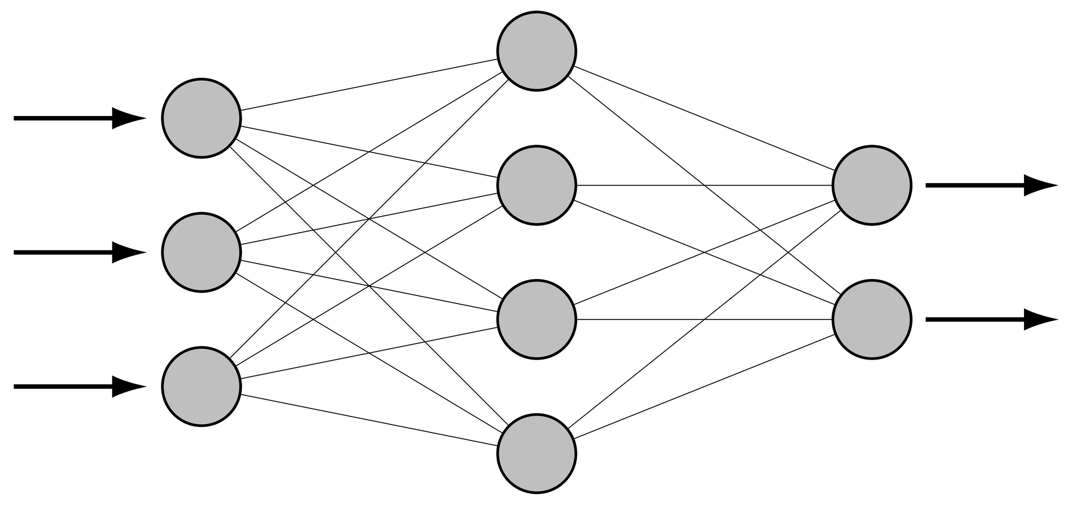
Sign Structure and Machine Learning

Tom Westerhout[†], Nikita Astrakhantsev, Konstantin S. Tikhonov, Mikhail I. Katsnelson, Andrey A. Bagrov

† Theory of Condensed Matter, Radboud University



Generalization vs. Expressibility

Expressibility is a capacity to accurately represent the data using a moderate number of parameters.
Generalization is the ability to correctly predict the results on inputs which were never encountered during training.

Introduction

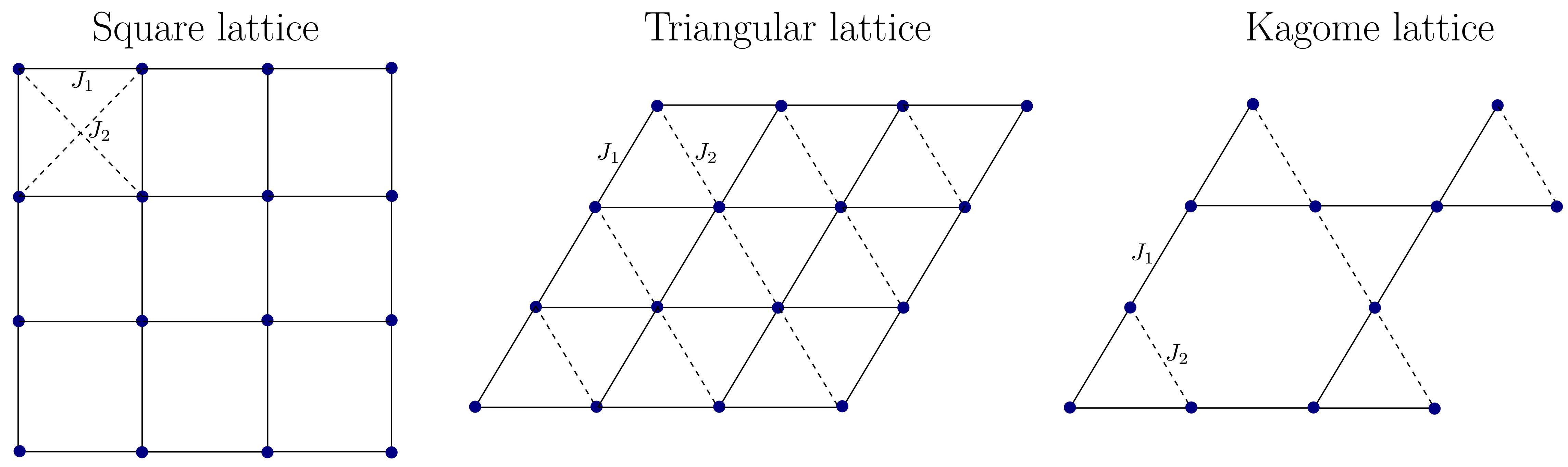
Neural networks can be used as a simple yet very unrestrictive ansatz for describing quantum many-body wavefunctions. They work well for simple systems such as one- and two-dimensional Heisenberg and transverse field Ising models, but have not yet been successful applied to frustrated systems.
Significant effort has been put into the search for neural quantum states (NQS) architectures that have good *expressibility*. Here, we study *generalization* of NQS and show that it might be the main reason why NQS fail in a number of physically interesting scenarios.

Supervised Learning

for $i \in \{1, \dots, epochs\}$ **do**
  for $(x, y) \in$ *training dataset* **do**
    $\hat{y} \leftarrow \psi_{\mathcal{W}}(x)$; // compute prediction
    compute loss $\mathcal{L}(\hat{y}, y)$;
    estimate gradient $\nabla\mathcal{L}$;
    update weights $\mathcal{W}$ using $\nabla\mathcal{L}$;
  end
  compute metrics on *validation dataset*;
end

Here, $\psi_{\mathcal{W}}$ is our wavefunction represented by a neural network and parametrized by some weights $\mathcal{W}$.

Systems



Heisenberg Hamiltonian

$$\hat{H} = J_1 \sum_{\langle a,b\rangle} \hat{\boldsymbol{\sigma}}_a \otimes \hat{\boldsymbol{\sigma}}_b + J_2 \sum_{\langle\langle a,b\rangle\rangle} \hat{\boldsymbol{\sigma}}_a \otimes \hat{\boldsymbol{\sigma}}_b$$

Sum runs over the unfrustrated sublattice (solid lines)

Sum runs over the sublattice which brings in frustrations (dashed lines)

Ground state wavefunction

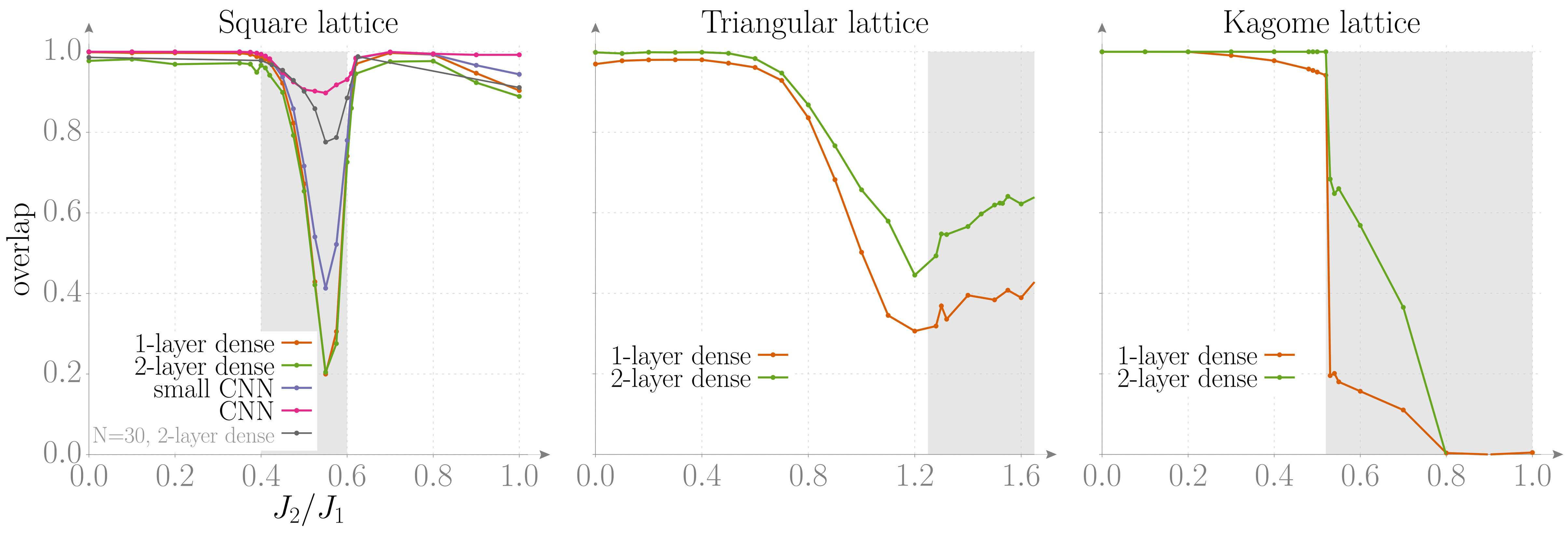
$$|\Psi_{GS}\rangle = \sum_{i=1}^{K} \psi_i |\mathcal{S}_i\rangle = \sum_{i=1}^{K} s_i |\psi_i|\, |\mathcal{S}_i\rangle$$

sign

amplitude

Varying Frustration Level

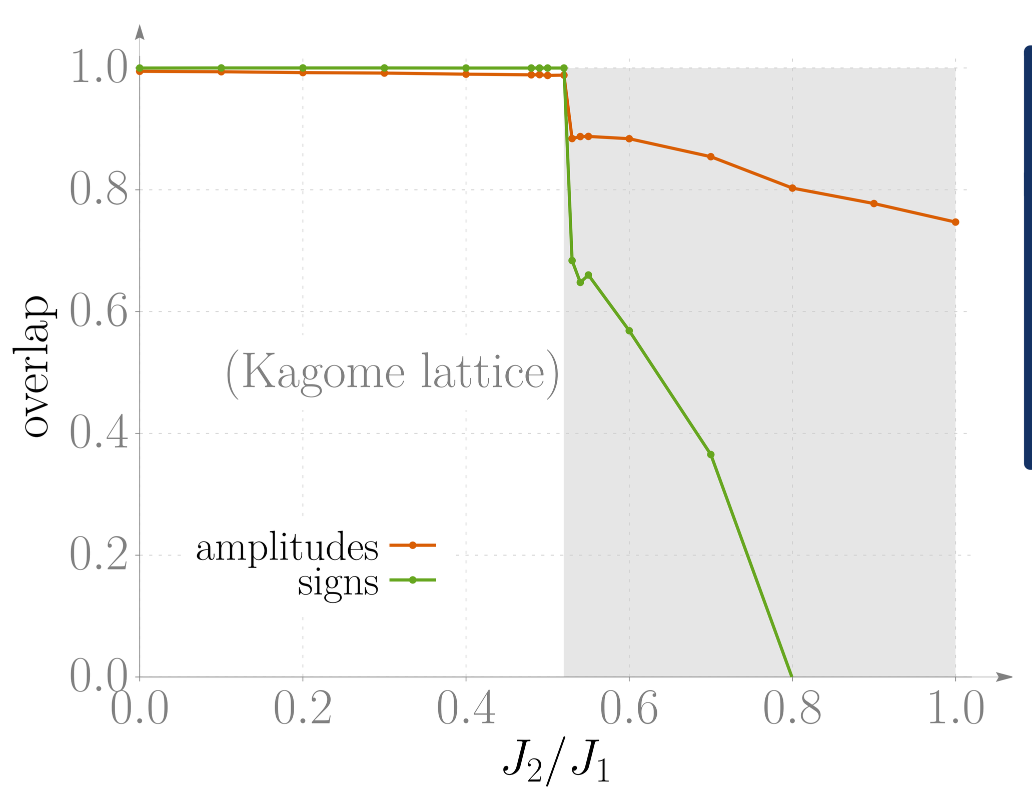
$J_2/J_1 = 0$ and $J_2/J_1 = \infty$ both correspond to unfrustrated regimes.
Perfect generalization means overlap = 1.



Result 1 & 2

- Generalization from a relatively small subset of the Hilbert space basis of the wavefunction sign structure is not granted even when the ansatz is able to express the ground state with high accuracy. Very well known to machine learning practitioners, this fact is also valid for spin systems, in both frustrated and ordered phases.
- Construction and training of a network to achieve good generalization, a task which is relatively simple in the ordered phase, becomes much harder in frustrated region.

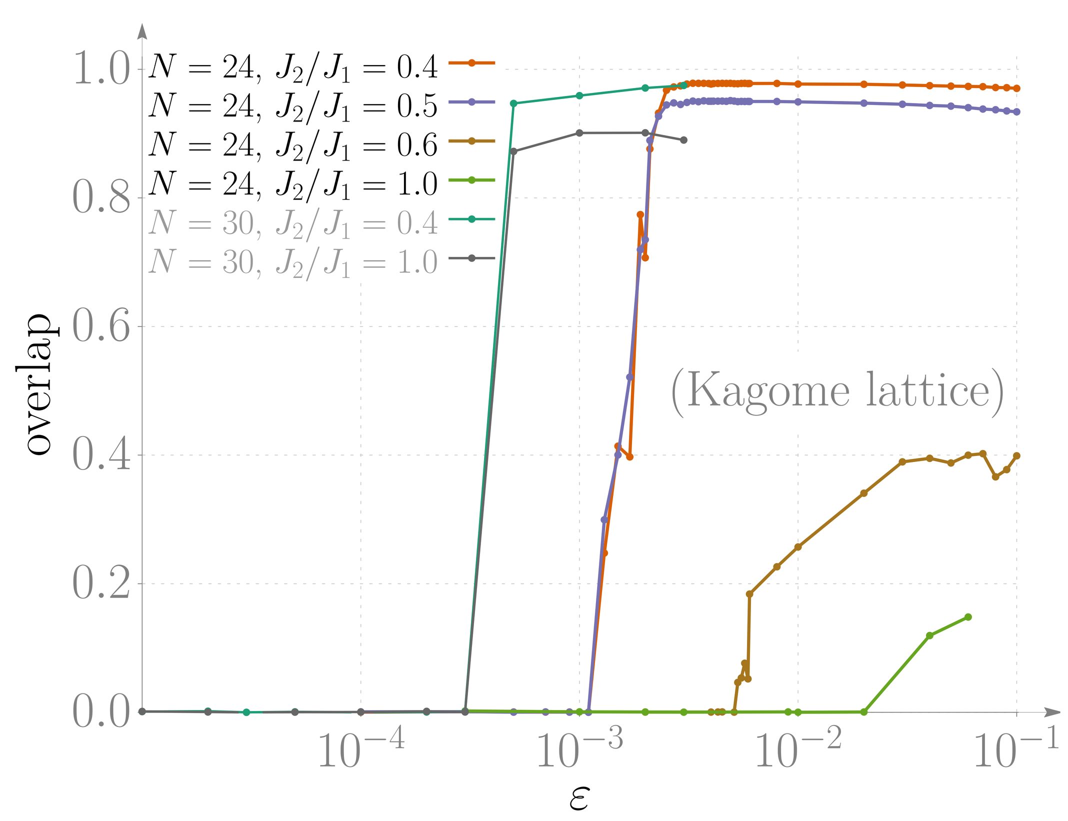
Amplitudes vs. Signs



Result 3

Predicting wavefunction amplitudes is substantially easier than predicting signs.

More Data?



Result 4

Generalization quality depends on the size of the training dataset in an abrupt way exhibiting a sharp increase at some critical fraction ε.

https://arxiv.org/pdf/1907.08186.pdf