UNIVERSITÀ DEGLI STUDI DI MILANO
DIPARTIMENTO DI FISICA

On the lattice-based neural network quantum states

Carlo Barbieri

Next Generation Ab Initio Nuclear Theory — ECT*, July 14-18, 2025

Neural Network Quantum States (NQS):

$$\frac{\langle \Psi_V | H | \Psi_V \rangle}{\langle \Psi_V | \Psi_V \rangle} = E_V \geq E_0$$

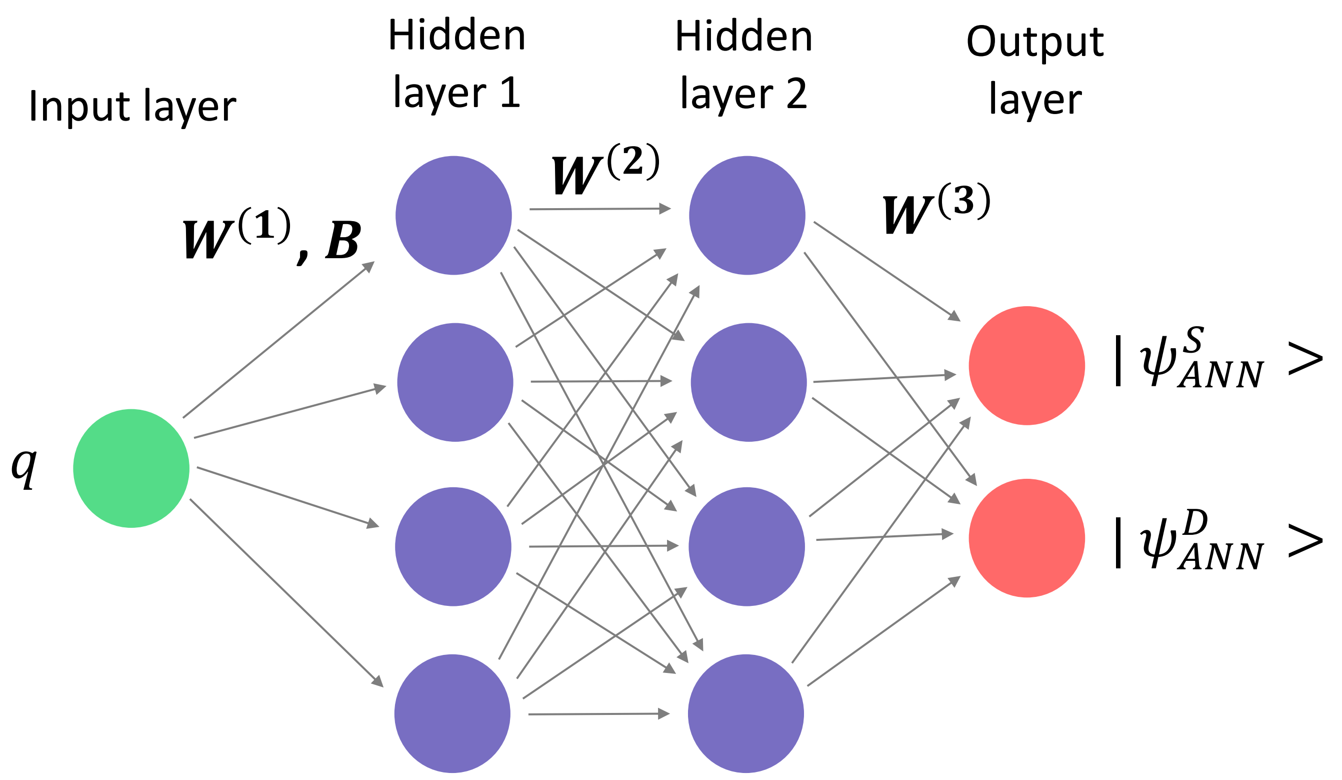
$$|\Psi_V^J\rangle = \prod_{i<j<k} \left(1 - \sum_{\mathrm{cyc}} u(r_{ij}) u(r_{jk})\right) \prod_{i<j} f(r_{ij}) |\Phi\rangle$$

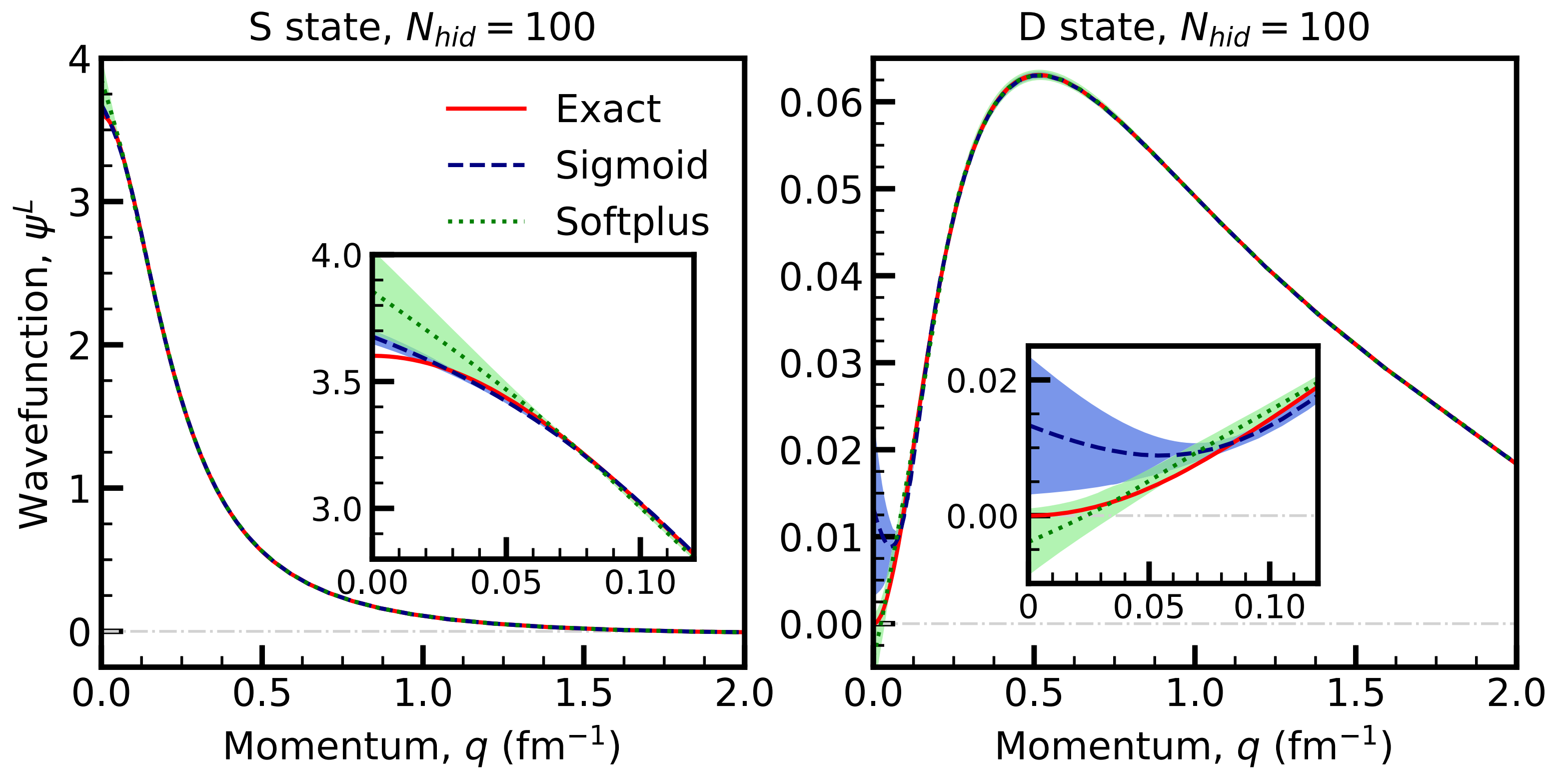
neural networks

And the NQS is…. a VMC with a NN trial wave function!

Early NQS for nuclear physics

The **deuteron**:





Keeble, Rios, Phys. Lett. B **809**, 135743 (2020)
Rozalén Sarmiento, Keeble, EPJ Plus **139**, 189 (2024)

Light nuclei:

	Λ	VMC-ANN	VMC-JS	GFMC	GFMC$_c$
^{2}H	4 fm^{-1}	$-2.224(1)$	$-2.223(1)$	$-2.224(1)$	-
	6 fm^{-1}	$-2.224(4)$	$-2.220(1)$	$-2.225(1)$	-
^{3}H	4 fm^{-1}	$-8.26(1)$	$-7.80(1)$	$-8.38(2)$	$-7.82(1)$
	6 fm^{-1}	$-8.27(1)$	$-7.74(1)$	$-8.38(2)$	$-7.81(1)$
^{4}He	4 fm^{-1}	$-23.30(2)$	$-22.54(1)$	$-23.62(3)$	$-22.77(2)$
	6 fm^{-1}	$-24.47(3)$	$-23.44(2)$	$-25.06(3)$	$-24.10(2)$

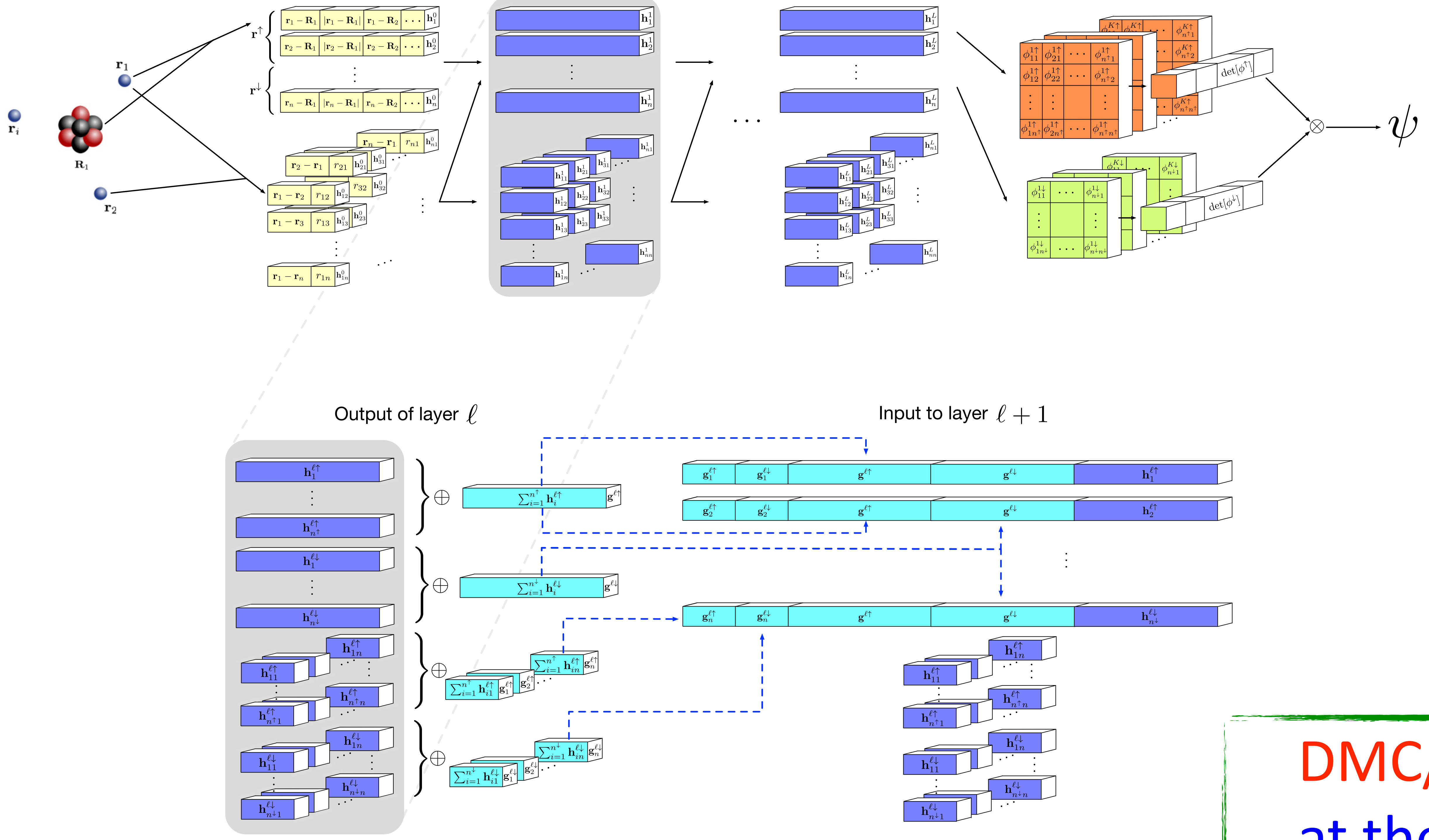
C. Adams, G. Carleo, A. Lovato, N. Rocco,
Phys. Rev. Lett. **127**, 022502 (2021)

Some example of NQS in quantum chemistry

Slater determinants of permutation-equivariant functions: $\phi_i^k(\mathbf{x}_j; \{\mathbf{x}_{/j}\})$

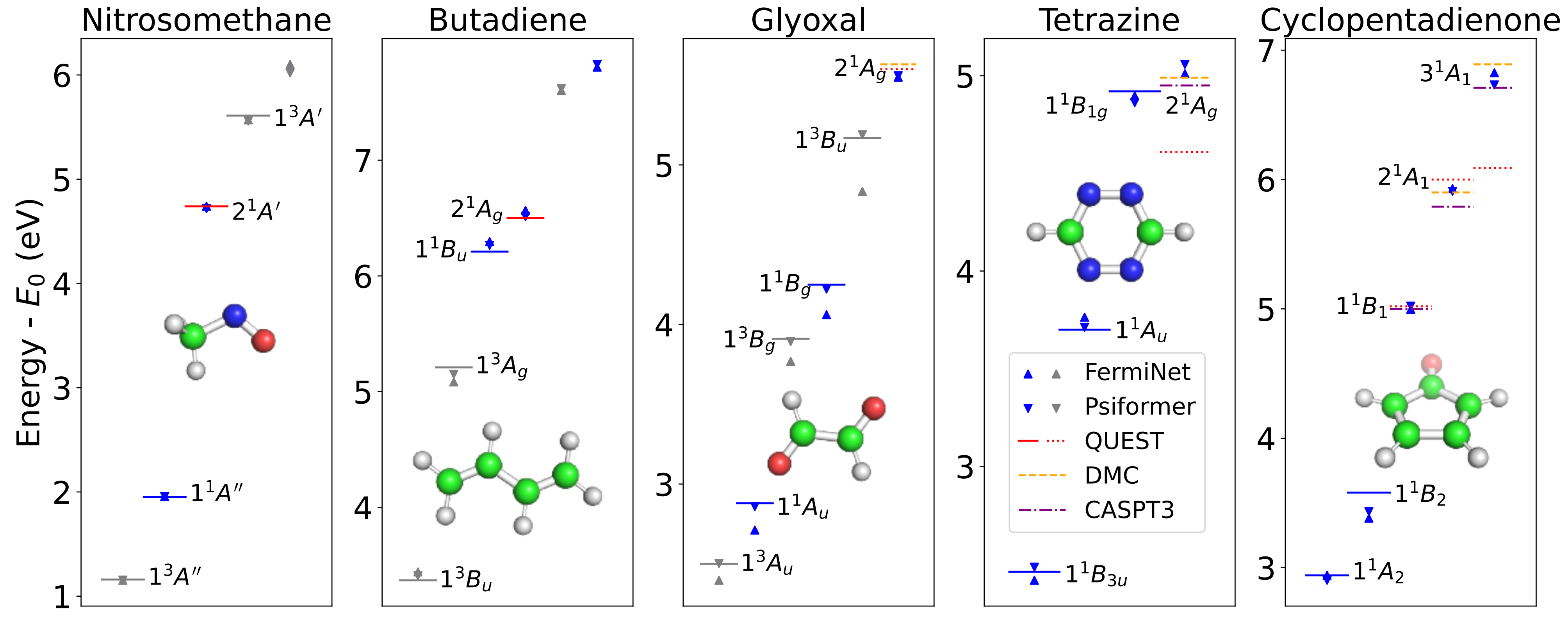
FermiNet NQS:

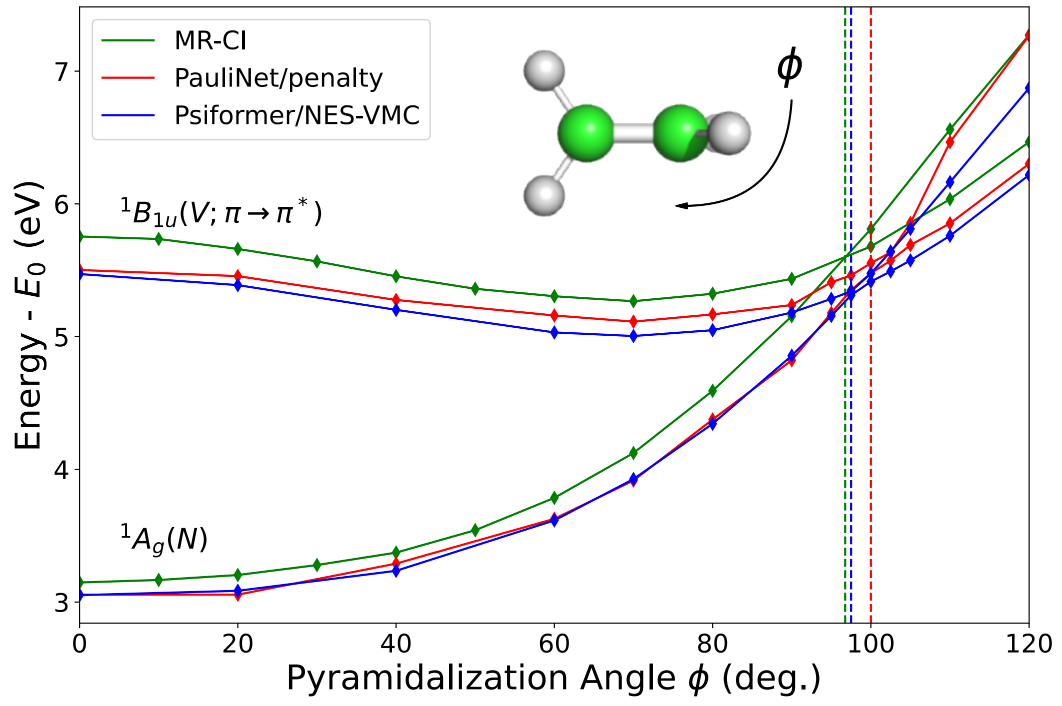
D. Pfau et al., Phys. Rev. Res. **2**, 033429 (2020)



FermiNet for excited states:

D. Pfau et al., Science **385**, 6711 (2024)



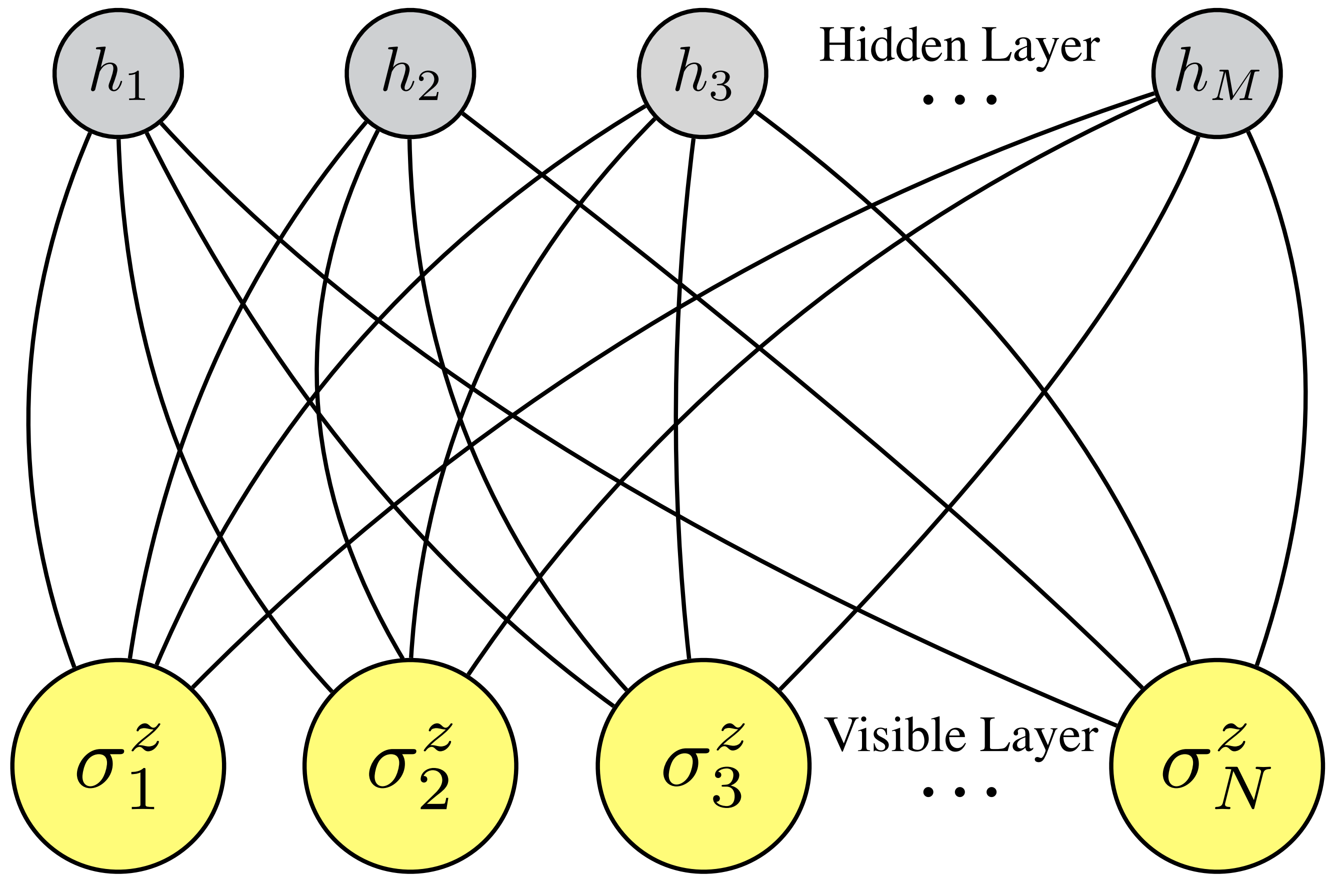


DMC/GFMC accuracy at the cost of VMC !

And many other architectures for correlated fermions: see e.g., J. Hermann et al., Nature Reviews Chemistry **7**, 692–709 (2023)

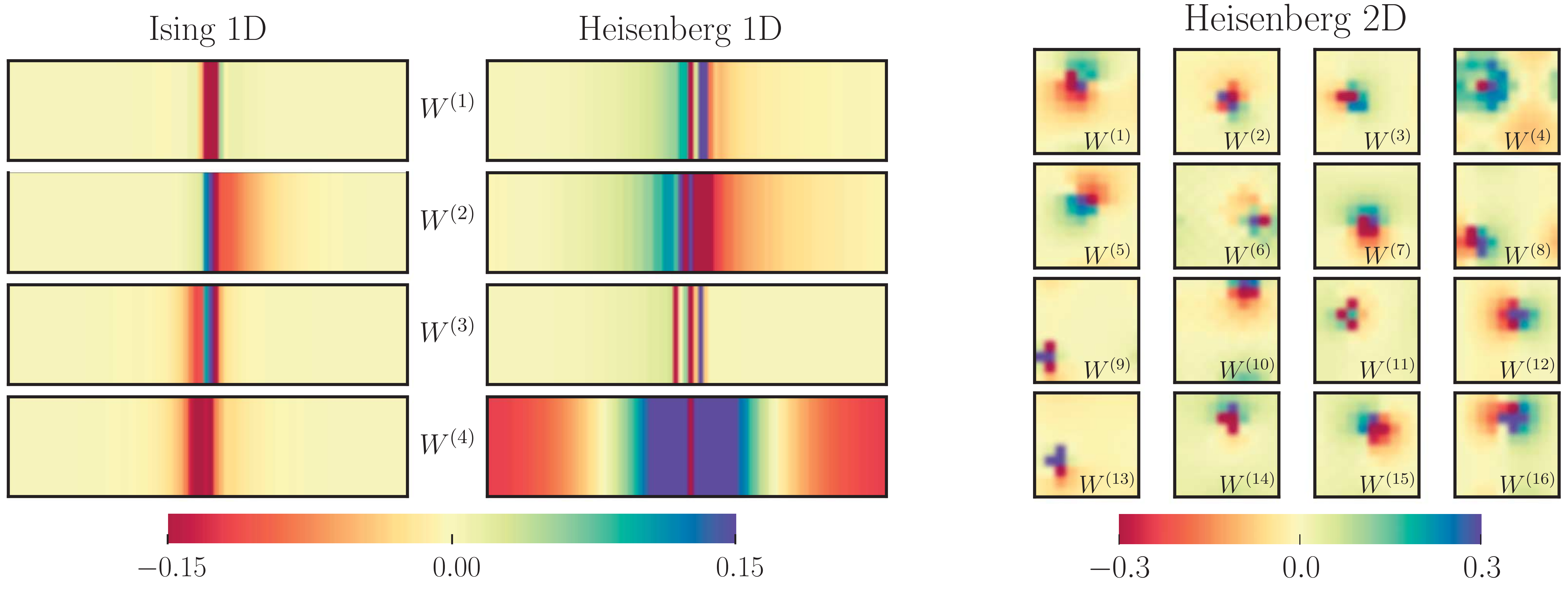
NQS for spin lattices — RBMs

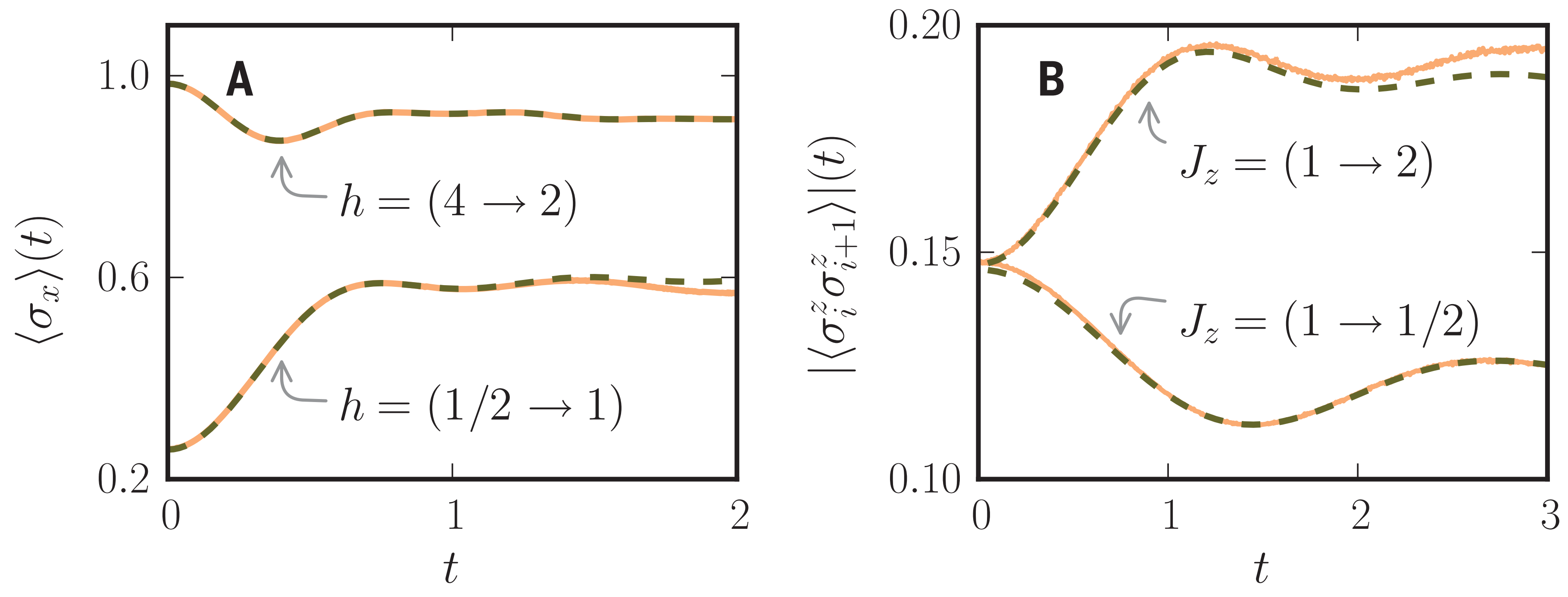
Restricted Boltzmann Machines:



$$E^{\mathcal{W}} = \frac{\langle \psi^{\mathcal{W}} | \hat{H} | \psi^{\mathcal{W}} \rangle}{\langle \psi^{\mathcal{W}} | \psi^{\mathcal{W}} \rangle}$$

$$\Psi_M(\mathcal{S};\mathcal{W}) = \sum_{\{h_i\}} e^{\sum_j a_j \sigma_j^z + \sum_i b_i h_i + \sum_{ij} W_{ij} h_i \sigma_j^z}$$





[G. Carleo & M. Troyer, Science **355**, 602 (2017)]
[K. Choo, et al.,Nature Comm., **11**, 2368, (2020).]

Fig. 4. Many-body unitary time evolution with NQS. NQS results (solid lines) for the time evolution induced by a quantum quench in the microscopic parameters of the models we study (the transverse

Why a NQS on the Lattice?

Advantages of a lattice-NQS for nuclei:

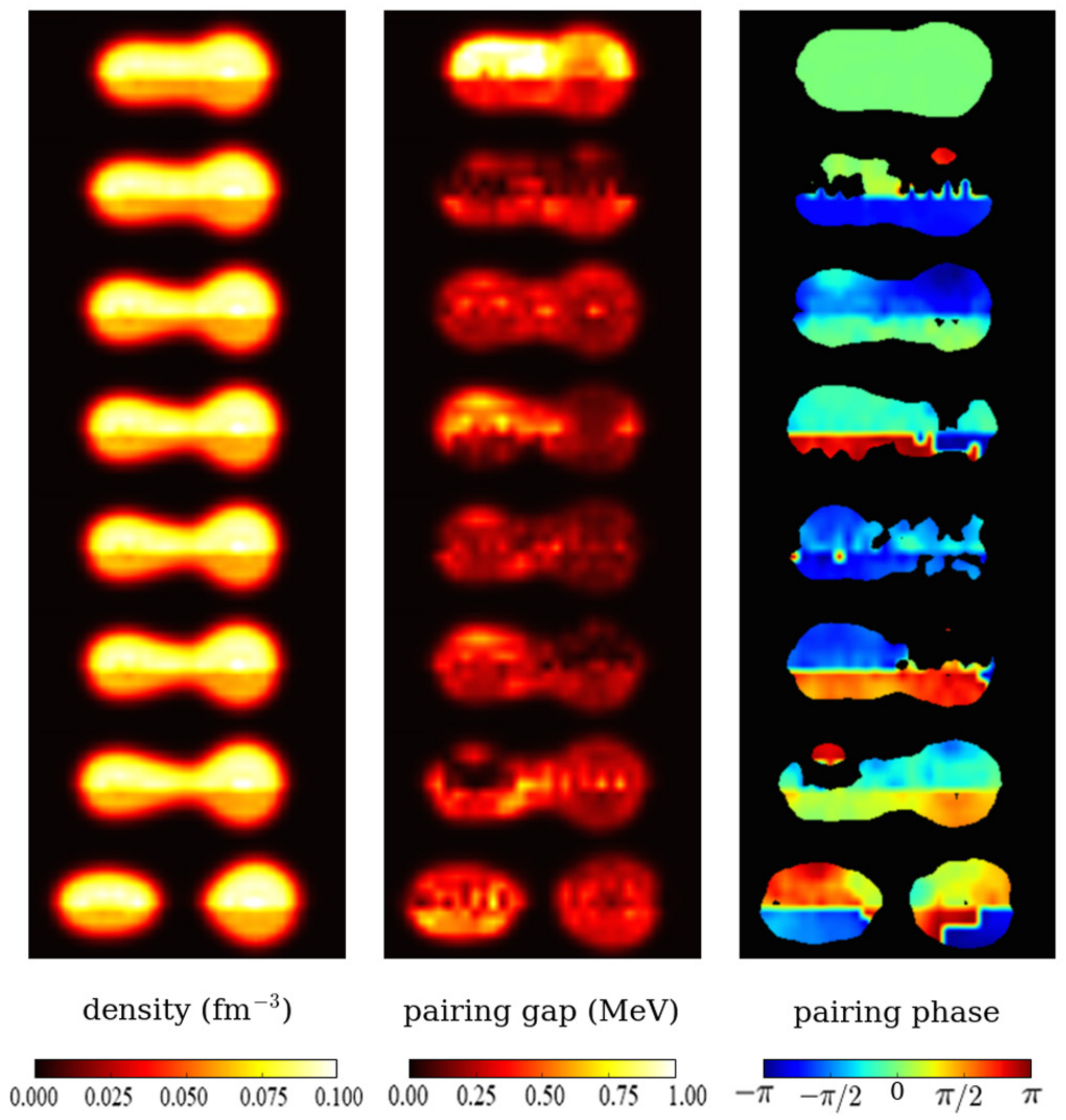
- It is hopeless — too costly to work in practice.

- Fermi-Pauli statistics comes for free (Fock space).

- Not tied to spherical or partially deformed ansätze (full deformation, etc...).

- Transfer learning (train few-nucleon first).

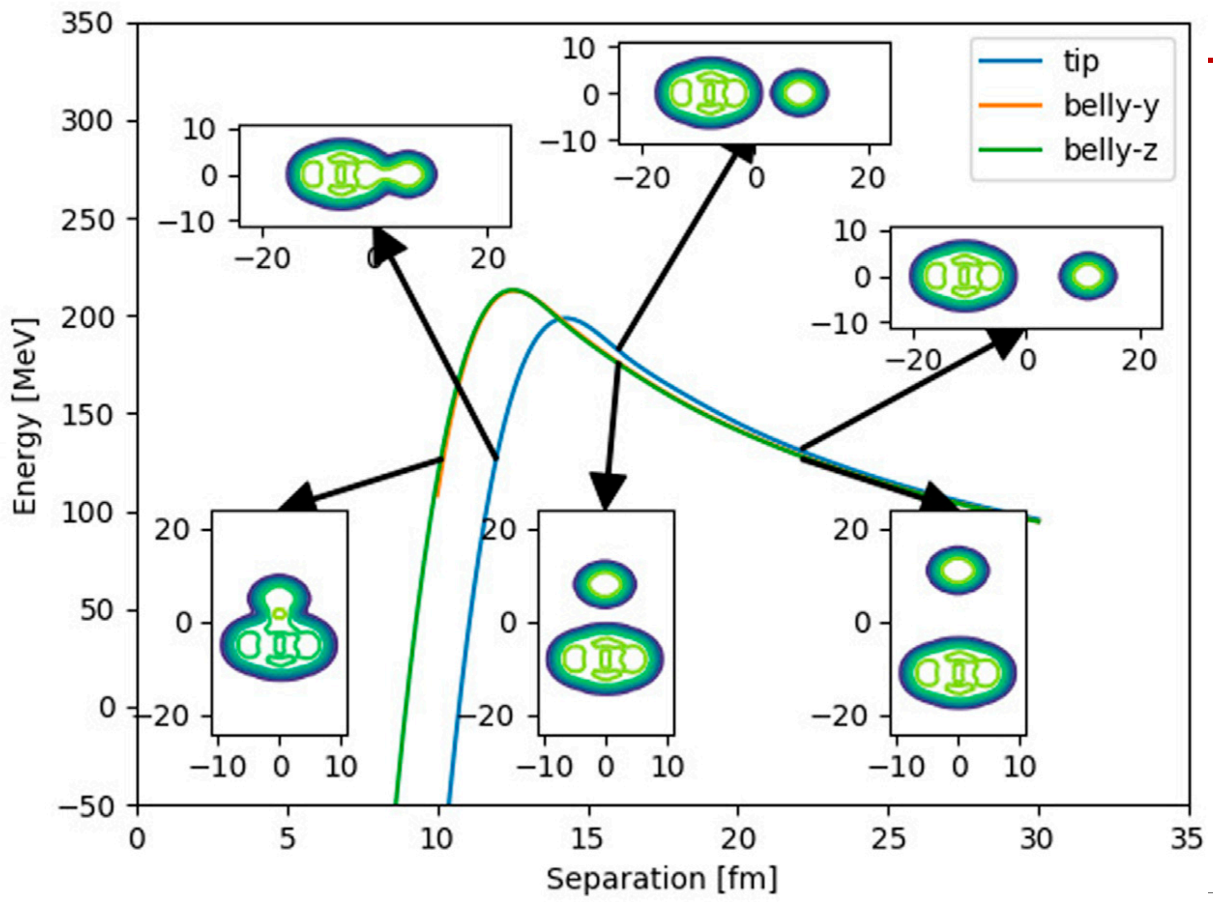
- Many-body dynamics.



Bulgac et al., Phys Rev C **100**, 034615 (2019)

Fission of ^{240}Pu:

- time dependent DFT inspired, in 3D
- 30 x 30 x 60 fm^3 box
- 24 x 24 x 48 = 27,000 pts mesh



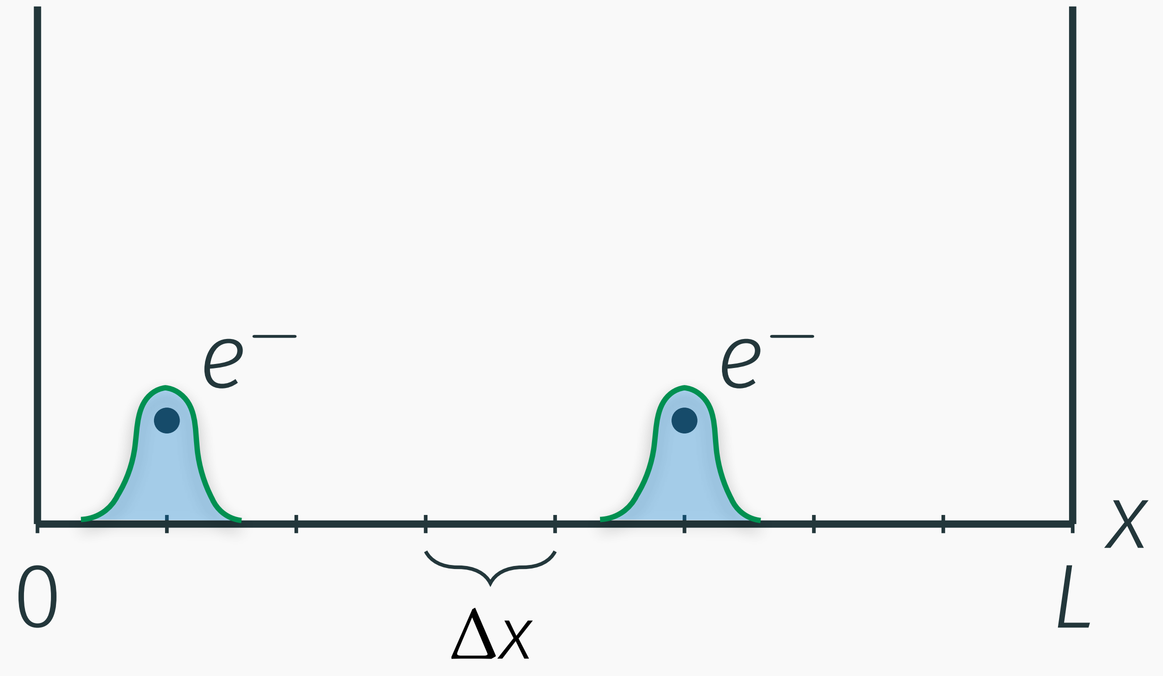
Mean-field simulations of Es-254 + Ca-48 heavy-ion reactions

P. Stevenson et al., Frontiers **10**, 1019285 (2022)

NQS for
fermions confined in a box

Confined fermions w/ a discrete coordinate space mesh

- Discretise coordinate space
- Use occupation number to locate particles



$\Longleftrightarrow$ $(0, 1, 0, 0, 0, 1, 0, 0, 0)$

no need to worry about antisymmetrization!

- Use a Fock space basis to represent particle configurations:

$$|\psi\rangle = \prod_i \psi^\dagger(x_i)\,|0\rangle \;= |n_0{=}0,\, n_1{=}1,\, n_2{=}0,\, n_3{=}0,\, n_4{=}0,\, n_5{=}1,\, \ldots n_L{=}0\rangle$$

- Can be mapped into a system of spins (*with **fixed magnetisation***):

$$\langle x|\psi\rangle \to \langle S|\psi\rangle \Leftrightarrow \begin{cases} c_{\uparrow\uparrow\uparrow\ldots} \doteq \langle\uparrow\uparrow\uparrow\ldots|\psi\rangle = \psi(\uparrow\uparrow\uparrow\ldots) \\ c_{\downarrow\uparrow\uparrow\ldots} \doteq \langle\downarrow\uparrow\uparrow\ldots|\psi\rangle = \psi(\downarrow\uparrow\uparrow\ldots) \\ \quad\vdots \\ c_{\downarrow\downarrow\downarrow\ldots} \doteq \langle\downarrow\downarrow\downarrow\ldots|\psi\rangle = \psi(\downarrow\downarrow\downarrow\ldots) \end{cases}$$

Can be solved as in Carleo and Troyer, Science **355**, 602 (2017)

NQS representation

- Use a Restricted Boltzmann Machine with complex parameter to represent the w.f.:

$$\mathcal{P}(\mathrm{v} \cap \mathrm{h}) = \frac{1}{\mathcal{Z}} \exp\left(\mathrm{a}^{\mathsf{T}}\mathrm{v} + \mathrm{b}^{\mathsf{T}}\mathrm{h} + \mathrm{h}^{\mathsf{T}}\underline{\underline{W}}\,\mathrm{v}\right)$$

with:

$$\begin{cases} \mathrm{v} \in \{-1, 1\}^{N_v} \\ \mathrm{h} \in \{-1, 1\}^{N_h} \end{cases} \qquad \begin{cases} \mathrm{a} \in \mathbb{C}^{N_v} \\ \mathrm{b} \in \mathbb{C}^{N_h} \\ W \in \mathrm{Mat}_{N_h \times N_v}(\mathbb{C}) \end{cases}$$

- Marginalize w.r.t. the hidden nodes:

$$\langle x|\psi\rangle \rightarrow \mathcal{P}(\mathrm{v}) = \sum_{\{\mathrm{h}\}} \mathcal{P}(\mathrm{v} \cap \mathrm{h})$$

$$\begin{cases} \mathrm{a} = \mathrm{a}^{(0)} + \Delta\mathrm{a} = \mathrm{a}^{(0)} + \sum_{i=1}^{N_h} \mathrm{a}^{(i)} \\ 1/\mathcal{Z} = \exp\left(\sum_{i=1}^{N_h} K^{(i)}\right) \\ \mathrm{w}^{(i)} = \begin{pmatrix} W_{1,i} \\ W_{2,i} \\ \vdots \\ W_{N_v,i} \end{pmatrix} \in \mathbb{C}^{N_v}, \\ \omega_i(\mathrm{v}) = \mathrm{w}^{(i)\mathsf{T}}\mathrm{v} + b_i = \sum_j W_{ji} v_j + b_i, \end{cases}$$

Restricted Boltzmann Machine

$$\psi(\mathbf{v}) = 2^{N_h} \exp\left(\mathrm{a}^{(0)\mathsf{T}}\mathrm{v}\right) \prod_{i=1}^{N_h} \left[\exp\left(K^{(i)} + \mathrm{a}^{(i)\mathsf{T}}\mathrm{v}\right) \cosh(\omega_i(\mathrm{v}))\right]$$

<— Note that **a**$^{(i)}$ are site dependent and no hidden nodes are necessary for a single particle (N_h=0).

Energy minimization

- The Hamiltonian for N_v fermions fermions will be:

$$\mathcal{H} = T + V = \sum_i \frac{-\hbar^2}{2m_i(\Delta x)^2} \sum_j \left[\psi^\dagger_{j+1}\psi_j - 2\psi^\dagger_j\psi_j + \psi^\dagger_{j-1}\psi_j\right] + V$$

\+ appropriate conditions at the walls.

- Sample E_{loc} from $|\psi(x)|^2 \sim \mathrm{RBM}^2$ using MCMC:

$$E_{\mathrm{loc}}(\mathsf{x}) = \int d\mathsf{x}'\, \mathcal{H}_{\mathsf{x}\mathsf{x}'} \frac{\psi(\mathsf{x}')}{\psi(\mathsf{x})}$$

$$\frac{\langle\psi|\mathcal{H}|\psi\rangle}{\langle\psi|\psi\rangle} = \langle E_{\mathrm{loc}}\rangle_{|\psi(\mathsf{x})|^2}$$

Use gradient descent w/ SR:

$$D_k(\mathsf{x};\theta) = \frac{\partial_{\theta_k}\psi^\theta(\mathsf{x})}{\psi^\theta(\mathsf{x})}$$

$$\partial_{\theta_k}\langle\mathcal{H}\rangle_\psi = \langle G_k\rangle_{|\psi(\mathsf{x})|^2}$$

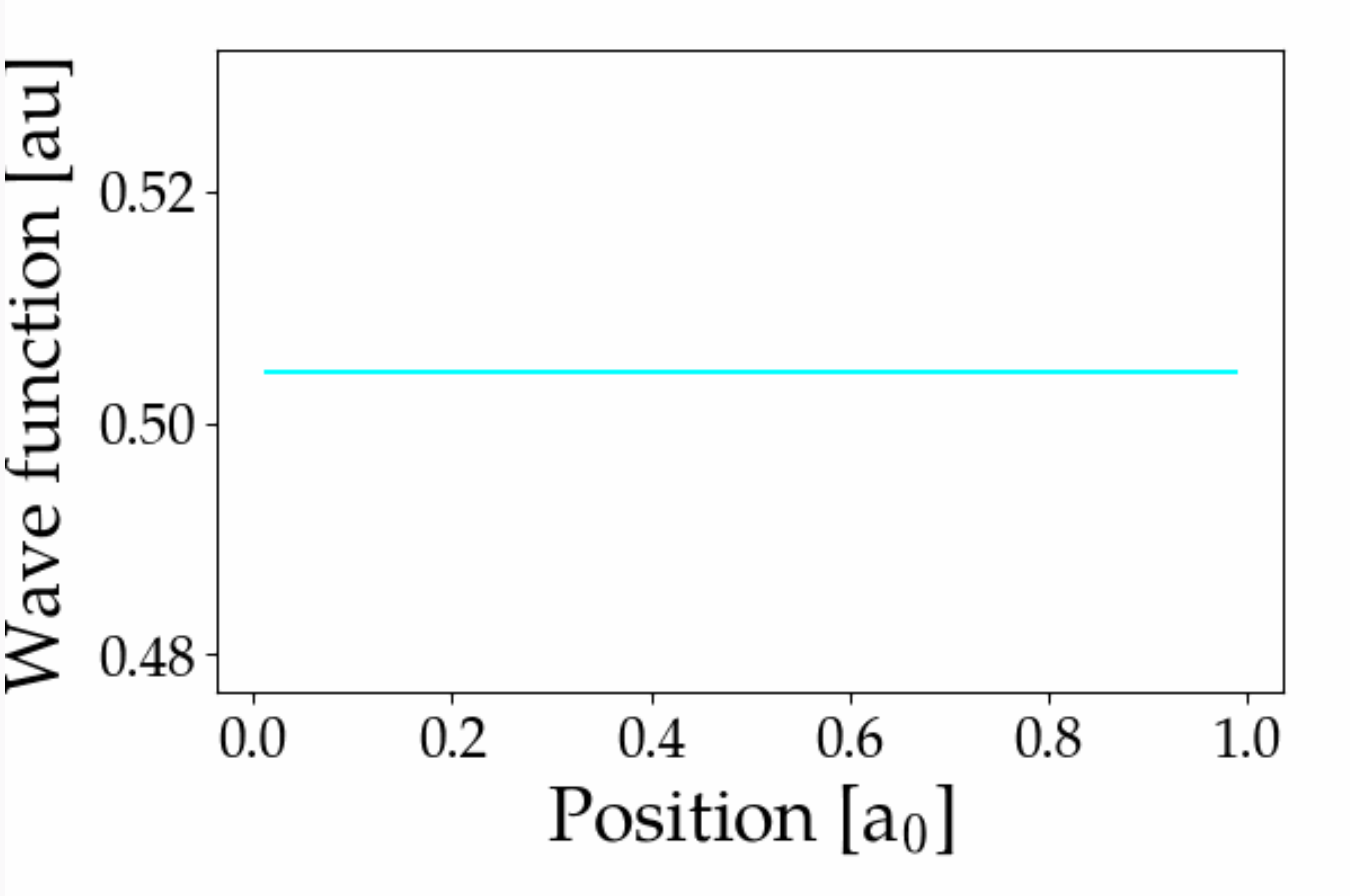
$$G_k(\mathsf{x};\theta) = 2\mathrm{Re}\left[D^*_k(\mathsf{x};\theta)\left(E_{\mathrm{loc}}(\mathsf{x}) - \langle E_{\mathrm{loc}}\rangle_{|\psi(\mathsf{x})|^2}\right)\right]$$

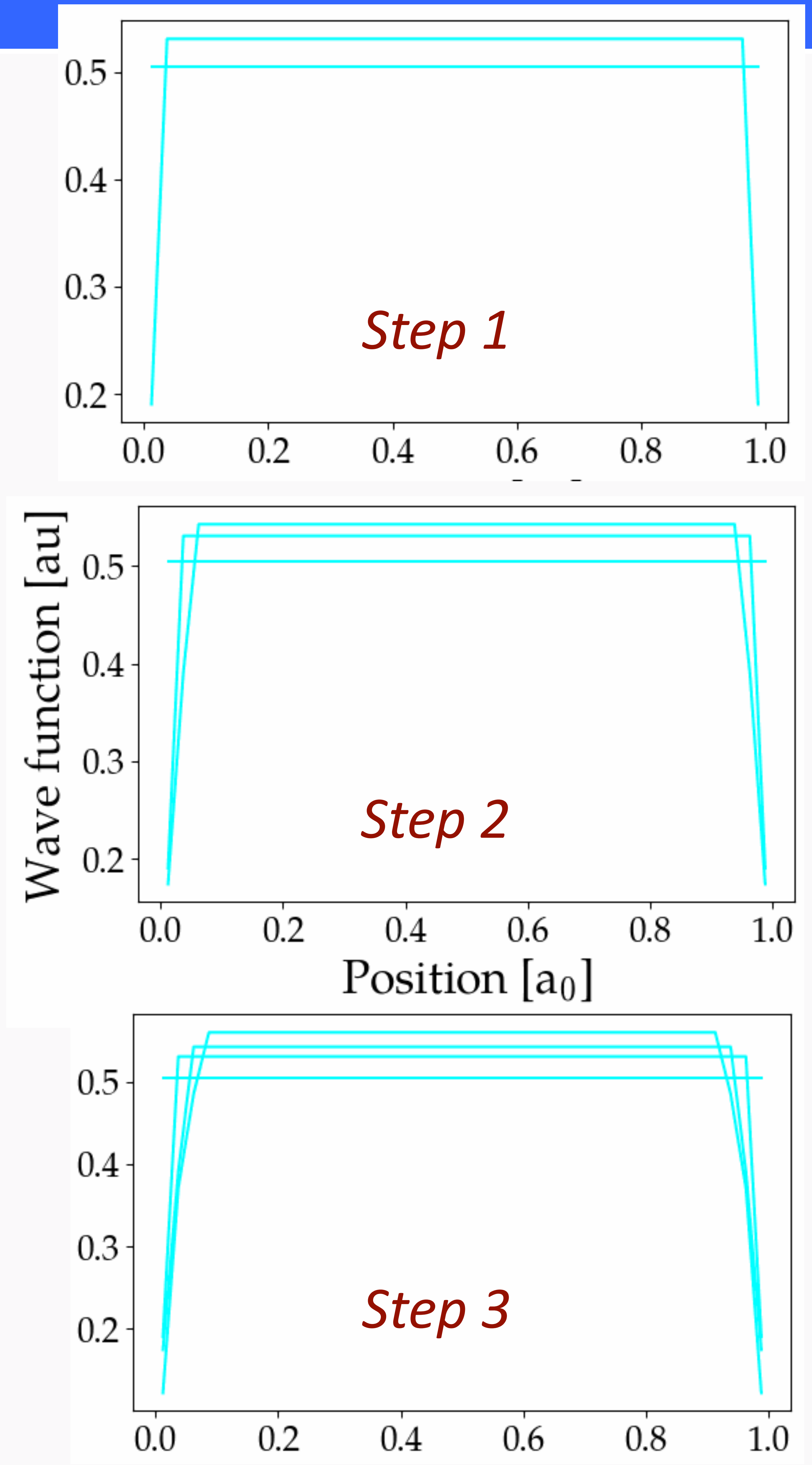
Stochastic Reconfiguration

$$\theta^{(t+1)} = \theta^{(t)} - \eta^{(t)} S^{-1} \nabla_\theta \langle\mathcal{H}\rangle_\psi,$$

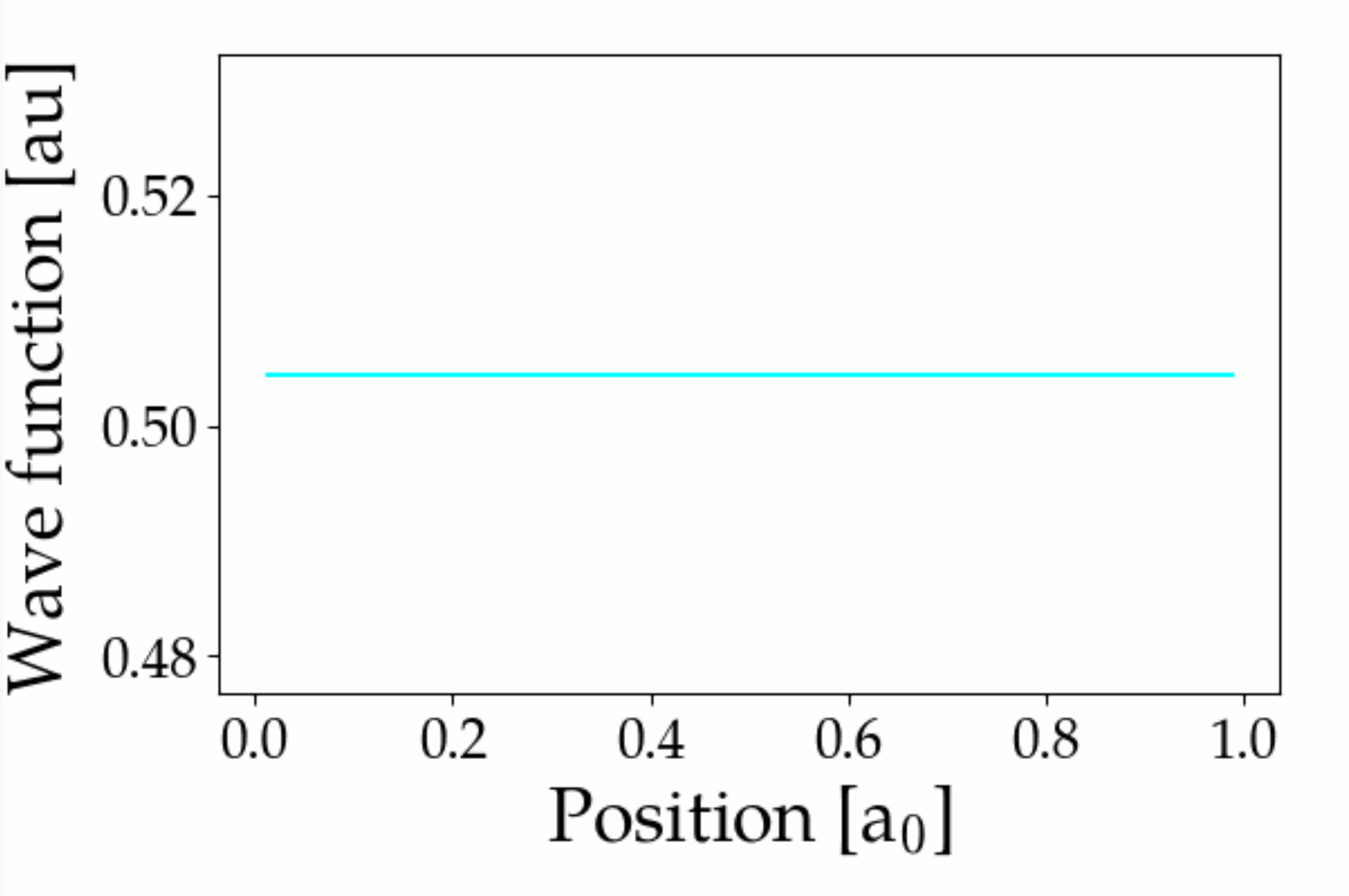
$$S_{ij} = \langle D^*_i\rangle\langle D_j\rangle - \langle D^*_i D_j\rangle,$$

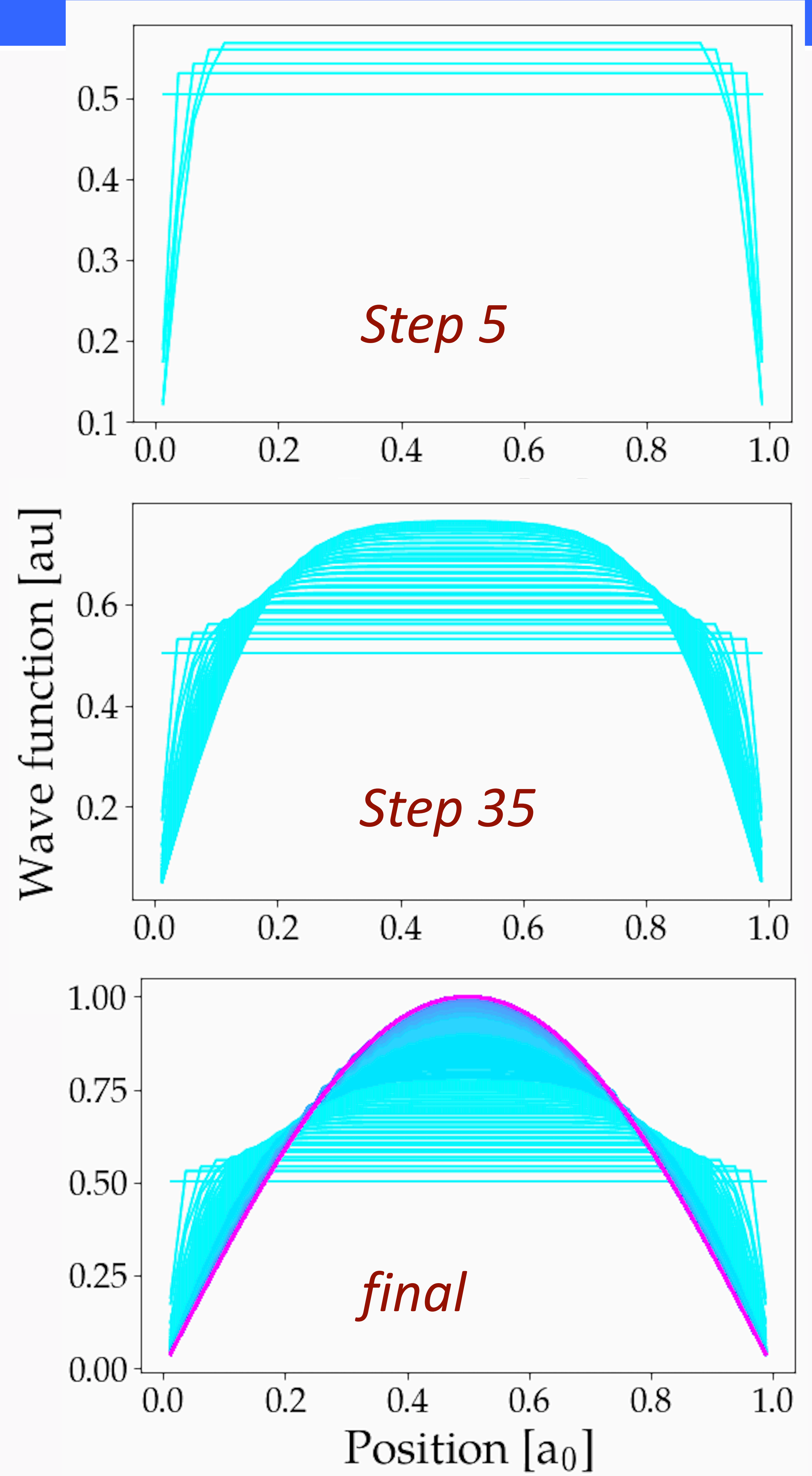
One fermion — optimization





One fermion – optimization





Two fermions — optimization

Convergence with N_v=10

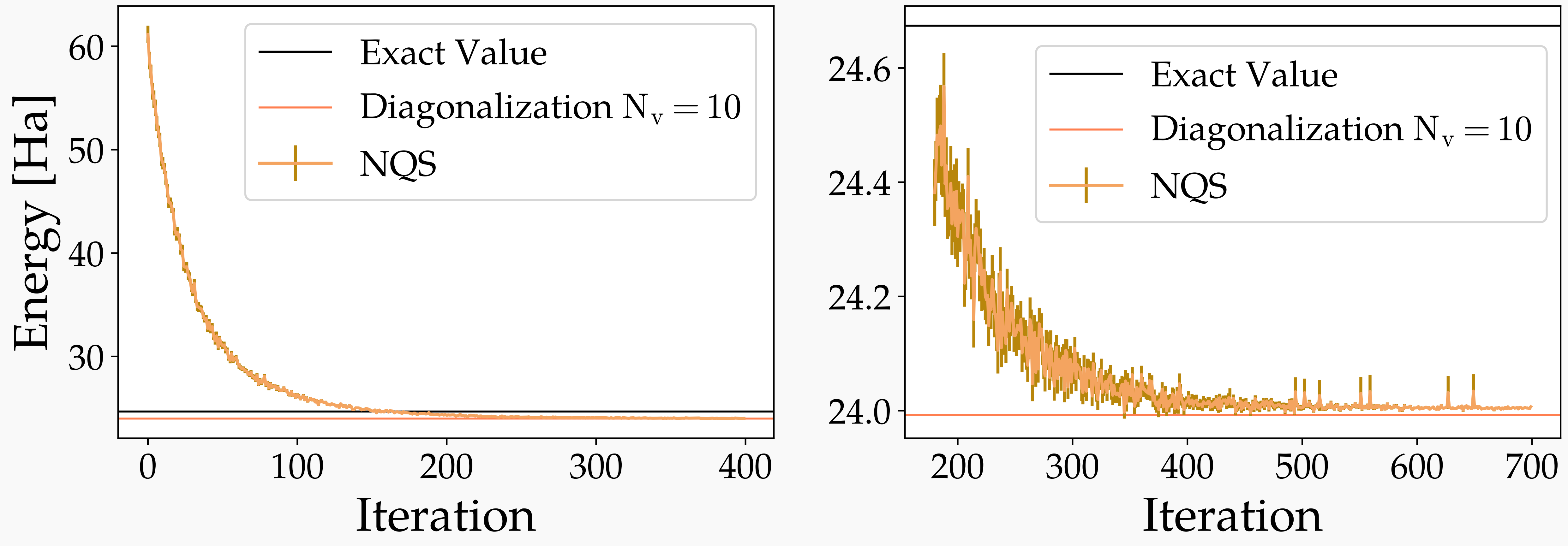


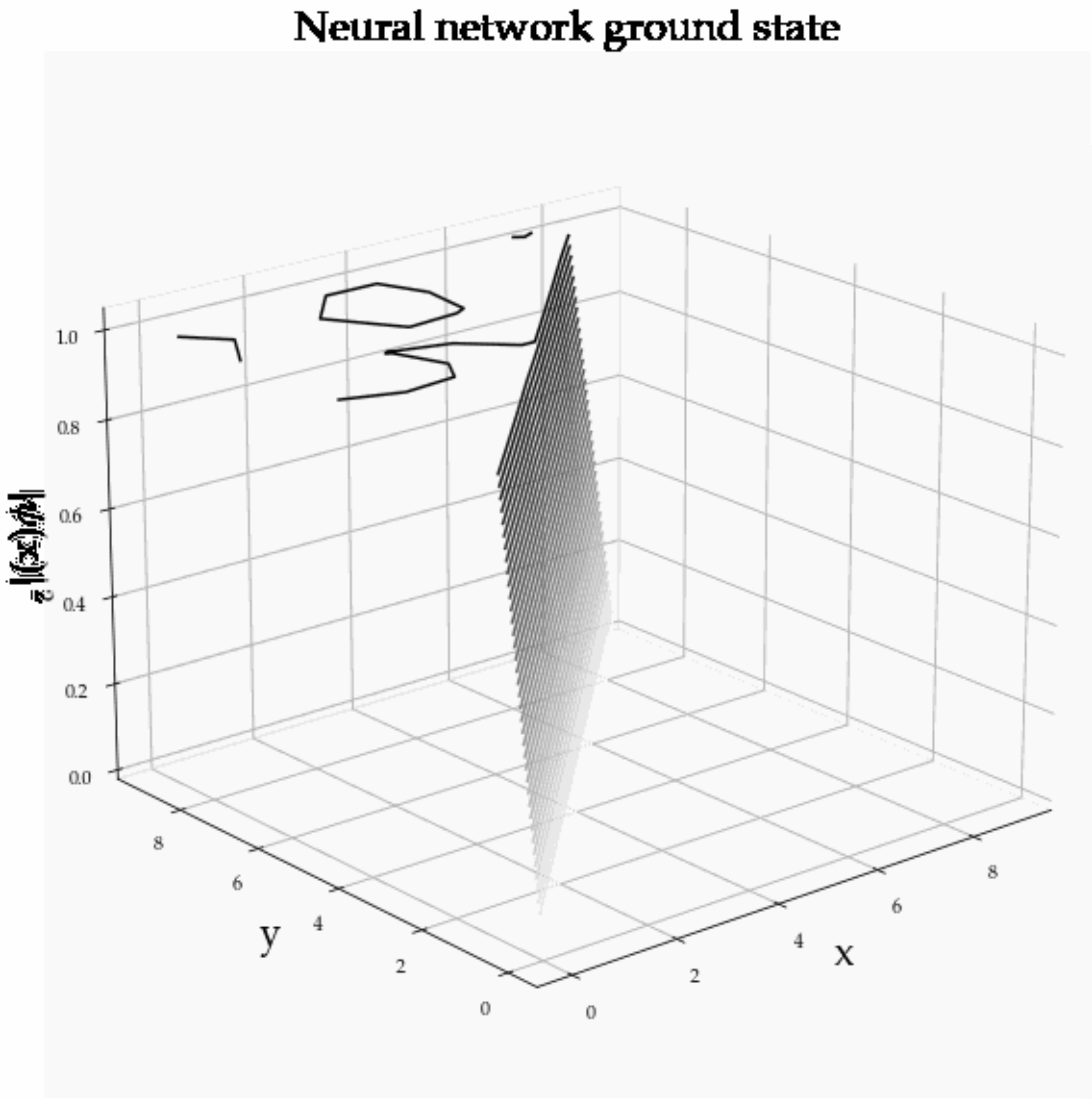
Figure 1: Solution for the two particles non-interacting problem with an RBM built with only 10 hidden nodes. This simulation proves that it is possible to have a satisfying approximation also with relatively few hyper parameters.

Hyperparameter optimization

N_h	$\langle t \rangle$ / iteration	$\Delta E / E$
40	2.75 min	2%
30	1.8 min	2.6%
20	1 min	0.03%
10	16.2 s	0.05%
5	4 s	60%

Table 1: Efficiency and quality of convergence of RBMs with various N_h to model the two non-interacting fermions in a box. N_v is set to 10 for every simulation.

Two fermions — NQS wave function



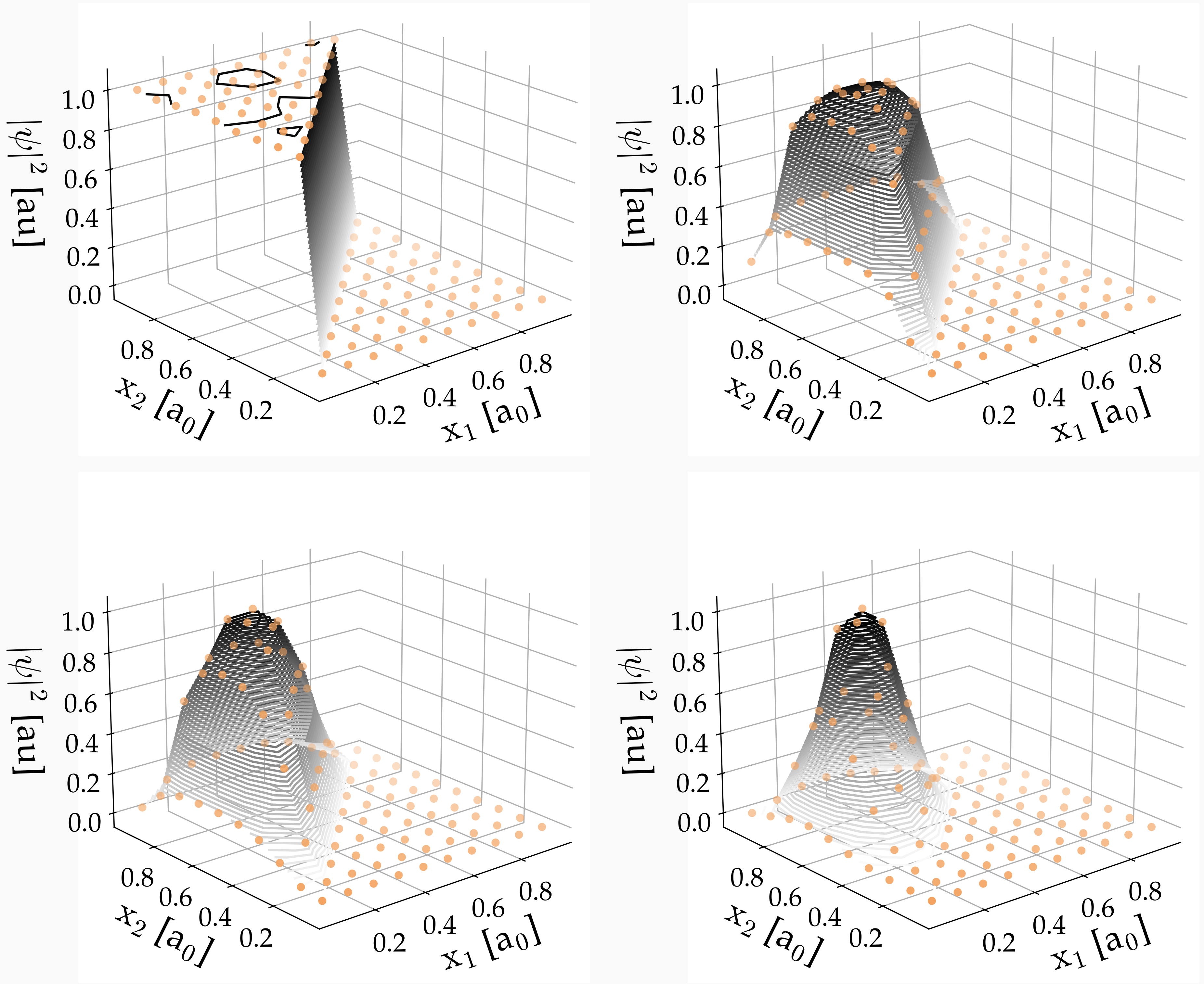


Figure 2: *Physical* learning process of an RBM with $N_h = 20$ from the starting wave function (upper left). The NN seems to be learning boundaries and antisimmetry in the following iterations (number 15, 30 and 300 are reported).

3D lattice 1 and 2 electrons — finite size (FS) tests

Soft-shoulder interactions (Ryberg atoms):

$$\mathcal{V}(r) = \frac{\mathcal{V}_0}{\left(\frac{r}{R}\right)^6 + 1}$$

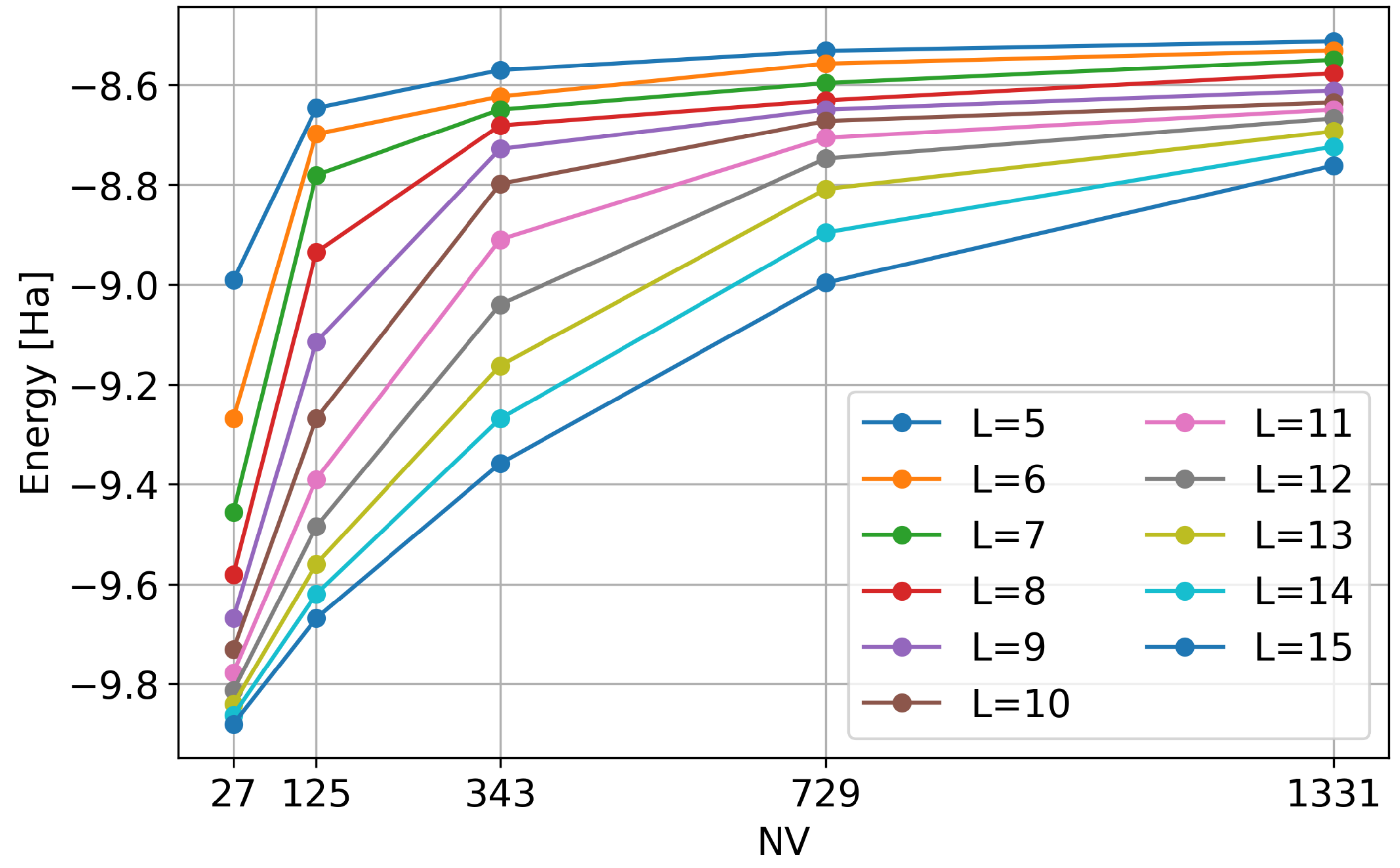


Figure 7.3: Energy as a function of NV, with different trap length values and with $\mathcal{V}_0 = -10$.

Hydrogen atom (Coulomb force):

Long range interaction and cusps…

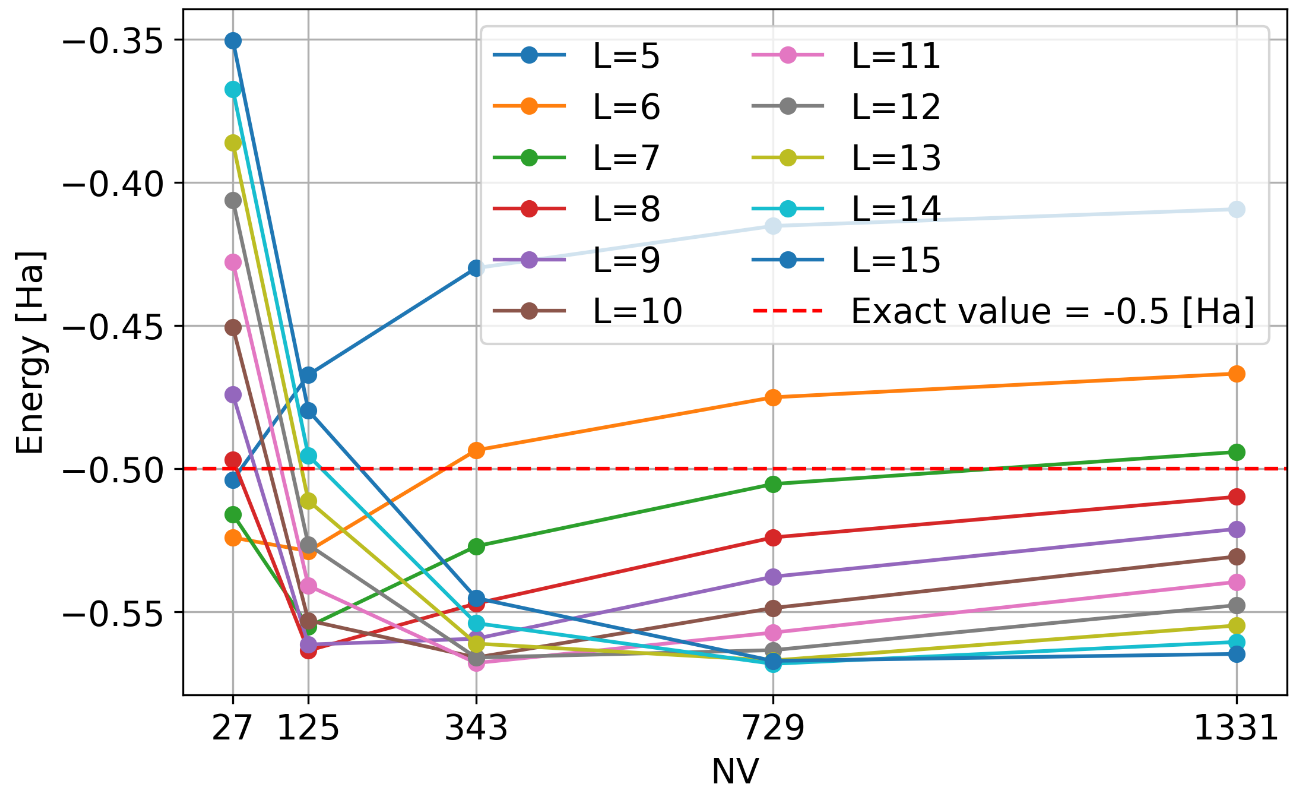
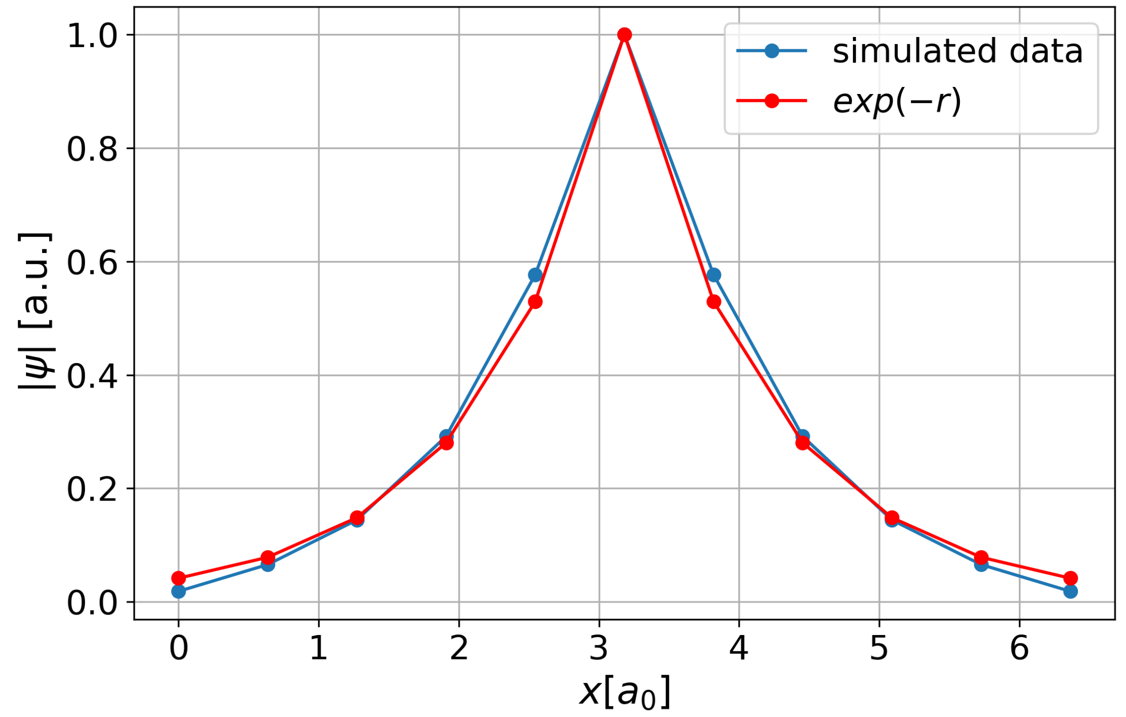


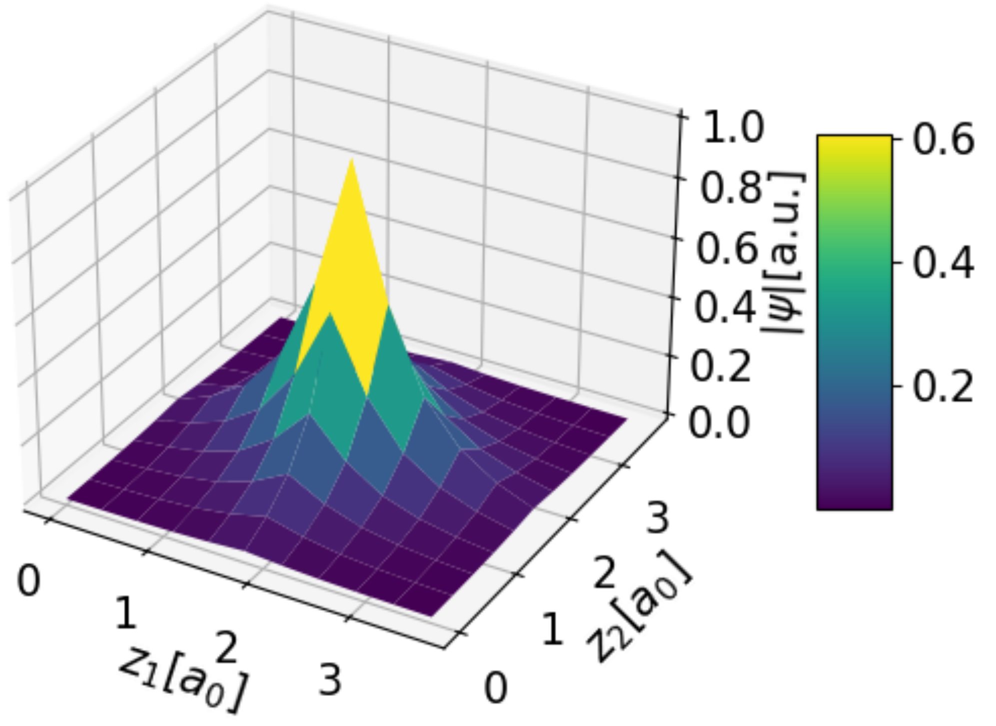
Figure 7.1: Energy as a function of the number of visible nodes for the hydrogen atom, with different trap length values.

Good cusp and 1-body wave function in spite of finite size (FS) effects !!

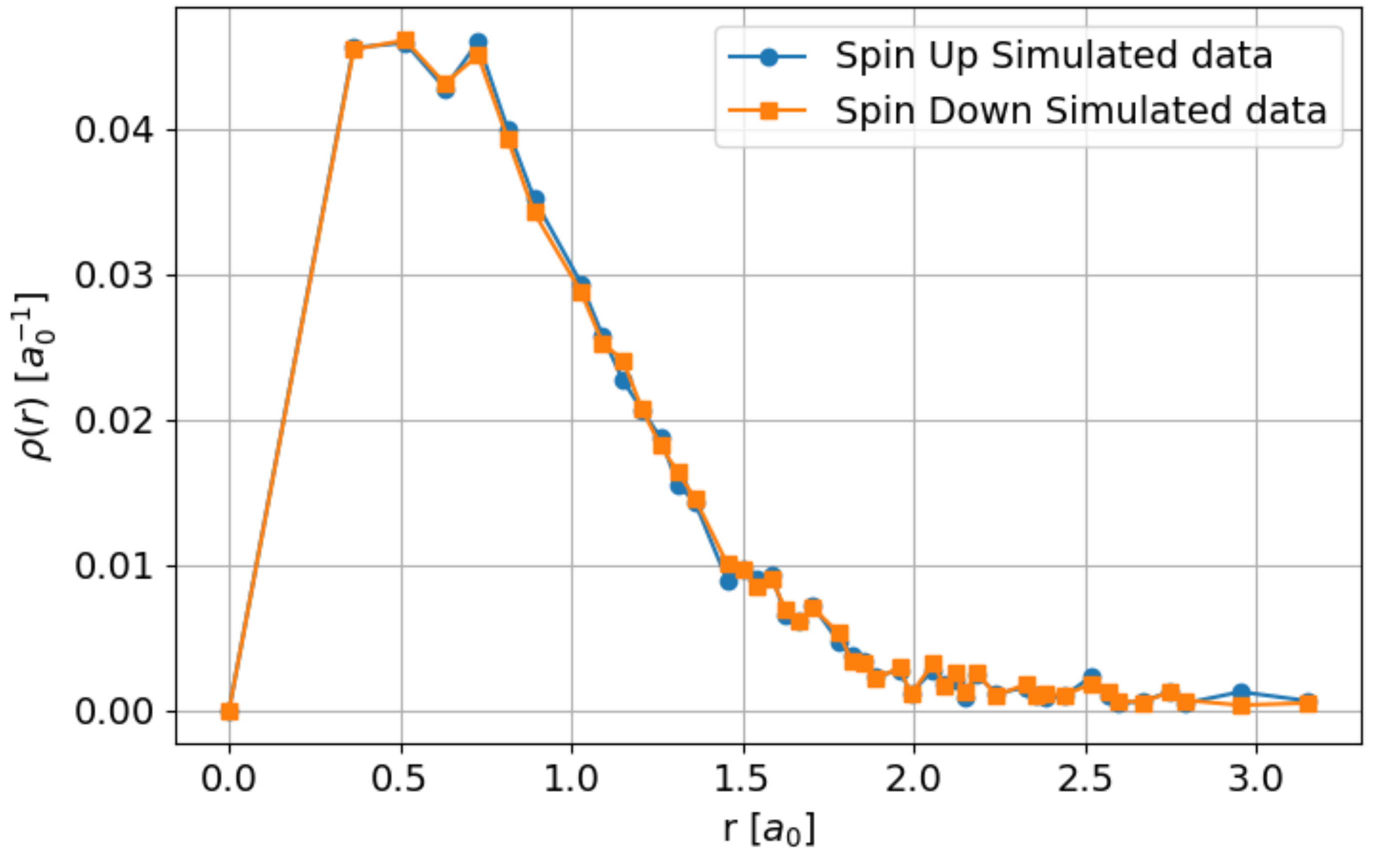


Examples with 2 electrons (Helium atoms)

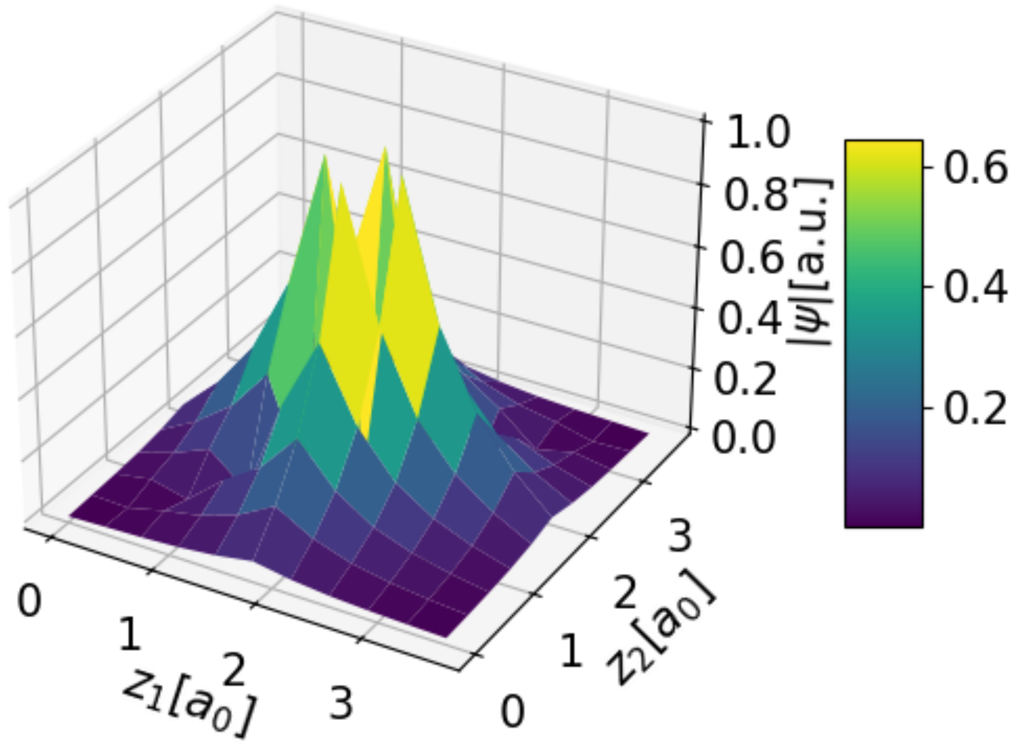
Inverted spins (He g.s.):



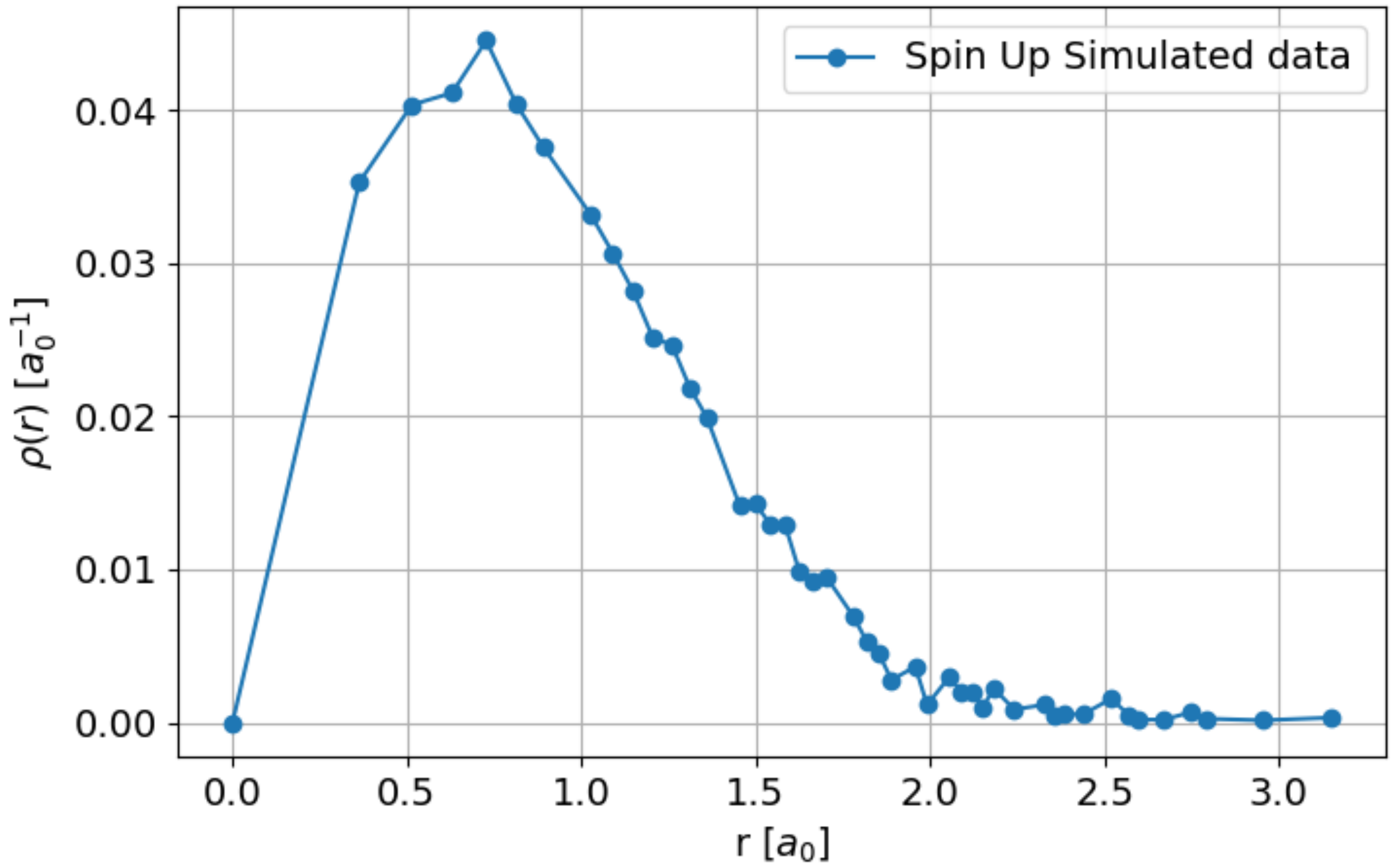
(d) Motion of two electrons along the z-axis with $x_1 = y_1 = x_2 = y_2 = 5$.



Two spins up (He g.s.):



(d) Motion of two electrons along the z-axis with $x_1 = y_1 = x_2 = y_2 = 5$.



Take home messages

- NQS on a lattice are highly challenging. And still at the infancy stage, anyway.
- BUT: they have great potentials (exotic structures, dynamics, etc.)
- Few fermions on small lattices are under reach—proof of principle.
- Transfer learning to heavy systems?? Maybe…
- NQS based VMC only exploits the variational principle (…but see PINN, next talk).

All merits goes to:

L. Lazzarino, **G. Paravizzini, E. Redaelli, G. Borroni**

(all BSc thesis projects…)

Thank you for your attention!!!