

Neural Network Ansatz Solids

How to use this document and notes

Blue text are guiding ideas which outline the purpose of the text below it

Green text are questions

Red text are notes, writing I am unsure of or which needs to be edited / expanded. Generally can be thought of as ‘filler’

The variational monte carlo and neural network Ansatz background sections are largely copied from a previous work and are put here for easy access and understanding. The most relevant and important section currently is the Periodic Systems section (in VMC). The introduction is entirely new and is forming into a sketch of how to frame the paper.

We have a framework which finds the ground-state wave function of a solid. The experiment we have run so far is:

- Lithium solid
- Single primitive cell simulation cell
- Ewald summation potential
- No corrective surface term (Equation 26)
- Periodic boundary conditions as described by the minimum image convention

The bulk energy of the Lithium solid should be -7.54 and our experiment computes -8.1.

The primitive vectors of the BCC unit cell used are

$$P = \begin{pmatrix} -1.0 & 0.0 & 0.0 \\ -1/3 & 0.942 & 0.0 \\ -1/3 & -0.47 & 0.81 \end{pmatrix} \quad (1)$$

Length of a primitive cell vector = 5.608132296572055 Bohr

Other tried:

$$P = \begin{pmatrix} -1/2 & 1/2 & 1/2 \\ 1/2 & -1/2 & 1/2 \\ 1/2 & 1/2 & -1/2 \end{pmatrix} \quad (2)$$

rescaled to atomic units taking the length of a side of the unit cell as 7.4077 Bohr.

1 Introduction

- What are precise electronic structure simulations, why are they important with the example systems

The wave function is an abstract mathematical object mapping configurations of particle states (including properties such as position, momentum, and spin) to amplitudes. Modelling the wave function of a given system is useful because observables such as the particle density and the energy can be computed. The wave functions we want to be able to model are the ground-state wave functions of (fermionic) systems comprised of electrons and nuclei. Finding good approximations to the wave function of these kinds of systems is known as a quantum many-body electronic structure problem and it is valuable to solve because good models of ground-state wave functions of a molecules or solids can inform the design of, for example, pharmaceuticals and batteries, respectively.

- general high level discussion about how hard the problem is followed

However, it is a hard problem. The size of the state space scales exponentially in the degrees of freedom of the system making exact simulation (currently) impossible. Approximate simulations of these fermionic systems suffer the “sign problem”, which is, in general, an NP-hard problem [9] where, for example in Monte Carlo [6] based methods, sampling errors resulting from negative probabilities arise when approximating the partition function. is the reference to approximating the partition function exclusive to Monte Carlo? how does the approximation of the partition function relate to the sampling of configurations in euclidean space?.

- by a sketch of the methods people currently use and some discussion / pros / cons leading into the

importance of the choice of ansatz

There are numerous approaches to this problem including Quantum Monte Carlo (QMC) [3] which scales as $\mathcal{O}(n_e^3)$ or $\mathcal{O}(Z^{5.5})$, where n_e is the number of electrons and Z is the maximum atomic number of the system what are the citations for these, full configuration interaction which is ... but ..., density functional theory which is ... but ..., tensor networks which is ... but ... and coupled cluster methods which are ... but These methods are derived from different approximations and all contain the same fundamental elements: Some way to represent a wave function (called the Ansatz) and some way to make that Ansatz close to the ground-state wave function of the target system..

- what are neural network Ansatz and why they might be a good solution to this problem, development of neural network Ansatz

The Ansatz should be able to accurately represent the true ground-state wave function by capturing the correlations between particles and their degrees of freedom. Neural networks are flexible parameterised functions that are remarkably good at capturing correlations in data and have already been effectively used with QMC methods to find an accurate representations of ground-state wave functions of atoms and molecules [5, 7]. These neural network representations of wave functions draw from traditional Slater-Jastrow-backflow Ansätze and have potential upside over other methods due to their flexibility (the neural network can be easily adapted) and efficiency (they are implemented on highly optimised frameworks with GPU backends).

- specifically outline what problem this paper tackles, how it tackles it, and why we decided to do it this way

In this work, we look to the problem of applying these neural network methods to solid systems where nuclei are fixed in a repeating pattern and electrons can move freely. solids are interesting because.... Currently, researchers can simulate ... and specific problems in the field are

- contributions

2 Related Work

- list of related work and how this work should be understood in relation to those

3 Background

3.1 Variational Monte Carlo

Outline of the schrodinger equation for a fermionic system this section is copied from previous paper

The Schrödinger equation governs the behaviour of quantum systems. In order to solve the it, apart from a handful of analytically solvable models, for example the hydrogen atom [2], one mostly needs to incorporate approximate techniques and numerical methods that scale unfavorably with the increasing system size (e.g. number of electrons). Many techniques and approximations have been introduced to address this problem, and in this work we have a particular focus on real space Monte Carlo approaches.

A fermionic system can be described by the time-independent Schrödinger equation

$$\hat{H}\psi(X) = E\psi(X). \quad (3)$$

where $\psi(X)$ is the wave function, X is a configuration of the system,

$$X = (r_{1,x}, r_{1,y}, r_{1,z}, \dots, r_{n_e,z}; R_{1,x}, \dots, R_{n_n,z}) \quad (4)$$

where $r_{i,j}$ and $R_{i,j}$ correspond to coordinate $j = x, y, z$ of the i -th electron and nuclei, respectively, \hat{H} is the Hamiltonian of the system for which, under the Born-Oppenheimer approximation, the position of nuclei are frozen and we ignore nuclei kinetic energy terms giving the Hamiltonian,

$$\hat{H} = -\frac{1}{2}\hat{\nabla}^2 + V(X). \quad (5)$$

where $\hat{\nabla}^2$ is the electron kinetic energy operator of the Hamiltonian and $V(X)$ is the potential energy of configuration X . The kinetic energy operator is the multidimensional ($3n_e$ where n_e is the number of electrons) Laplacian of the wave function

$$\hat{\nabla}^2 = \sum_{i=1}^{n_e} \sum_{j=x,y,z} \frac{\partial^2}{\partial r_{i,j}^2}. \quad (6)$$

and finally E is the energy of the eigenstate $\psi(X)$. outline variational monte carlo this section is copied from previous paper Approximately solving the Schrödinger equation, Equation 3.1, is a subroutine in finding the minimum energy of the system, i.e.

the ground state energy E_0 . Since the Hamiltonian is a bounded operator, one may use a variational principle [8]

$$E_0 \leq \frac{\langle \psi | \hat{H} | \psi \rangle}{\langle \psi | \psi \rangle} = \frac{\int dX \psi^*(X) \hat{H} \psi(X)}{\int dX \psi^*(X) \psi(X)}, \quad (7)$$

detailing that the expectation value of the Hamiltonian \hat{H} with respect to a state $\psi(X)$ is bounded from below by the ground state energy E_0 . Finding the best approximation to the ground state $\psi_0(X)$ can be done with a parameterized Ansatz, a so-called trial wave function $\psi(X; \theta)$, which is iteratively optimized until a satisfactory accuracy (in terms of energy) is achieved. Different Ansätze have varying capacities to express wave functions, resulting in different possible minimal energy wave functions. The greater the capacity of an Ansatz to model the true wave function, the better the approximation to the ground state, in general.

A popular class of variational methods - Variational Monte Carlo (VMC) - relies on random sampling of the configuration space in order to estimate expectation value of the Hamiltonian (computing loss function) as

$$\mathcal{L}(\theta) = \frac{\langle \psi(\theta) | \hat{H} | \psi(\theta) \rangle}{\langle \psi(\theta) | \psi(\theta) \rangle} = \frac{\int dX |\psi(X; \theta)|^2 E_L(X; \theta)}{\int dX |\psi(X; \theta)|^2}, \quad (8)$$

where $E_L(X; \theta) = \psi^{-1}(X; \theta) \hat{H} \psi(X; \theta)$ is local energy, which for molecular/atomic Hamiltonians is convenient to express in log-domain as

$$E_L(X'; \theta) = -\frac{1}{2} \left[\hat{\nabla}^2 \log |\psi(X; \theta)| \Big|_{X'} + \left(\hat{\nabla} \log |\psi(X; \theta)| \Big|_{X'} \right)^2 \right] + V(X'). \quad (9)$$

$(\hat{\nabla} \cdot)^2$ is the inner product of the nabla operator $\left(\frac{\partial}{\partial r_{1,x}}, \dots, \frac{\partial}{\partial r_{n_e,z}} \right)$ with itself.

The integral (3.1) is an expectation value of the sampled configurations X ,

$$\int dX |\psi(X; \theta)|^2 E_L(X; \theta) = \mathbb{E}_{X \sim p(X; \theta)} [E_L(X; \theta)]. \quad (10)$$

The expectation in Equation (10) is approximated by a Monte-Carlo estimate,

$$\mathbb{E}_{X \sim p(X; \theta)} [E_L(X; \theta)] \approx \frac{1}{N} \sum_{i=1}^N E_L(X_i; \theta), \quad (11)$$

where we introduce configuration probability $p(X; \theta) \propto |\psi(X; \theta)|^2$. Samples (also referred to as walkers and are represented by X) are generated from the wave function distribution via the Metropolis Hastings Monte Carlo method. In order to update the parameters θ and improve the wave function, one needs to compute gradients of the loss function with respect to θ denoted $\Delta \mathcal{L}(\theta)$. The parameters θ of the wave function are optimized using some form of gradient descent and computed via

$$\Delta \mathcal{L}(\theta) = \mathbb{E}_X \left[(E_L(X; \theta) - \mathbb{E}_X [E_L(X; \theta)]) \hat{\nabla} \log |\psi(X; \theta)| \right] \quad (12)$$

and estimated through sampling of the configuration space. This procedure allows us to get close to the ground state $\psi_0(X)$, however it strongly relies on the parameterized Ansatz and ease of computing the gradients $\Delta \mathcal{L}(\theta)$. From now on, we will omit θ parameters where it is clear from the context.

3.2 Periodic Systems

Outline the approximation of the system as periodic why would we need to make approximations for a solid system

It is computationally difficult to compute the local energy, Equation 3.1 for large systems (~ 30 electrons), in particular for methods involving Slater determinants computing the Laplacian scales as $\mathcal{O}(n_e^4)$ this is different to the $\mathcal{O}(n_e^3)$ value I see quoted elsewhere, I am definitely confused on the scaling because they talk about quantum monte carlo scaling at n^3 in general which does not need a slater determinant which is where the quartic estimation comes from. Additionally, the potential energy of a general system of interacting nuclei and electrons is given by

$$V(X) = \sum_{i>j}^{n_e} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} + \sum_{i,I}^{n_e, n_n} \frac{Z_I}{|\mathbf{r}_i - \mathbf{R}_I|} + \sum_{I>J}^{n_n} \frac{Z_I Z_J}{|R_I - R_J|} \quad (13)$$

where n_n is the number of nuclei, Z_I is the atomic number of nuclei I , and \mathbf{r}_i and \mathbf{R}_I are the position vectors of the electron i and nuclei I , respectively. There is something about this sum not being ‘absolutely convergent’ I don’t know what the words are but basically it depends on the order of the sum? Or is this only for the ewalds summation as a result of the cut off?. The size of the systems able to be

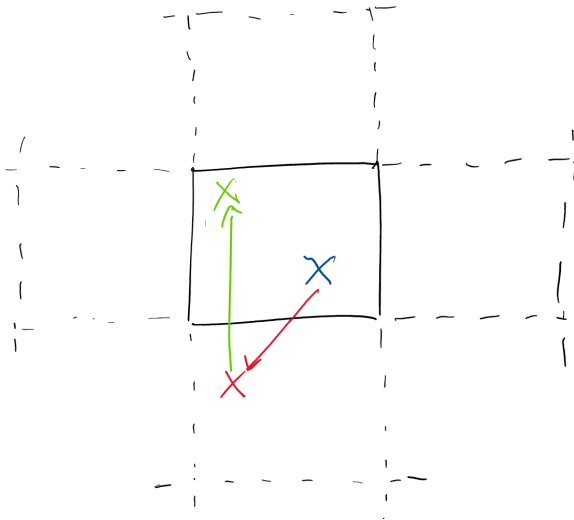


Figure 1: Electrons that leave the simulation cell and enter image cells are moved back into the simulation cell to the corresponding location.

studied remains small without further approximations.

[what approximations are made for a solid system](#)

For a solid system the periodicity can be exploited to approximate the true Hamiltonian. First, we can restrict the system to a small set of particles in a cell (the simulation cell) that sit in an infinitely repeating set of image cell of this system in all directions. [what is a better way of expressing this?](#) All the image cells are identical to the simulation cell with respect to the particles, their properties and the wave function [this is a strange thing to say because the image cells wave function only depend on the images of the simulation particles. Is there a better way to express this?](#)

[describe the minimum image convention](#)
https://en.wikipedia.org/wiki/Periodic_boundary_conditions

The continuity of the wave function is enforced by the minimum image convention, which is comprised of two parts. The first is restricting the electrons to the simulation cell, see Figure 1. The second is defining the interactions between the particles. In the situation outlined in Figure 10 particles interact with an image cell particle instead of the simulation cell particle. Practically, this results in a change of the inputs to the network, altering

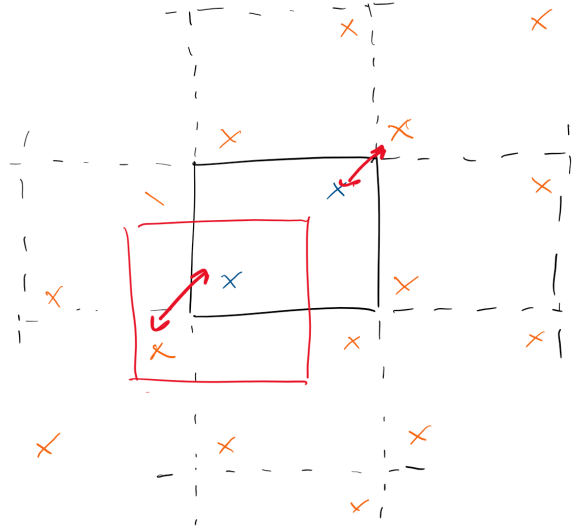


Figure 2: Particles over half the distance (in terms of the basis vectors) from other particles in the simulation cell will interact instead with particles from the image cells. In the situation outlined here, the particle in the bottom left hand corner of the simulation cell interacts with the particle in the top right hand corner of an image cell. A pseudo-simulation box has been drawn in red around the particle to show how the distance to the image cell particle is less than the the distance to the simulation cell particle.

(Equation 29),

$$\mathbf{h}_{ij}^{0\alpha\beta} = (\mathbf{r}_i - \mathbf{r}_j, \|\mathbf{r}_i - \mathbf{r}_j\|) \quad (14)$$

to

$$\mathbf{h}_{ij}^{0\alpha\beta} = (\mathbf{r}'_{ij}, \|\mathbf{r}'_{ij}\|) \quad (15)$$

where

$$\mathbf{r}'_{ij} = \mathbf{r}_i - \mathbf{r}_j - \text{nint}(2\mathbf{r}_{ij}) \quad (16)$$

This equation only works in the case the system has been mapped to a unit cell (with orthogonal unit length basis vectors) and the origin is in the corner.

[Expand on the potential function including mapping to reciprocal space, volume computation](#)

The Hamiltonian of this new approximate system can be computed with the same kinetic energy terms as previously (of the particles in the simulation cell) and by computing the potential energy of each configuration via the Ewalds summation.

The total potential resulting from the interaction of charges in the system is

$$V(X) = \sum_{i=1}^N \sum_{j<i} \frac{q_i q_j}{|\mathbf{r}_{ij}|} \quad (17)$$

$$= \frac{1}{2} \sum_{i=1}^N \sum_{j \neq i} \frac{q_i q_j}{|\mathbf{r}_{ij}|} \quad (18)$$

$$(19)$$

where \mathbf{r}_{ij} are the vectors between two particles. This can be reduced by assuming all the particles in image cells are in the same relative positions as the the particles in the simulation cells and taking only the potential contribution from the simulation cell (giving the per cell potential energy)

$$V(X)/N_s = \frac{1}{2} \sum_{R_s} \sum_{i=1}^n \sum_{j=1}^n \frac{q_i q_j}{|\mathbf{r}_{ij} - R_s|} \quad (20)$$

where $R = \{R_0, R_1, \dots, R_{n_s}\}$ is the set of lattice vectors (displacements of the image cells) and the dash in the first sum indicates the exclusion of $i = j$ terms from the sum when $R = R_0$ (the simulation cell). By decomposing the sum into two sums of the erf and erfc functions in reciprocal and real space, respectively, and correcting for self-interaction terms and the effect of a non-neutral system we have for the interaction between particles i and j

$$\begin{aligned} V_E^{ij}(\mathbf{r}_{ij}) &= \sum_{R_s} \frac{\text{erfc}[\kappa|\mathbf{r}_{ij} - R_s|]}{|\mathbf{r}_{ij} - R_s|} \\ &+ \frac{4\pi}{\Omega} \sum_{G_s \neq 0} \frac{\exp(-G_s^2/4\kappa^2)}{G_s^2} \exp(-iG_s \cdot \mathbf{r}_{ij}) \\ &- \delta_{ij} \frac{2\kappa}{\sqrt{\pi}} - \frac{Q^2}{\kappa^2 \Omega} \end{aligned} \quad (21)$$

where κ is a hyperparameter determining the convergence of both sums and should be set to minimize the size of the sets required to make the sums converge. Ω is the volume of the simulation cell, $G = \{G_0, G_1, \dots, G_{n_r}\}$ is the set of reciprocal lattice vectors, δ_{ij} is the Kronecker delta and Q is the sum of all the charges in the simulation cell. **Where do the neutral correction and the self-interaction correction come from?**

how to compute the reciprocal lattice vectors The reciprocal lattice vectors are all combinations up

to some cut-off of the primitive reciprocal lattice vectors that are computed

$$b_0 = 2\pi \frac{a_1 \times a_2}{V} \quad (22)$$

$$b_1 = 2\pi \frac{a_2 \times a_0}{V} \quad (23)$$

$$b_2 = 2\pi \frac{a_0 \times a_1}{V} \quad (24)$$

where $V = a_0 \cdot (a_1 \times a_2)$. **Some people use the 2π some people don't #science.**

The change of basis

Given a row vector $\mathbf{r} = [r_x, r_y, r_z]$ and a basis where the basis vectors are arranged in the rows

$$P = \begin{pmatrix} e_x^0 & e_y^0 & e_z^0 \\ e_x^1 & e_y^1 & e_z^1 \\ e_x^2 & e_y^2 & e_z^2 \end{pmatrix} \quad (25)$$

discuss the surface term that we have not included

Finally, and this has not yet been implemented, there is a correction for surface effects due to the permittivity of the external medium.

$$V_{\text{surface}}(X) = \frac{2\pi}{3\Omega} \left| \sum_i^N q_i \mathbf{r}_i \right| \quad (26)$$

where it is not clear what the sum goes to. In Reference [11], Equation 6, the implication is that the sum goes to all particles including image cells but that can't make sense soooo.... giving final potential

$$V(X) = V_E(X) + V_{\text{surface}}(X) \quad (27)$$

What are finite size errors and how do we correct for them Finite-size errors are the errors introduced by limiting the size of the simulation cell.

The Ewalds interaction introduces finite size effects dependent on the size and shape of the simulation cell [4]. The model periodic Coulomb interaction produces smaller finite error effects [10]. The interaction only accounts for finite-size errors occurring from the Ewald interaction and **not those from the interacting part of the kinetic energy. It implies this paragraph 1 page 2 in [1] but it's hard to interpret.** There are many methods for estimating the finite-size errors, reference [1] seems to do a good job of exploring them.

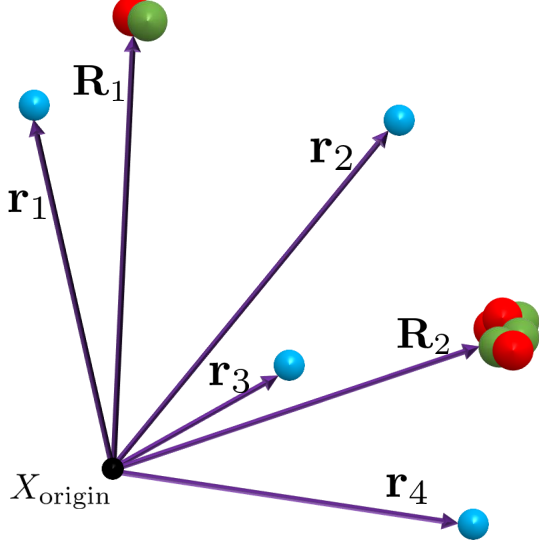


Figure 3: Outline of the system

3.3 Modelling solids

3.4 Neural network Ansatz

There are two sets of streams in the network referred to as the single and pairwise streams. These streams contain the data corresponding to single, $\mathbf{h}_i^{l\alpha}$, and pairwise, $\mathbf{h}_{ij}^{l\alpha\beta}$, electron features, respectively. These variables are indexed by l , the layer of the network, and i and j , the electron indexes. α is the spin of electron i and β is the spin of electron j .

The inputs to the network indexed by $l = 0$ are computed from the system X :

$$\mathbf{h}_i^{0\alpha} = (\mathbf{r}_i - \mathbf{R}_0, \|\mathbf{r}_i - \mathbf{R}_0\|, \mathbf{r}_i - \mathbf{R}_1, \|\mathbf{r}_i - \mathbf{R}_1\|, \dots, \mathbf{r}_i - \mathbf{R}_n, \|\mathbf{r}_i - \mathbf{R}_n\|), \quad (28)$$

$$\mathbf{h}_{ij}^{0\alpha\beta} = (\mathbf{r}_i - \mathbf{r}_j, \|\mathbf{r}_i - \mathbf{r}_j\|). \quad (29)$$

where \mathbf{r}_i and \mathbf{R}_j are the electron and atom position vectors, as shown in Figure 3, and $\|\cdot\|$ is the Euclidean norm. The data from the streams at each layer are transformed by a permutation equivariant function. These functions ensure that exchanging indexes of the data inputs (the electron positions) results in an equivalent exchange of index on the network outputs (before the Slater determinants). The values of the outputs do not change, but the in-

dexes they correspond to change and are the same as the change in indexes of the inputs. As the wave function must be anti-symmetric (required by fermionic systems) these permutation equivariant functions allow the output of the body of the neural network to be fed into a Slater determinant and have all the properties required for a fermionic Ansatz. The inputs to the single stream layer at layer l , $\mathbf{f}_i^{l\alpha}$, are constructed as

$$\mathbf{f}_i^{l\alpha} = \left(\mathbf{h}_i^{l\alpha}, \frac{1}{n_{\uparrow}} \sum_{\beta \neq \downarrow} \mathbf{h}_{ij}^{l\alpha\beta}, \frac{1}{n_{\downarrow}} \sum_{\beta \neq \uparrow} \mathbf{h}_{ij}^{l\alpha\beta} \right) \quad (30)$$

$$\mathbf{g}^l = \left(\frac{1}{n_{\uparrow}} \sum_{\alpha \neq \downarrow} \mathbf{h}_i^{l\alpha}, \frac{1}{n_{\downarrow}} \sum_{\alpha \neq \uparrow} \mathbf{h}_i^{l\alpha} \right), \quad (31)$$

This implementation requires less resources than the original, as the mean over spin terms in the layer, \mathbf{g}^l , are only operated on once, instead of n_e times, in the case that these data are not split and \mathbf{g}^l appears in all single electron streams.

And the updates on the single and pairwise streams at layer l are computed as

$$\mathbf{h}_i^{l\alpha} = \tanh(\mathbf{W}_{(l-1)} \mathbf{f}_i^{(l-1)\alpha} + \mathbf{Z}_{(l-1)} \mathbf{g}^{(l-1)} + \mathbf{b}_{(l-1)}) + \mathbf{h}_i^{(l-1)\alpha} \quad (32)$$

$$\mathbf{h}_{ij}^{l\alpha\beta} = \tanh(\mathbf{V}_{(l-1)} \mathbf{h}_{ij}^{(l-1)\alpha\beta} + \mathbf{c}_{(l-1)}) + \mathbf{h}_{ij}^{(l-1)\alpha\beta} \quad (33)$$

where \mathbf{W}_l , \mathbf{Z}_l , and \mathbf{V}_l are weights and \mathbf{b}_l , and \mathbf{c}_l are the biases. Residual connections are added to all layers where $\dim(\mathbf{h}^l) = \dim(\mathbf{h}^{(l-1)})$. Note, that we need to use a twice-differentiable activation function in order to compute the Laplacian as in Equation (6).

There are n_l of these parameterized layers. The outputs $\mathbf{f}_i^{L\alpha}$ are split into spin dependent data blocks. There is a linear transformation to map $\mathbf{h}_i^{L\alpha}$ to a scalars which are coefficients of the exponentials in the orbitals. Multiple determinants are generated in this way, indexed by k , which contribute to the modelling capacity of the network and the elements of the determinants are the product of the coefficient computed by the Fermi Net and the envelopes

$$\phi_{ij}^{\alpha k} = (\mathbf{w}_{Li}^{\alpha k} \mathbf{h}_j^{L\alpha} + d_{Li}^{\alpha k}) \times \quad (34)$$

$$\sum_m \pi_{im}^{\alpha k} \exp(-|\Sigma_{im}^{\alpha k}(\mathbf{r}_j^{\alpha} - \mathbf{R}_m)|) \quad (35)$$

$\Sigma_{im}^{\alpha k}$ control the anisotropic (direction dependent) behaviour of the envelope whereas the inverse exponential ensures the wave function decays to zero when the electrons are large distances from the nuclei.

The determinants are constructed

$$\det[\Phi^{\alpha k}] = \begin{vmatrix} \phi_{00}^{\alpha k} & \dots & \phi_{0n}^{\alpha k} \\ \vdots & & \vdots \\ \phi_{n0}^{\alpha k} & \dots & \phi_{nn}^{\alpha k} \end{vmatrix}. \quad (36)$$

and the amplitudes are computed from these

$$\psi(X) = \sum_k \det[\Phi^{\uparrow k}] \det[\Phi^{\downarrow k}]. \quad (37)$$

Equation (37) is a representation of the full determinant as a product of spin up and down determinants. This representation forces off-block-diagonal elements of the full determinant to zero, when

$$\begin{aligned} & (i \in \{1, \dots, n_{\uparrow}\} \wedge j \in \{n_{\uparrow} + 1, \dots, n\}) \\ & \vee (i \in \{n_{\uparrow} + 1, \dots, n\} \wedge j \in \{1, \dots, n_{\uparrow}\}) \\ & = (i \leq n_{\uparrow} \wedge j > n_{\uparrow}) \vee (i > n_{\uparrow} \wedge j \leq n_{\uparrow}), \end{aligned} \quad (38)$$

where i and j refer to the orbital index. Finally, the sign is split from the the amplitudes and the network outputs the amplitudes in the log-domain for numerical stability

$$\log |\psi(X)| = \log \left| \sum_k \det[\Phi^{\uparrow k}] \det[\Phi^{\downarrow k}] \right| \quad (39)$$

4 Methods

5 Results

6 Conclusion

References

- [1] Sam Azadi and WMC Foulkes. Systematic study of finite-size effects in quantum monte carlo calculations of real metallic systems. *The Journal of chemical physics*, 143(10):102807, 2015.
- [2] BH Bransden and CJ Joachain. Quantum mechanics,(2000). *Prentice Hall.*, 2000.
- [3] WMC Foulkes, Lubos Mitas, RJ Needs, and Guna Rajagopal. Quantum monte carlo simulations of solids. *Reviews of Modern Physics*, 73(1):33, 2001.
- [4] Louisa M Fraser, WMC Foulkes, G Rajagopal, RJ Needs, SD Kenny, and AJ Williamson. Finite-size effects and coulomb interactions in quantum monte carlo calculations for homogeneous systems with periodic boundary conditions. *Physical Review B*, 53(4):1814, 1996.
- [5] Jan Hermann, Zeno Schätzle, and Frank Noé. Deep neural network solution of the electronic Schrödinger equation. *arXiv:1909.08423*, 2019.
- [6] Nicholas Metropolis, Arianna W Rosenbluth, Marshall N Rosenbluth, Augusta H Teller, and Edward Teller. Equation of state calculations by fast computing machines. *The journal of chemical physics*, 21(6):1087–1092, 1953.
- [7] David Pfau, James S Spencer, Alexander G de G Matthews, and W Matthew C Foulkes. Ab-initio solution of the many-electron Schrödinger equation with deep neural networks. *arXiv:1909.02487*, 2019.
- [8] Attila Szabo and Neil S Ostlund. *Modern quantum chemistry: introduction to advanced electronic structure theory*. Courier Corporation, 2012.
- [9] Matthias Troyer and Uwe-Jens Wiese. Computational complexity and fundamental limitations to fermionic quantum Monte Carlo simulations. *Physical review letters*, 94(17):170201, 2005.
- [10] AJ Williamson, G Rajagopal, RJ Needs, LM Fraser, WMC Foulkes, Y Wang, and M-Y Chou. Elimination of coulomb finite-size effects in quantum many-body simulations. *Physical Review B*, 55(8):R4851, 1997.
- [11] Naomi JK Withey. *Computer simulations of dipolar fluids using Ewald summations*. PhD thesis, University of Reading, 2012.