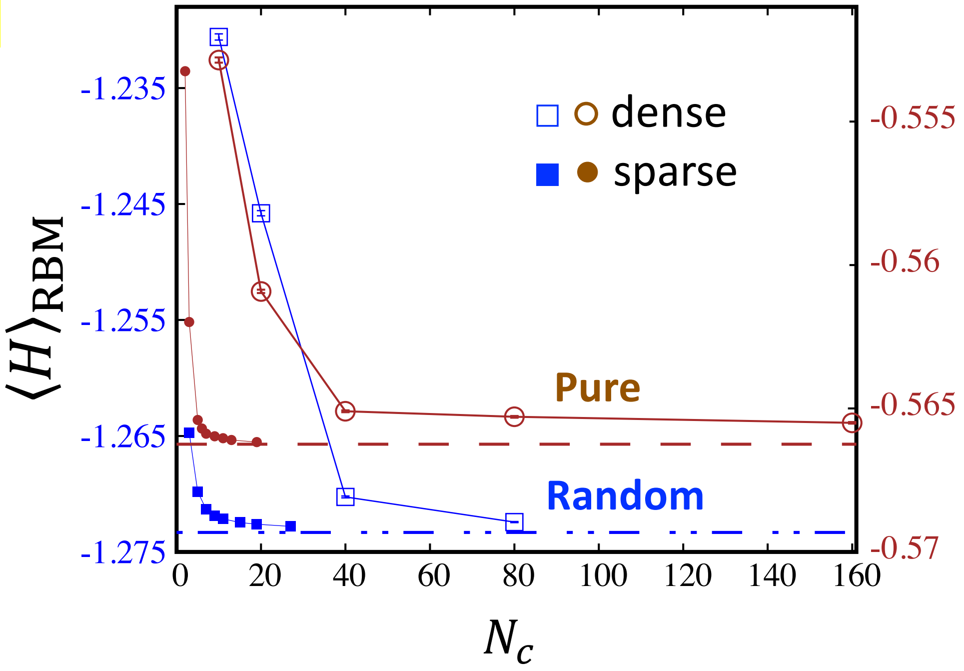
Connectivity $N_c = 2$

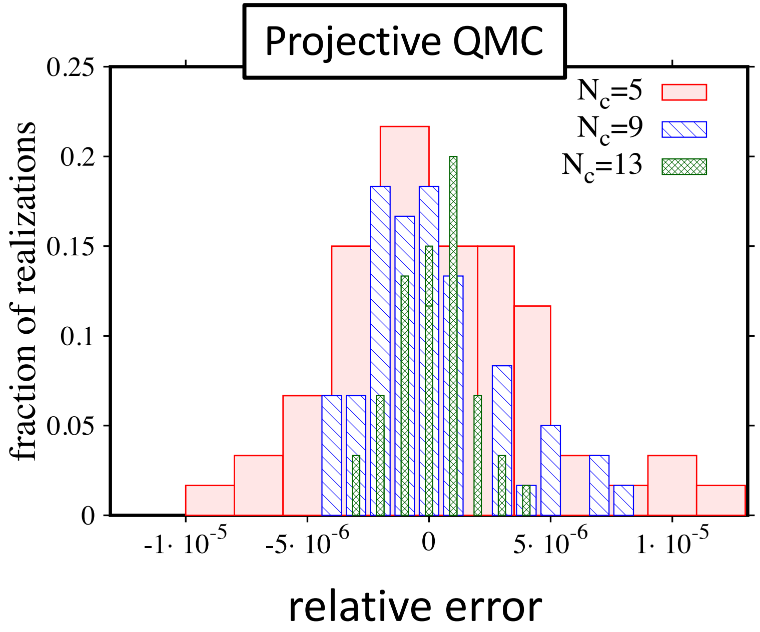
Connectivity $N_c = 3$

Dense

Shadow wf







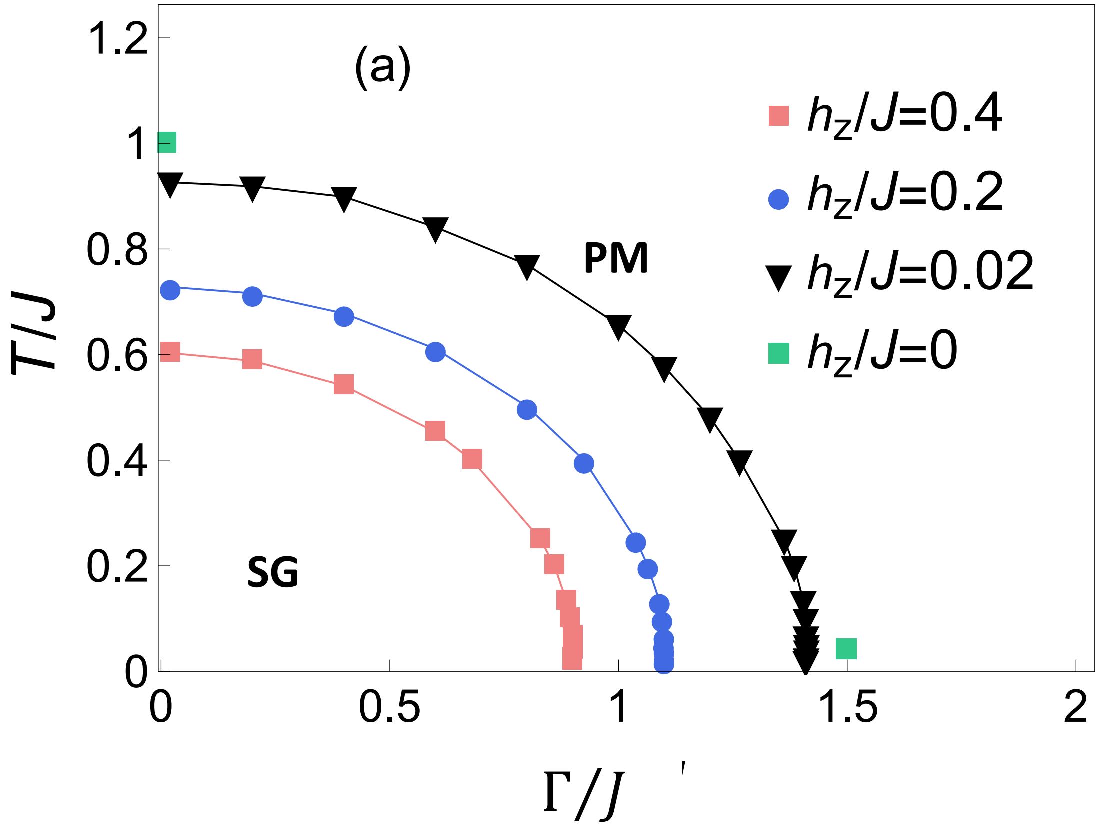
SP, Pieri, PRE (2020)

Quantum spin glasses

$$H = -\sum_{i<j} J_{i,j}\sigma_i^z\sigma_j^z - \Gamma\sum_i \sigma_i^x - h_z\sum_i \sigma_i^z$$

- 1D, n.n. interactions: no frustration, no spin glass
- Finite D (e.g., 2D or 3D): spin glass, RSB vs droplet?
- Sherrington-Kirkpatrick: $J_{i,j} \sim \mathrm{Norm}[\mu = 0, \sigma = J/\sqrt{N}]$

Exact MF solution (classical), replica-symmetry breaking!



Quantum SK: **dynamical MFT**

A. Kiss et al. arXiv:2306.07337 (2023)

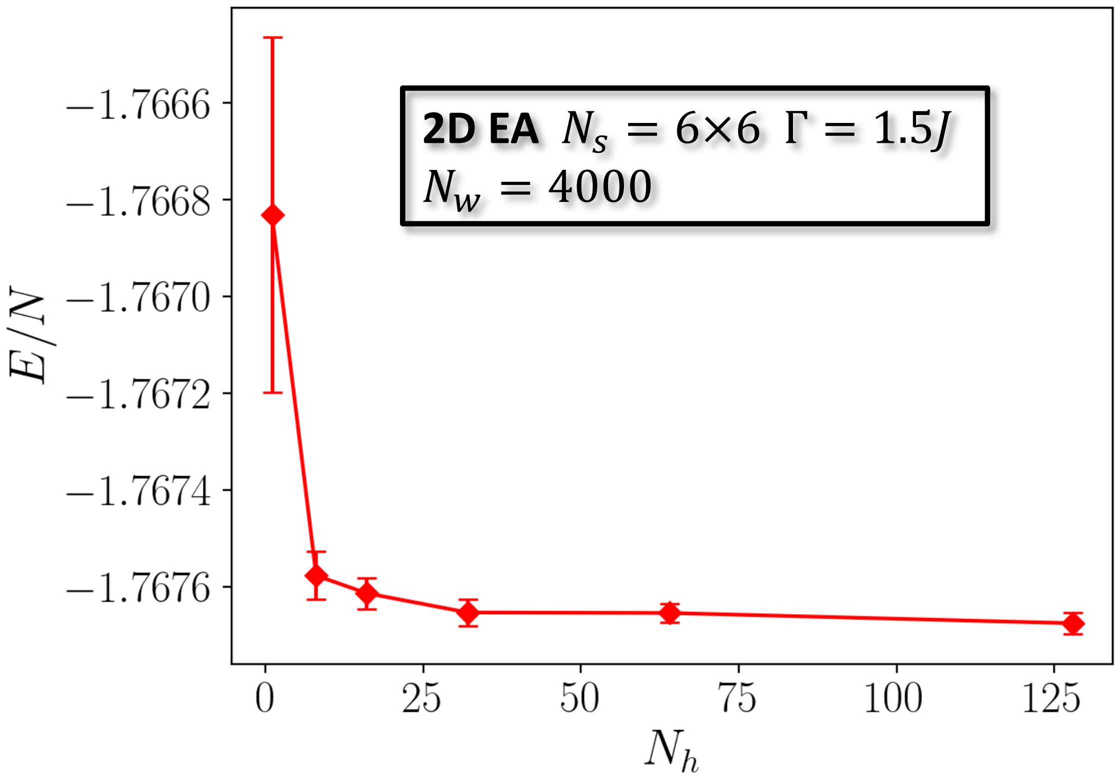
See also:

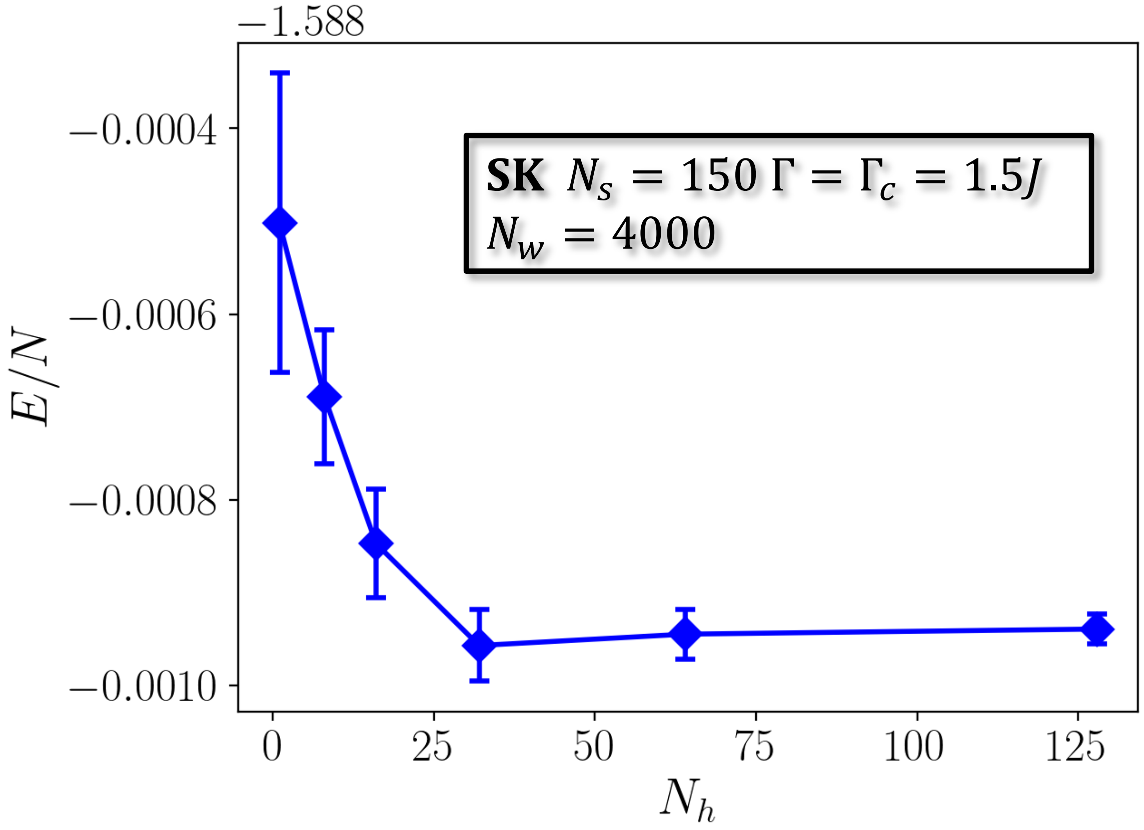
Leschke et al., PRL 127, 207204 (2021)

Schindler et al., PRL 129, 220401 (2022)

Population control bias

Self-learning DMC with N_w =4000 walkers guided by RBM with N_h hidden neurons





DMC for Quantum SK model: PRELIMINARY RESULTS!!!

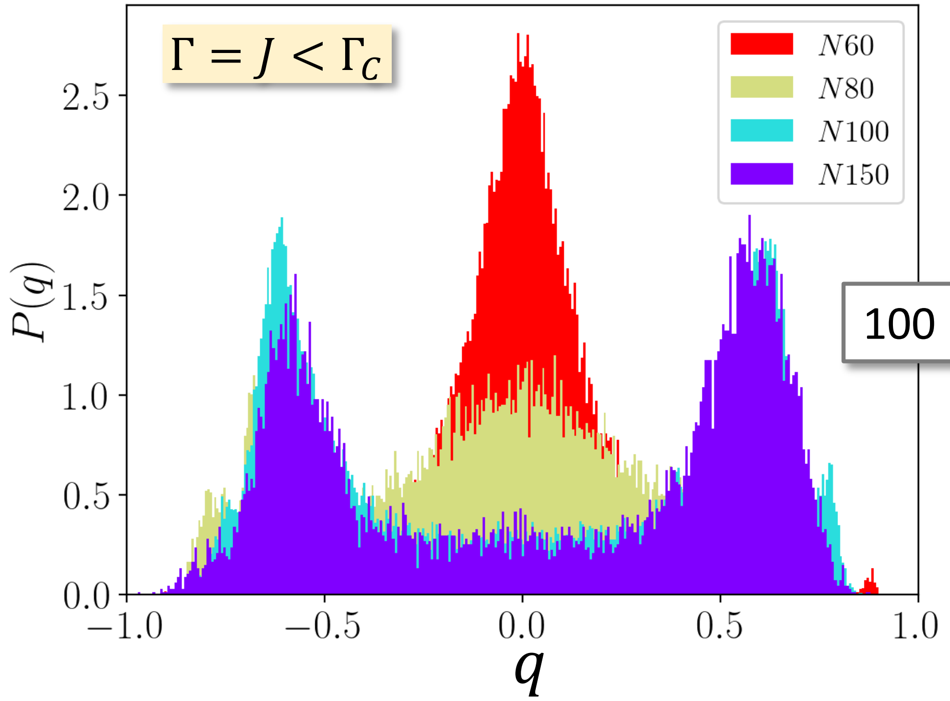
Edward-Anderson order parameter: $q_{EA} = \frac{1}{N}\sum_i \langle \sigma_i^Z \rangle^2$

Spin overlap: $q = \frac{1}{N}\sum_i \sigma_{iA}^Z \sigma_{iB}^Z$ Two copies: $H_{\text{TOT}} = H_A + H_B$ (separable RBM)

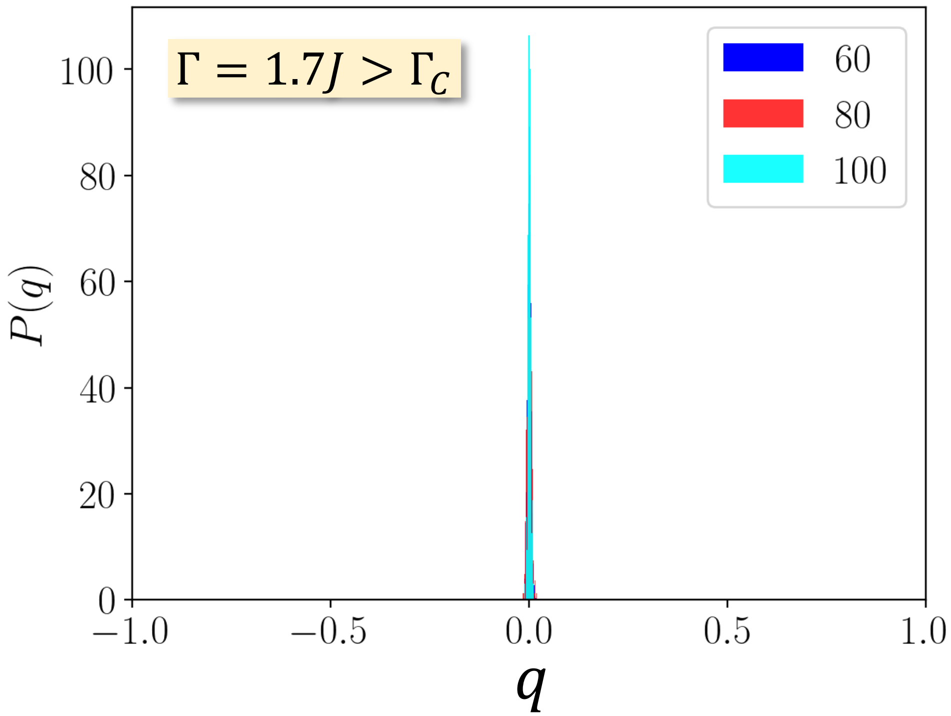
$\rightarrow \quad q_{EA} = \langle q \rangle$

Overlap distribution: Droplet picture $P(q) = \frac{1}{2}\delta(q - q_{EA}) + \frac{1}{2}\delta(q + q_{EA})$

Replica simmetry breaking $P(q = 0) > 0$



100 realizations



Summary:

- QMB physics via deep learning with NQS, DFT, supervised learning.
- Scalability matters.
- Sign problem is not the only problem.
- Promising results with neural DMC for quantum spin glasses.

- Scriva, Costa, McNaughton, SP, SciPost Phys. 15, 018 (2023)
- Costa, Fazio, SP, Phys. Rev. B (2023)
- Costa, Scriva, Fazio, SP, Phys. Rev. E 106 (4), 045309 (2022)
- Mujal, Martínez Miguel, Polls, Juliá-Díaz, S Pilati, SciPost Phys. 10 (3), 073 (2021)
- Saraceni, Cantori, SP, Phys. Rev. E 102, 033301 (2020)
- McNaughton, Milošević, Perali, SP, Phys. Rev. E 101 (5), 053312 (2020)
- SP, Pieri, Phys. Revs E 101 (6), 063308 (2020)
- SP, Inack, Pieri, Phys. Rev. E 100, 043301 (2019)
- SP, Pieri, Sci. Rep. 9, 5613 (2019)
- Inack, Giudici, Parolini, Santoro, SP, Phys. Rev. A 97 (3), 032307 (2018)