Learning disordered interactions between Rydberg atoms from experimental shapshots

Olivier Simard[1,2*], Anna Dawid, Joseph Tindall, Michel Ferrero, Anirvan Sengupta and Antoine Georges

[1] *Collège de France, 11 place Marcelin Berthelot, 75005 Paris, France*
[2] *CPHT, CNRS, École Polytechnique, IP Paris, F-91128 Palaiseau, France*

Objectives

- Can one infer experimentally realized Hamiltonian in cold atom systems in a scalable manner through measurements?
- What machine learning model and physical input should be used?
- Are there regions in the physical parameter space facilitating the learning process?

Introduction

We make use of the Principal Neighbor Aggregator (PNA) [1] graph neural network (GNN) to predict the relative nearest-neighbor (NN) atomic displacements in Rydberg arrays [2] from the local magnetization and spin-spin correlation functions. For practicability, we ought to

- use a limited amount of physical observables (correlators) for training,
- use a viable number of snapshot measurements,
- map out accurately and rapidly the effective Hamiltonian implemented.

Models and Methods

Transverse-field Ising model

The two-dimensional transverse-field Ising model (TFIM) reads

$$\hat{H} = \sum_{i \neq j} \underbrace{\frac{C_6}{R_{ij}^{\,6}}}_{\equiv J_{i,j}} \hat{\sigma}_i^z \hat{\sigma}_j^z + \hbar\Omega \sum_i \hat{\sigma}_i^x + \hbar\delta \sum_i \hat{\sigma}_i^z, \quad (1)$$

where i, j runs over all atomic positions, $C_6/\hbar \simeq 5.420158 \times 10^6 \text{rad} \cdot \mu\text{s}^{-1} \cdot \mu\text{m}^6$ is the dipole-dipole interaction term [3]. Ω stands for the transverse Ising field (Raby frequency), δ for the detuning field, $\hat{\sigma}^{z(x)}$ denotes the (off)-diagonal real Pauli matrix.
The physical input consists in the local magnetization M_i dressing the graph nodes

$$M_i = \langle \hat{\sigma}_i^z \rangle_{0;\vec{\Omega}} \quad (2)$$

and spin-spin correlation functions $\chi_{i,j}$ weighting the graph edges

$$\chi_{i,j;\vec{\Omega}} = \langle \hat{\mathbf{S}}_i^z \cdot \hat{\mathbf{S}}_j^z \rangle_{0;\vec{\Omega}}, \quad (3)$$

for a sequence of transverse field $\vec{\Omega}$, with

$$\hat{\mathbf{S}}_l^z = \frac{1}{2} \sum_{\mu,\nu} \hat{\Psi}_{l,\mu}^\dagger \sigma_{\mu,\nu}^z \hat{\Psi}_{l,\nu}\,,$$

where $\hat{\Psi}_l \equiv (\hat{c}_{l,\uparrow}, \hat{c}_{l,\downarrow})$. $\hat{c}_{l,\sigma}^{(\dagger)}$ is the electronic annihilation (creation) operator, on site l with spin σ. The indices μ, ν span over the spin space and the Hadamard operator defined in the Z basis is applied onto the wave function to measure in the X basis.

PNA GNN

We minimize the squared L^2 modulus of the difference between the predictions $\hat{\mathbf{y}}$ and the truths $\mathbf{y}$:

$$\text{MSE}(\mathbf{y}, \hat{\mathbf{y}}) = \|\mathbf{y} - \hat{\mathbf{y}}\|^2.$$

The coefficient of determination is

$$R^2 = 1 - \frac{\sum_i (y_i - \hat{y}_i)^2}{\sum_i (y_i - \bar{y}_i)^2}, \quad (4)$$

and the mean absolute error (MAE)

$$\text{MAE} = \sum_{i=1}^{N} \frac{|\hat{y}_i - y_i|}{N}. \quad (5)$$

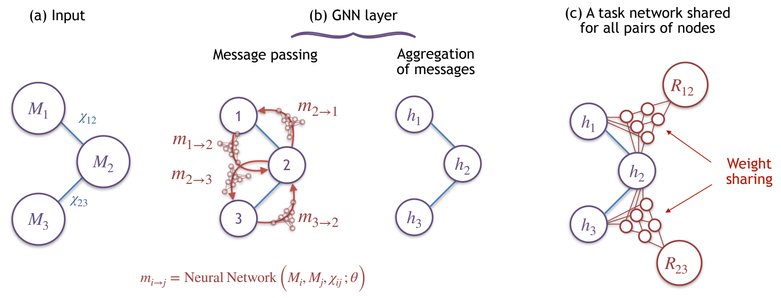

Figure: Elements of a GNN architecture. a) Graph input consisting of the spin-spin correlation functions. b) GNN hidden layer updates assisted by aggregated message passing between neighboring nodes. c) Task neural network predicting distances between nodes based off physical observables.

Case #1	M and Ω-history
Case #2	M, χ^{NN} and Ω-history
Case #3	M, χ^{NN}, χ^{NNN} and Ω-history
Case #4	M, χ^{NN}, NNN edges and Ω-history
Case #5	χ^{NN}, NNN edges and Ω-history
Case #6	M, χ^{NN}, χ^{NNN} and Ω-history in Z+X bases

Table: Summary of the training cases used in this work.

Results

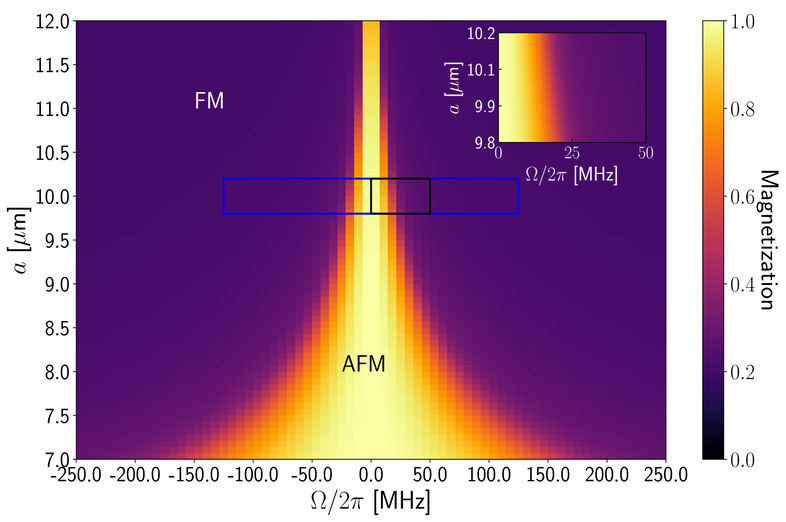

Figure: Phase diagram at $\delta = 0$. The blue box encloses the Ω-history. The inset plot is a zoom within the black box.

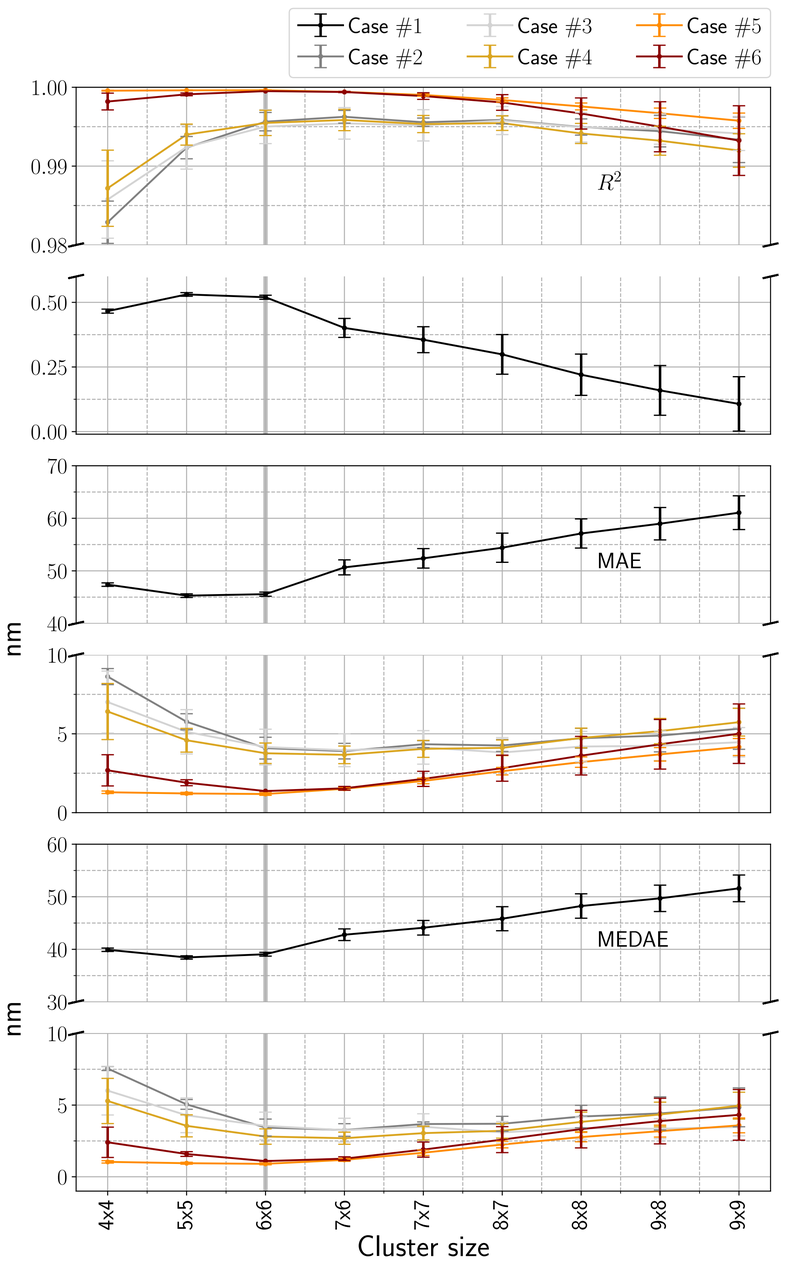

Figure: Top panel: R^2 for various cluster sizes in the test dataset. Middle panel: Mean absolute error of predictions. Bottom panel: Median of the absolute error (MEDAE) of predictions vs truths. The vertical grey line indicates the extrapolating threshold. Cases are detailed in the table.

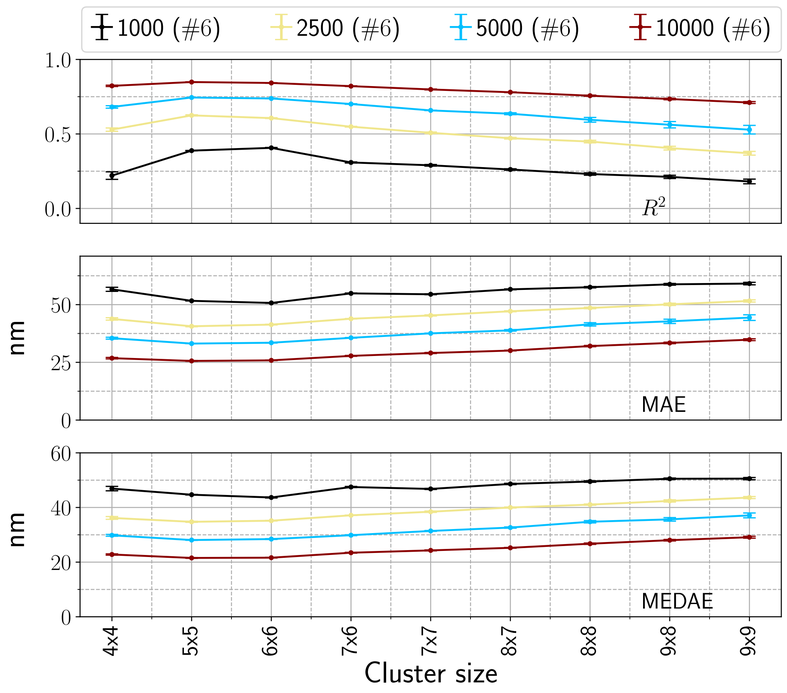

Figure: A view of the metrics so far considered (see annotations) as a function of the cluster size. The case #6 is considered. The dataset has been generated using a the same finite number of snapshots.

Conclusion

- The amount of cluster sizes in training dataset matters when extrapolating to larger sizes.
- Constraining the GNN physical input to NN spin-spin correlation functions is sufficient to perform scalable Hamiltonian learning.
- Training with more 'weightless' graph edges generates an architectural benefit, akin to skip connections.
- Compiling the dataset samples across an array $\vec{\Omega}$ helps gathering up the information from the more relevant values of Ω in the phase diagram.
- Overcompleting the prediction-vs-truth correspondance by including longer-range correlation functions improves the outcome of the model at larger cluster sizes.

References

[1] Gabriele Corso, Luca Cavalleri, Dominique Beaini, Pietro Liò, and Petar Veličković.
Principal neighbourhood aggregation for graph nets, 2020.

[2] A. Browaeys and T. Lahaye.
Many-body physics with individually controlled Rydberg atoms.
Nat. Phys., 16(2):132–142, 2020.

[3] Pascal Scholl, Michael Schuler, Hannah J. Williams, Alexander A. Eberharter, Daniel Barredo, Kai-Niklas Schymik, Vincent Lienhard, Louis-Paul Henry, Thomas C. Lang, Thierry Lahaye, Andreas M. Läuchli, and Antoine Browaeys.
Quantum simulation of 2d antiferromagnets with hundreds of rydberg atoms.
Nature, 595(7866):233–238, July 2021.

* Email: olivier.simard@polytechnique.edu

Acknowledgements