

Variational Wave Functions for Molecules and Solids

Matthew Foulkes

Department of Physics, Imperial College London

Autumn School on Correlated Electrons: Topology, Entanglement,
and Strong Correlations

21–25 September 2020, Forschungszentrum Jülich, Germany

Why?

To get chemistry right at room temperature, we need total energies to better than

Chemical Accuracy

$$1 \text{ kcal/mole} = 43 \text{ meV/particle}$$

- At room temperature, $k_B T \approx 25 \text{ meV}$.
- Energy differences between competing structures can be $< 10 \text{ meV}$.
- Energy differences between different magnetic structures and different correlated states can be much smaller again.

Our Grand Unified Theory

The Many-Electron Schrödinger Hamiltonian

$$-\frac{1}{2} \sum_i \nabla_{\mathbf{r}_i}^2 - \sum_i \sum_l \frac{Z_l e^2}{|\mathbf{r}_i - \mathbf{d}_l|} + \sum_{i>j} \frac{e^2}{|\mathbf{r}_i - \mathbf{r}_j|} + \sum_{l>j} \frac{Z_l Z_j e^2}{|\mathbf{d}_l - \mathbf{d}_j|}$$

The many-electron Schrödinger equation is NP hard, so the best we can realistically hope for are approximate results.

Energy and Length Scales

Energy scales

Total electronic	$> 10^2$ eV
Chemical bond	few eV
Chemical reaction	10^{-1} eV
Correlations and magnetism	$< 10^{-2}$ eV

Length scale for accuracy of 0.1 eV

$$E/E_F = 10^{-2}$$

$$\lambda/\lambda_F \approx 10$$

$$\lambda \sim 20 \text{ \AA}$$

Density Functional Theory

can sometimes achieve chemical accuracy – if you choose the right functional. But how can you tell in advance?

Perturbation Theory

works well for weakly-correlated materials. But what about strong correlations and non-Fermi-liquid ground states?

Guessing the Wavefunction

was the method responsible for most of our few successes in understanding strongly-correlated non-Fermi-liquid systems:

- BCS
- FQHE
- Bethe Ansatz
- RVB
- Kitaev Model
- ...

Parameterized wavefunction guessing

- Hartree-Fock
- Configuration interaction
- Slater-Jastrow
- Coupled cluster
- Pfaffian/geminal
- DMRG, matrix product states, tensor product states
- RPA
- ...

Variational Monte Carlo

- Guess $\Psi_T(x_1, x_2, \dots, x_N)$, where $x_i \equiv (\mathbf{r}_i, \sigma_i)$.
- Use Monte Carlo integration to evaluate

$$\begin{aligned} E_T[\Psi_T] &= \int \Psi_T^*(x_1, \dots, x_N) \hat{H} \Psi_T(x_1, \dots, x_N) dx_1 \dots dx_N \\ &= \int \left(\frac{\hat{H} \Psi_T(x_1, \dots, x_N)}{\Psi_T(x_1, \dots, x_N)} \right) |\Psi_T(x_1, \dots, x_N)|^2 dx_1 \dots dx_N \end{aligned}$$

where $\int dx \equiv \sum_{\sigma} \int d^3r$.

- Adjust Ψ_T to minimise $E_T[\Psi_T]$.

Non-Interacting Systems

If the Hamiltonian is a sum of terms, one for each electron,

$$\hat{H} = \sum_{i=1}^N \left(-\frac{1}{2} \nabla_{r_i}^2 + V(x_i) \right) = \sum_{i=1}^N \hat{h}_i,$$

the Schrödinger equation

$$\hat{H}\Psi(x_1, x_2, \dots, x_N) = E\Psi(x_1, x_2, \dots, x_N)$$

has separable solutions:

$$\Psi(x_1, x_2, \dots, x_N) = \phi_1(x_1)\phi_2(x_2) \dots \phi_N(x_N).$$

Check

- Substituting $\Psi(x_1, x_2, \dots, x_N) = \phi_1(x_1)\phi_2(x_2) \dots \phi_N(x_N)$ into the Schrödinger equation with $\hat{H} = \hat{h}_1 + \hat{h}_2 + \dots + \hat{h}_N$ gives

$$\begin{aligned} &(\hat{h}_1\phi_1)\phi_2 \dots \phi_N + \phi_1(\hat{h}_2\phi_2) \dots \phi_N \\ &+ \dots + \phi_1\phi_2 \dots (\hat{h}_N\phi_N) = E\phi_1\phi_2 \dots \phi_N \end{aligned}$$

- Dividing by $\phi_1\phi_2 \dots \phi_N$ gives

$$\frac{\hat{h}_1\phi_1(x_1)}{\phi_1(x_1)} + \frac{\hat{h}_2\phi_2(x_2)}{\phi_2(x_2)} + \dots + \frac{\hat{h}_N\phi_N(x_N)}{\phi_N(x_N)} = E$$

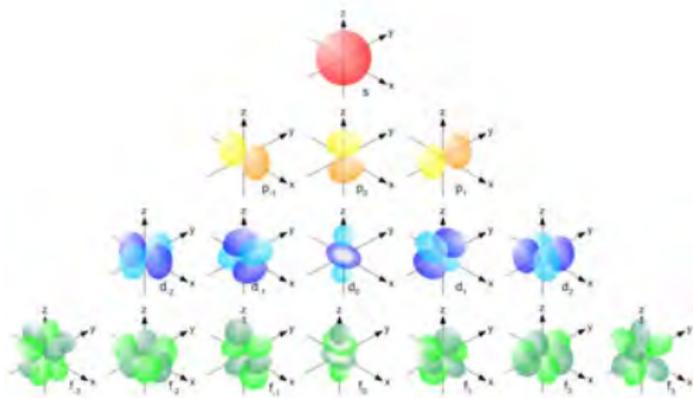
- The first term depends only on x_1 , the second only on x_2 , and so on, but the sum must be the constant E . This is only possible if every term is constant:

$$\hat{h}_1\phi_1 = \epsilon_1\phi_1, \quad \hat{h}_2\phi_2 = \epsilon_2\phi_2, \quad \dots, \quad \hat{h}_N\phi_N = \epsilon_N\phi_N$$

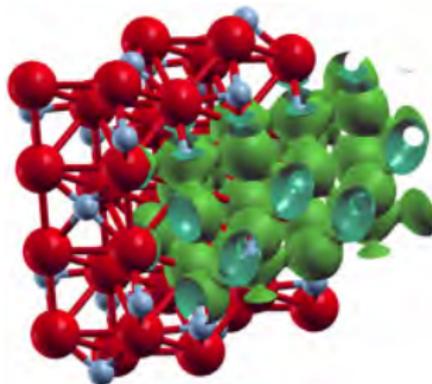
with $E = \epsilon_1 + \epsilon_2 + \dots + \epsilon_N$.

One-Electron Orbitals

Functions such as $\phi_i(x)$, obtained by solving a non-interacting or mean-field Schrödinger equation, are called *one-electron orbitals*.



Hydrogenic orbitals



Fe₂N

- The product state $\Psi(x_1, x_2, \dots, x_N) = \phi_1(x_1)\phi_2(x_2) \dots \phi_N(x_N)$ is not antisymmetric, but we can build an antisymmetric linear combination of degenerate solutions with the N electrons distributed among the N one-electron orbitals in different ways:

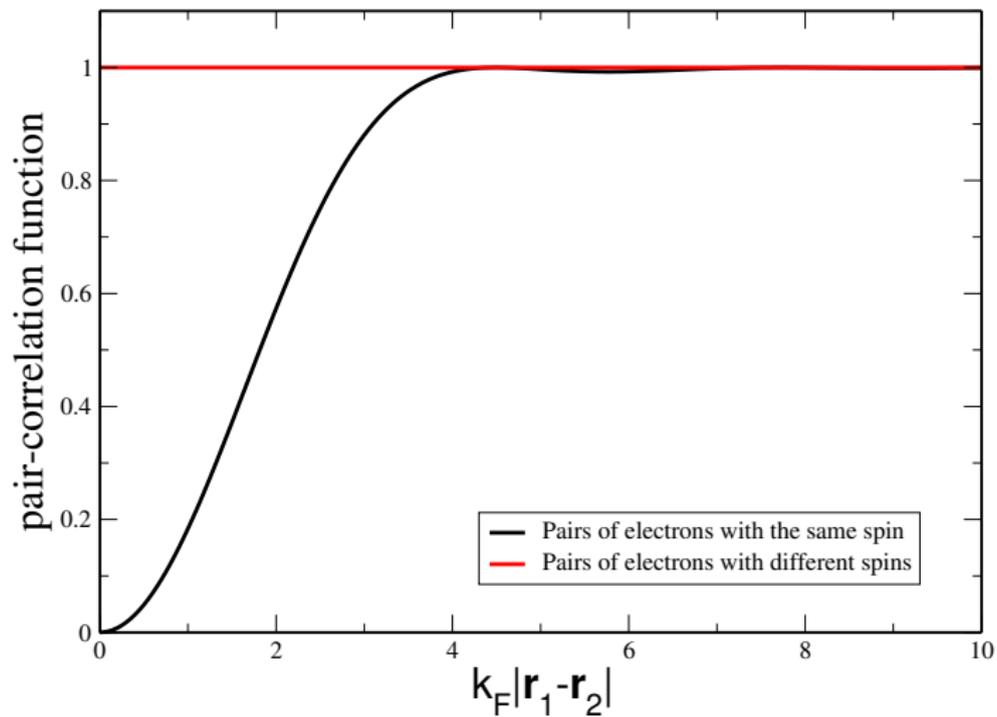
$$\Psi(x_1, x_2, \dots, x_N) = \frac{1}{\sqrt{N!}} \sum_P (-1)^{\xi_P} \phi_{P1}(x_1)\phi_{P2}(x_2) \dots \phi_{PN}(x_N)$$

This may look more familiar written as a

Slater Determinant

$$\Psi(x_1, x_2, \dots, x_N) = \frac{1}{\sqrt{N!}} \begin{vmatrix} \phi_1(x_1) & \phi_1(x_2) & \cdot & \cdot & \cdot & \phi_1(x_N) \\ \phi_2(x_1) & \phi_2(x_2) & \cdot & \cdot & \cdot & \phi_2(x_N) \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \phi_N(x_1) & \phi_N(x_2) & \cdot & \cdot & \cdot & \phi_N(x_N) \end{vmatrix}$$

- If two of $\phi_1, \phi_2, \dots, \phi_N$ are the same, two rows are the same and $\Psi = 0$. This is the Pauli exclusion principle.
- If two electrons of the same spin approach the same point in space, two columns become the same and $\Psi \rightarrow 0$. The antisymmetry helps keep spin-parallel electrons apart.



Interacting Electrons

- The interacting Hamiltonian

$$-\frac{1}{2} \sum_i \nabla_{\mathbf{r}_i}^2 - \sum_i \sum_I \frac{Z_I e^2}{|\mathbf{r}_i - \mathbf{d}_I|} + \sum_{i>j} \frac{e^2}{|\mathbf{r}_i - \mathbf{r}_j|} + \sum_{I>J} \frac{Z_I Z_J e^2}{|\mathbf{d}_I - \mathbf{d}_J|}$$

is *not* separable and Slater determinants are not energy eigenfunctions.

- The Schrödinger equation cannot be solved exactly for systems of more than a few electrons, even using the world's most powerful (classical) computers.

Slater Determinants as Basis Functions

- Given a complete set of one-electron orbitals, $\{\phi_1(x), \phi_2(x), \dots\}$, the set of all products of the form

$$\phi_{i_1}(x_1)\phi_{i_2}(x_2)\dots\phi_{i_N}(x_N)$$

is a complete basis for the N -particle Hilbert space.

- Any antisymmetric wavefunction can be expanded as a linear combination of anti-symmetrized products:

$$\Psi(x_1, x_2, \dots, x_N) = \sum_{\mathbf{i}} C_{\mathbf{i}} D_{\mathbf{i}}(x_1, x_2, \dots, x_N)$$

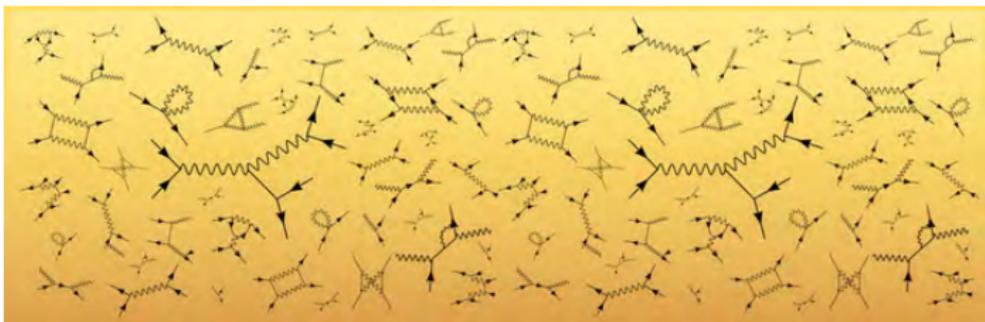
where the sum is over all distinct determinants and $\mathbf{i} = (i_1, i_2, \dots, i_N)$ identifies the orbitals $\phi_{i_1}, \phi_{i_2}, \dots, \phi_{i_N}$ appearing in $D_{\mathbf{i}}$.

Slater Determinants and Second Quantization

- Whenever you use second-quantized notation, you are in fact manipulating Slater determinants:

$$\hat{c}_p^\dagger |D_{i_1, i_2, \dots, i_N}\rangle = |D_{p, i_1, i_2, \dots, i_N}\rangle$$
$$\hat{c}_p |D_{p, i_1, i_2, \dots, i_N}\rangle = |D_{i_1, i_2, \dots, i_N}\rangle$$

- The anti-commutation relations follow from the antisymmetry of the determinants on which the creation and annihilation operators act.



Exchange and Correlation

- Interacting electrons avoid each other:

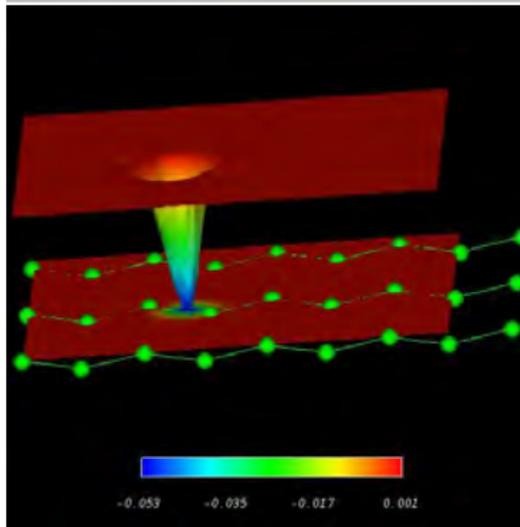
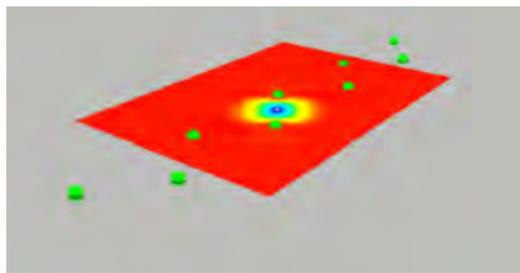
$$n(\mathbf{r}, \sigma; \mathbf{r}', \sigma') d^3 r d^3 r' = \underbrace{n(\mathbf{r}, \sigma | \mathbf{r}', \sigma')}_{\text{conditional probability}} d^3 r \times n(\mathbf{r}', \sigma') d^3 r'$$

- The exchange-correlation hole is defined by:

$$n(\mathbf{r}, \sigma | \mathbf{r}', \sigma') = n(\mathbf{r}, \sigma) + \underbrace{n_{\text{xc}}(\mathbf{r}, \sigma | \mathbf{r}', \sigma')}_{\text{change in density at } r}$$

- Sum rule:

$$\sum_{\sigma} \int n_{\text{xc}}(\mathbf{r}, \sigma | \mathbf{r}', \sigma') d^3 r = -1$$



The Hartree-Fock Approximation

- Many-electron eigenfunctions can in principle be written as linear combinations of (huge numbers of) determinants, but the number required grows exponentially with the number of electrons.
- In the HF approximation, the variational principle is used to find the *single* Slater determinant that best approximates the ground state.
- The overlap of the HF determinant with the ground state is exponentially small in large systems.

Configuration-Interaction Methods

$$\Psi(x_1, x_2, \dots, x_N) = \sum_i C_i D_i(x_1, x_2, \dots, x_N)$$

- Choose $M (\geq N)$ one-electron basis functions.
- Approximate Ψ as a linear combination of the ${}^M C_N$ distinct N -electron determinants.

The vector \mathbf{C} that minimizes

$$E(\mathbf{C}) = \frac{\langle \Psi | \hat{H} | \Psi \rangle}{\langle \Psi | \Psi \rangle} = \frac{\sum_{i,j} C_i^* \langle D_i | \hat{H} | D_j \rangle C_j}{\sum_{\mathbf{k}} C_{\mathbf{k}}^* C_{\mathbf{k}}}$$

satisfies

$$\sum_j H_{ij} C_j = E C_i,$$

where

$$H_{ij} = \langle D_i | \hat{H} | D_j \rangle$$

is an $M C_N \times M C_N$ matrix.

Scaling with System Size

- Suppose a He atom can be described using just four spin orbitals:

$$\phi_{1s}(\mathbf{r})\chi_{\uparrow}(\sigma), \phi_{1s}(\mathbf{r})\chi_{\downarrow}(\sigma), \phi_{2s}(\mathbf{r})\chi_{\uparrow}(\sigma), \phi_{2s}(\mathbf{r})\chi_{\downarrow}(\sigma)$$

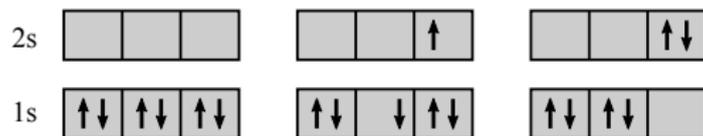
- The FCI basis for N electrons ($N/2$ helium atoms) contains

$${}^M C_N = {}^{2N} C_N = \frac{(2N)!}{N!N!}$$

determinants.

- The Hamiltonian matrix for a collection of just 10 helium atoms ($N = 20$) has more than 10^{11} rows and columns and 10^{22} elements!

Truncated CI Methods



Some of the determinants contributing to the CI singles and doubles wavefunction of 3 He atoms

- Basis set size ($\sim N^2(M - N)^2$) is more manageable.
- No more than two of the $N/2$ He atoms can be found in excited configurations at the same time. What if $N/2 = 100$?
- Energy of $N/2$ well-separated He atoms is larger than $N/2$ times the energy of one He atom. CISD is not *extensive/size consistent*.
- Fraction of correlation energy recovered $\rightarrow 0$ as $N/2 \rightarrow \infty$.

Are CI Methods Useful?

- Full- and truncated-CI methods can produce remarkably accurate results for small molecules.
- CI methods are not useful for solids.
- The full CI quantum Monte Carlo method can be used to sample the components of the vector \mathbf{C} statistically without storing the whole thing.

Coupled-Cluster Methods

- In second-quantized notation, the Hamiltonian of two well-separated He atoms is

$$\hat{H} = \sum_{i,j} \hat{c}_i^\dagger h_{ij} \hat{c}_j + \frac{1}{2} \sum_{i,j,k,l} \hat{c}_i^\dagger \hat{c}_j^\dagger V_{ijkl} \hat{c}_l \hat{c}_k$$

where

$$h_{ij} = \int \phi_i^*(x) \left(-\frac{1}{2} \nabla^2 + V_{\text{nuc}}(\mathbf{r} - \mathbf{d}_A) + V_{\text{nuc}}(\mathbf{r} - \mathbf{d}_B) \right) \phi_j(x) dx$$

$$V_{ijkl} = \iint \phi_i^*(x) \phi_j^*(x') \frac{1}{|\mathbf{r} - \mathbf{r}'|} \phi_k(x) \phi_l(x') dx dx'$$

- If $|\mathbf{d}_A - \mathbf{d}_B|$ is large enough, all matrix elements involving orbitals on different atoms can be neglected:

$$\hat{H} \approx \hat{H}_A + \hat{H}_B$$

- Let $|\psi_A\rangle = \hat{\psi}_A^\dagger|0\rangle$ and $|\psi_B\rangle = \hat{\psi}_B^\dagger|0\rangle$ be the atomic ground states:

$$\hat{H}_A \hat{\psi}_A^\dagger|0\rangle = E_{\text{atom}} \hat{\psi}_A^\dagger|0\rangle$$

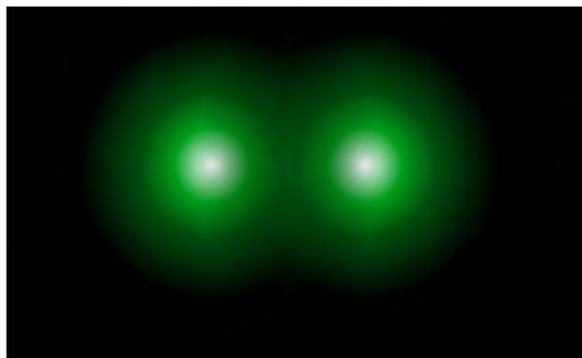
$$\hat{H}_B \hat{\psi}_B^\dagger|0\rangle = E_{\text{atom}} \hat{\psi}_B^\dagger|0\rangle$$

$\hat{\psi}_A^\dagger$ and $\hat{\psi}_B^\dagger$ are linear combinations of products of pairs of creation operators. They commute with each other because they create one-electron orbitals on different atoms.

- Then

$$\begin{aligned} \hat{H} \hat{\psi}_A^\dagger \hat{\psi}_B^\dagger|0\rangle &\approx (\hat{H}_A + \hat{H}_B) \hat{\psi}_A^\dagger \hat{\psi}_B^\dagger|0\rangle \\ &= \hat{\psi}_B^\dagger \hat{H}_A \hat{\psi}_A^\dagger|0\rangle + \hat{\psi}_A^\dagger \hat{H}_B \hat{\psi}_B^\dagger|0\rangle \\ &= \hat{\psi}_B^\dagger E_{\text{atom}} \hat{\psi}_A^\dagger|0\rangle + \hat{\psi}_A^\dagger E_{\text{atom}} \hat{\psi}_B^\dagger|0\rangle \\ &= 2E_{\text{atom}} \hat{\psi}_A^\dagger \hat{\psi}_B^\dagger|0\rangle \quad \checkmark \end{aligned}$$

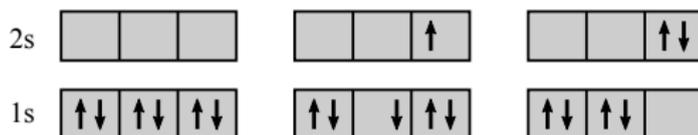
Ground state of two well-separated atoms is $\hat{\Psi}_A^\dagger \hat{\Psi}_B^\dagger |0\rangle$



- Antisymmetrized product of atomic ground states.
- Because of the product form, the number of excitations on atom A does not affect or limit the number of excitations on atom B . Results are size consistent.

Excitation Operators

- Any determinant in the FCI basis can be created from the HF determinant D_0 by making a number of electron-hole excitations of the form $\hat{X}_i^a = \hat{c}_a^\dagger \hat{c}_i$.



- \hat{c}_i annihilates an orbital ϕ_i appearing in D_0
- \hat{c}_a^\dagger creates an orbital ϕ_a that did not appear in D_0
- Creation and annihilation operators for the same orbital never appear, so the \hat{X} operators commute.

Coupled-Cluster Wavefunction

$$|\Psi\rangle = \left[\prod_{a,i} (1 + t_i^a \hat{X}_i^a) \right] \left[\prod_{b>a,j>i} (1 + t_{ij}^{ab} \hat{X}_{ij}^{ab}) \right] \dots |D_0\rangle$$

Coupled-Cluster Wavefunction

$$|\Psi\rangle = \left[\prod_{a,i} (1 + t_i^a \hat{X}_i^a) \right] \left[\prod_{b>a,j>i} (1 + t_{ij}^{ab} \hat{X}_{ij}^{ab}) \right] \dots |D_0\rangle$$

- Terminology:
 - Keep only t_i^a : CC singles.
 - Keep only t_i^a and t_{ij}^{ab} : CC singles and doubles
- For two well-separated atoms, all amplitudes involving orbitals on both atoms are zero. After moving all terms involving atom A to the front of the product, we regain the separable form $\hat{\Psi}_A^\dagger \hat{\Psi}_B^\dagger |0\rangle$. The results are extensive at all truncation levels.

Coupled-Cluster Wavefunction

$$|\Psi\rangle = \left[\prod_{a,i} (1 + t_i^a \hat{X}_i^a) \right] \left[\prod_{b>a,j>i} (1 + t_{ij}^{ab} \hat{X}_{ij}^{ab}) \right] \dots |D_0\rangle$$

- The product of an excitation operator with itself is always zero, so

$$1 + t\hat{X} = 1 + t\hat{X} + \frac{1}{2!}(t\hat{X})^2 + \dots = e^{t\hat{X}}$$

- This allows us to write the CC wavefunction as an exponential:

Exponential Form

$$|\Psi\rangle = \exp \left(\sum_{a,i} t_i^a \hat{X}_i^a + \sum_{a>b,i>j} t_{ij}^{ab} \hat{X}_{ij}^{ab} + \dots \right) |D_0\rangle$$

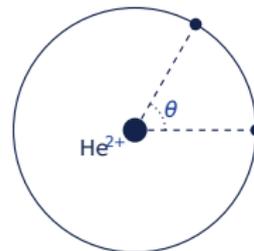
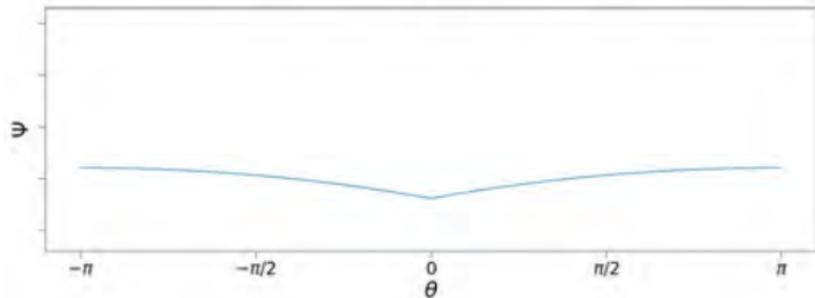
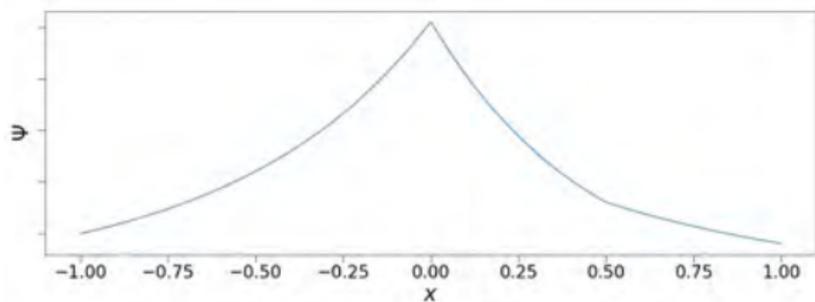
Are CC Methods Useful?

- CC is usually the best method available for medium to large molecules.
- Shows promise in solids, too, but scaling is steep (N^6 for CCSD; N^7 for CCSD(T)) and progress has been slow.
- Struggles with “strongly correlated” problems such as bond breaking.

Water Molecule – Errors relative to CI

method	$R = R_{\text{ref}}$	$R = 2R_{\text{ref}}$
HF	0.217822	0.363954
CCSD	0.003744	0.022032
CCSDT	0.000493	-0.001405
CCSDTQ	0.000019	-0.000446

Cusps



$$\left. \frac{\partial \ln |\Psi|}{\partial r_1} \right|_{r_1=0} = -2.0;$$

$$\left. \frac{\partial \ln |\Psi|}{\partial r_{12}} \right|_{r_{12}=0} = 0.5$$

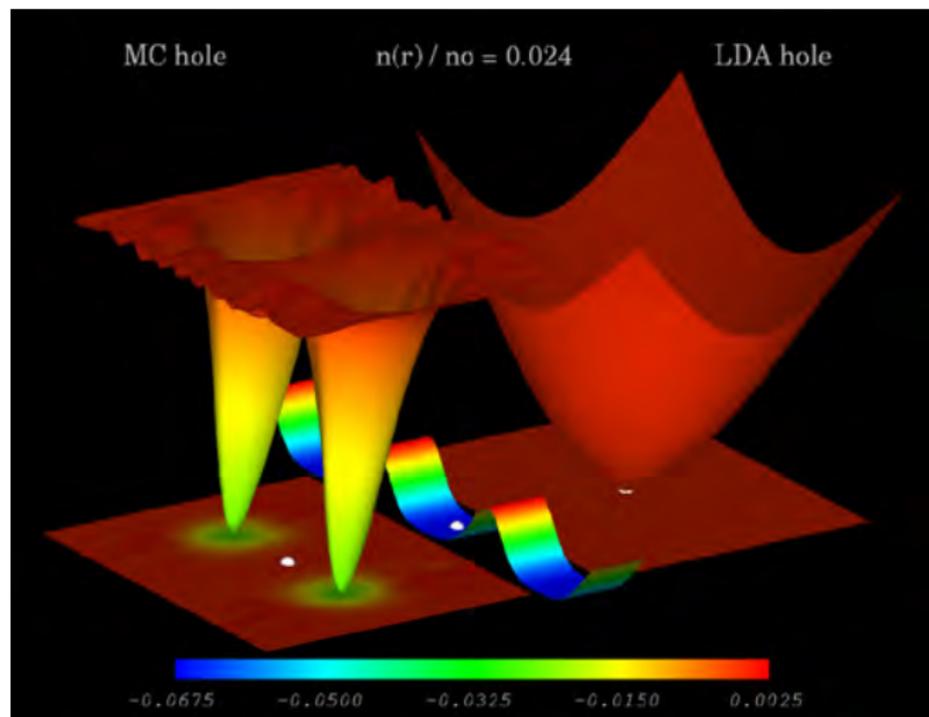
Slater-Jastrow Wavefunctions

Slater-Jastrow wavefunctions offer an easy way to build in the cusps. They also account for some other aspects of correlation.

$$\Psi_T(x_1, x_2, \dots, x_N) = \exp \left(- \sum_{i>j} u(x_{ij}) \right) D$$

$$D = \begin{vmatrix} \psi_1(x_1) & \psi_1(x_2) & \dots & \psi_1(x_N) \\ \psi_2(x_1) & \psi_2(x_2) & \dots & \psi_2(x_N) \\ \cdot & \cdot & \dots & \cdot \\ \cdot & \cdot & \dots & \cdot \\ \psi_N(x_1) & \psi_N(x_2) & \dots & \psi_N(x_N) \end{vmatrix}$$

They work surprisingly well in weakly-correlated solids. Systems of thousands of electrons can be studied using VMC.



Exchange-correlation hole in the sine-wave electron gas

Limitations

- VMC with a SJ trial function is not normally capable of chemical accuracy, even in weakly correlated systems. (DMC is much better.)
- How to extend the SJ form when it doesn't work?
 - Linear combinations of determinants are often used, but this approach does not scale well.
 - No help when Fermi-liquid physics is very wrong.

Observation 1

Artificial neural networks are flexible and efficient function approximators.

Neural QMC

Perhaps we can represent $\Psi(x_1, x_2, \dots, x_N)$ as a deep neural network?

Observation 2

- Nothing requires the orbitals in a Slater determinant

$$D = \begin{vmatrix} \psi_1(x_1) & \psi_1(x_2) & \dots & \psi_1(x_N) \\ \psi_2(x_1) & \psi_2(x_2) & \dots & \psi_2(x_N) \\ \cdot & \cdot & \dots & \cdot \\ \cdot & \cdot & \dots & \cdot \\ \psi_N(x_1) & \psi_N(x_2) & \dots & \psi_N(x_N) \end{vmatrix}$$

to be functions of the coordinates of a single electron.

- The only requirement is that exchanging any two input variables, x_i and x_j , exchanges two columns.

- This allows us to replace the single-electron orbitals $\psi_i(x_j)$ by multi-electron functions

$$\psi_i(x_j; x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_N) = \psi_i(x_j; \{x_{/j}\})$$

- D remains antisymmetric as long as $\psi_i(x_j; \{x_{/j}\})$ is invariant under any change in the order of the arguments after x_j :

$$D = \begin{vmatrix} \psi_1(x_1, \{x_{/1}\}) & \psi_1(x_2, \{x_{/2}\}) & \dots & \psi_1(x_N, \{x_{/N}\}) \\ \psi_2(x_1, \{x_{/1}\}) & \psi_2(x_2, \{x_{/2}\}) & \dots & \psi_2(x_N, \{x_{/N}\}) \\ \cdot & \cdot & \dots & \cdot \\ \cdot & \cdot & \dots & \cdot \\ \psi_N(x_1, \{x_{/1}\}) & \psi_N(x_2, \{x_{/2}\}) & \dots & \psi_N(x_N, \{x_{/N}\}) \end{vmatrix}$$

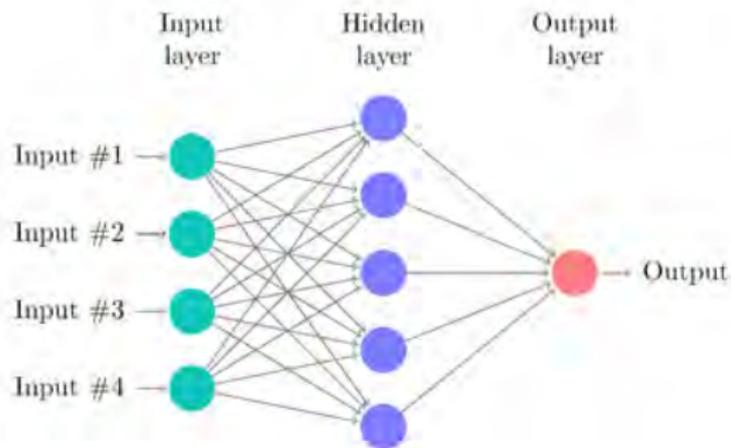
FermiNet

The construction of these permutation-equivariant functions with a neural network is the main innovation of FermiNet.

A Simple Neural Network

$$\mathbf{x}^{\ell+1} = A(W^{\ell} \mathbf{x}^{\ell} + \mathbf{b}^{\ell})$$

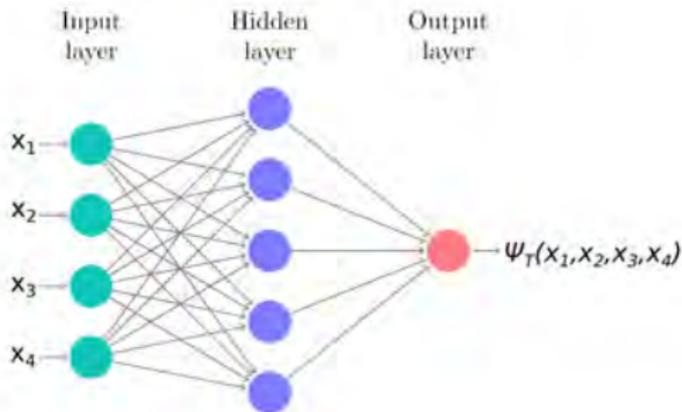
Learn network parameters using automatic differentiation and back propagation.



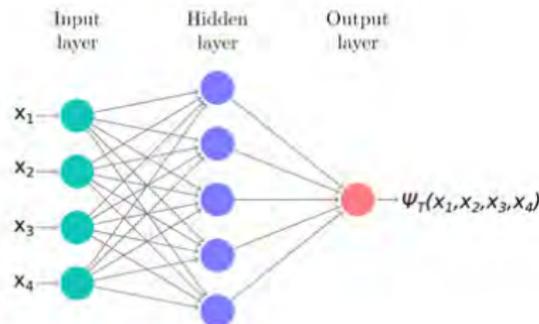
A Simple Neural Network

$$\mathbf{x}^{\ell+1} = A(W^{\ell} \mathbf{x}^{\ell} + \mathbf{b}^{\ell})$$

Learn network parameters using automatic differentiation and back propagation.



Neural QMC



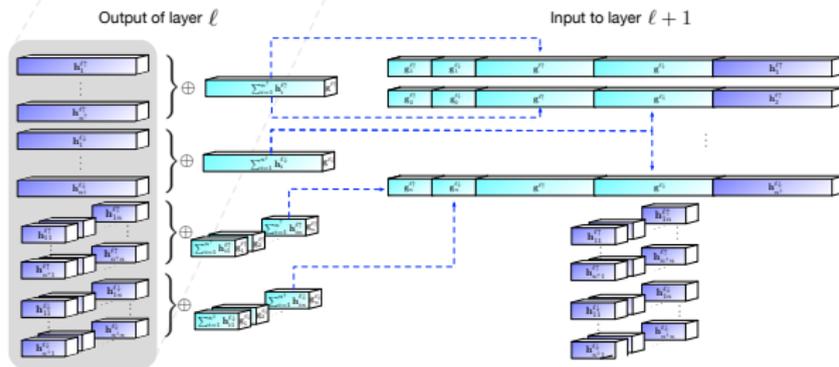
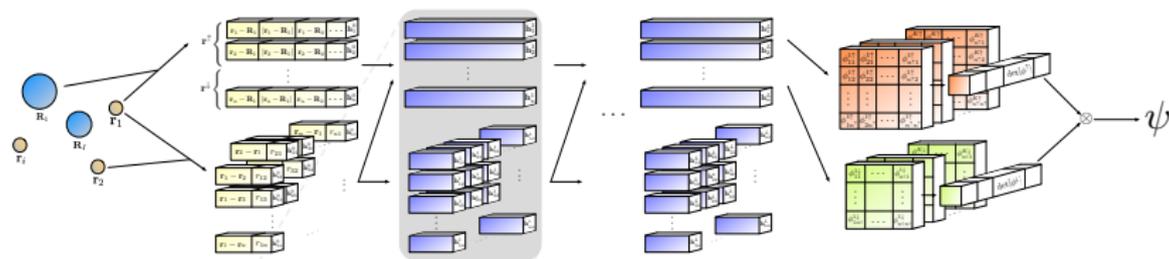
- Sample points $X = (x_1, x_2, x_3, x_4)$ from current $|\Psi_T(X)|^2$ using the Metropolis algorithm.
- Use values of Ψ_T and its derivatives at the sampled points to estimate

$$E_T[\Psi_T] = \int \left(\frac{\hat{H}\Psi_T(x_1, \dots, x_4)}{\Psi_T(x_1, \dots, x_4)} \right) |\Psi_T(x_1, \dots, x_4)|^2 dx_1 \dots dx_4$$

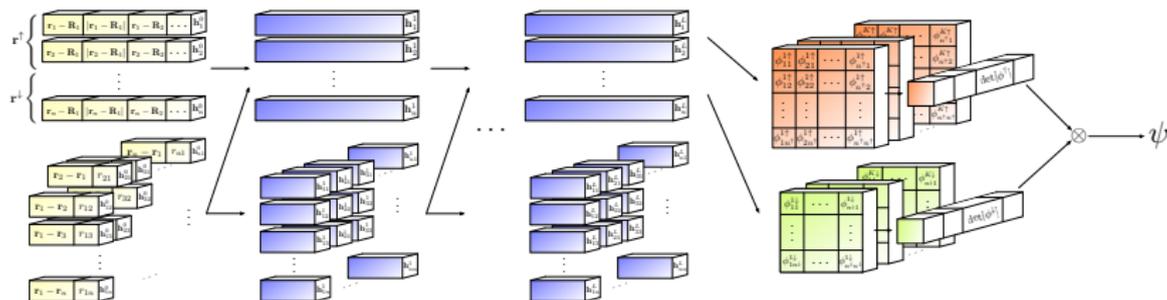
and its first derivatives.

- Adjust network parameters to lower $E_T[\Psi_T]$.

FermiNet



FermiNet Layers



- \mathbf{h}_i^0 is a vector with elements $\mathbf{r}_i - \mathbf{d}_l$ and $|\mathbf{r}_i - \mathbf{d}_l|$ (for all ions l).
 - The elements of \mathbf{h}_i^0 depend on \mathbf{r}_i only.
 - Including $|\mathbf{r}_i - \mathbf{d}_l|$ helps the network learn the electron-nuclear cusps.
- As data propagates through the network

$$\mathbf{h}_i^0 \rightarrow \mathbf{h}_i^1 \rightarrow \dots \rightarrow \mathbf{h}_i^L$$

information about the positions of other electrons is mixed in such that

$$\mathbf{h}_i^L = \mathbf{h}_i^L(\mathbf{r}_i, \{\mathbf{r}_{/i}\})$$

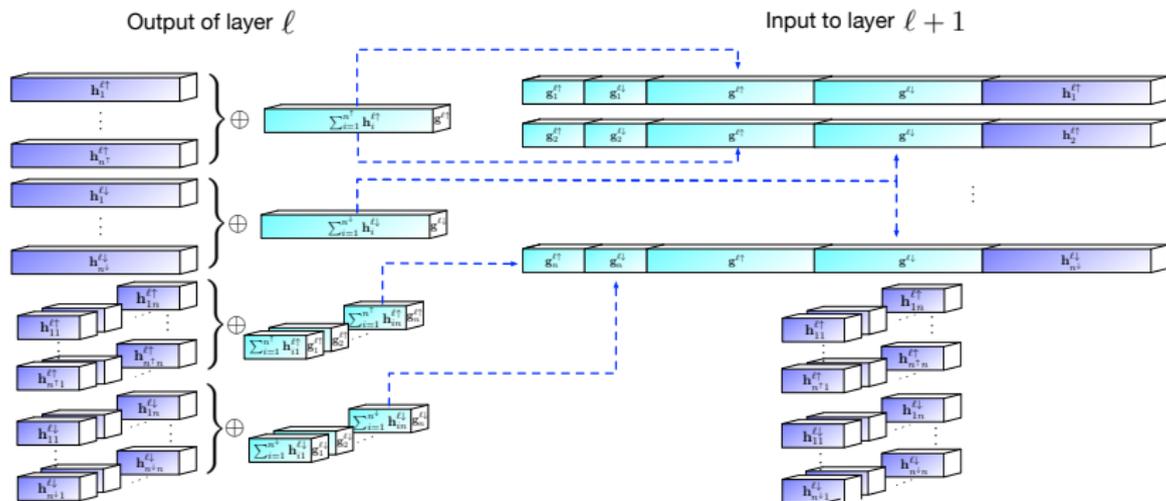
- A final linear transformation gives

$$\psi_j(\mathbf{r}_i, \{\mathbf{r}_{/i}\}) = \mathbf{w}_j \cdot \mathbf{h}_i^L(\mathbf{r}_i, \{\mathbf{r}_{/i}\}) + g_j$$

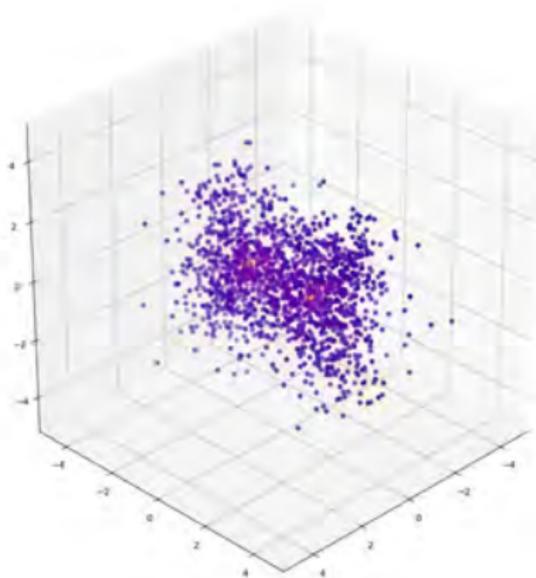
after which the determinant is evaluated.

- We often use linear combinations of these (non-orthogonal) generalised determinants, all optimised together.

FermiNet Layer Inputs

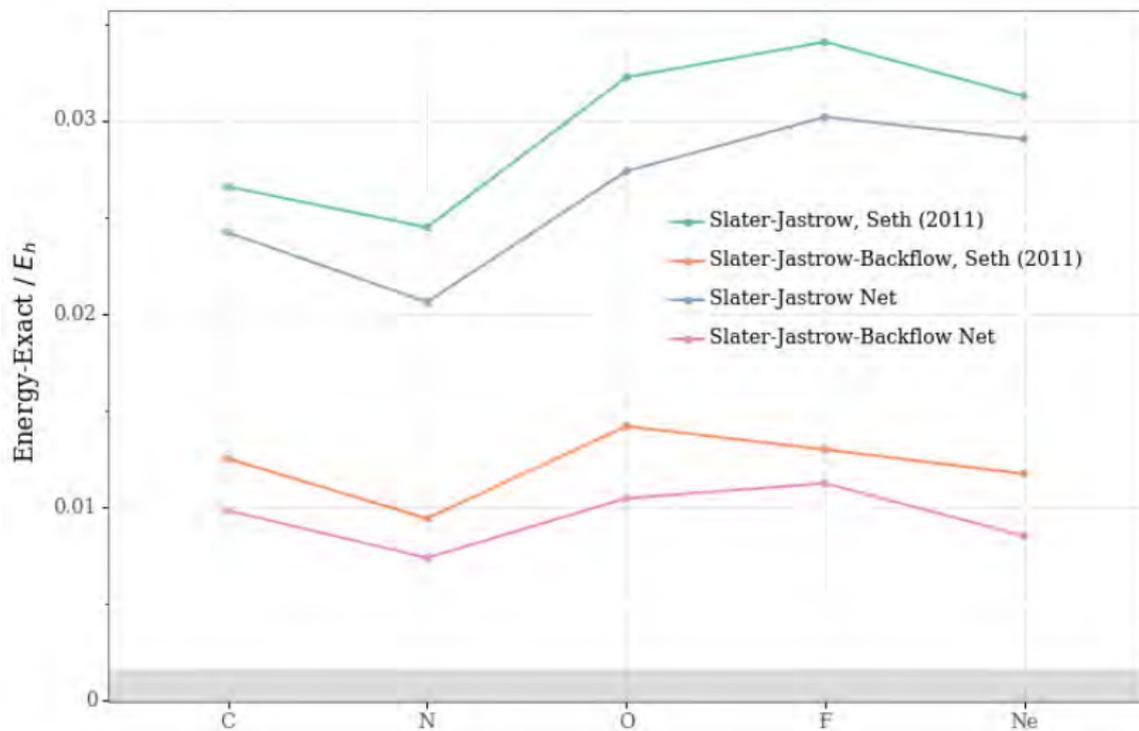


Ethene



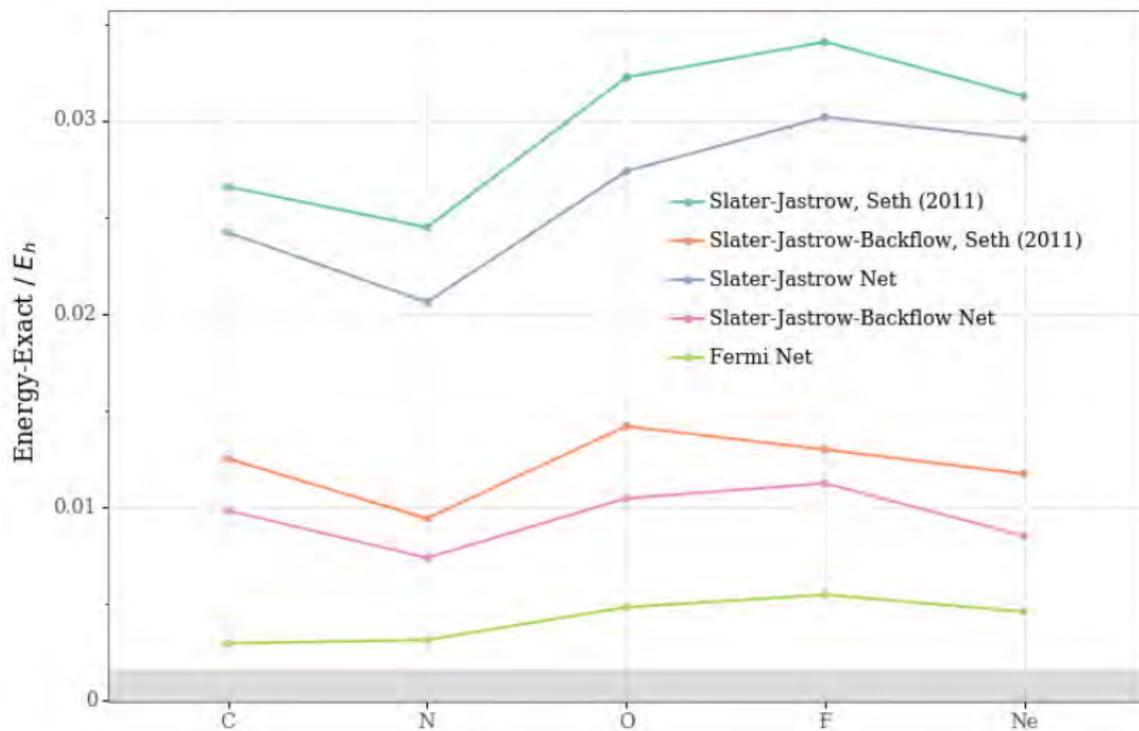
FermiNet versus Slater-Jastrow

Single Determinant

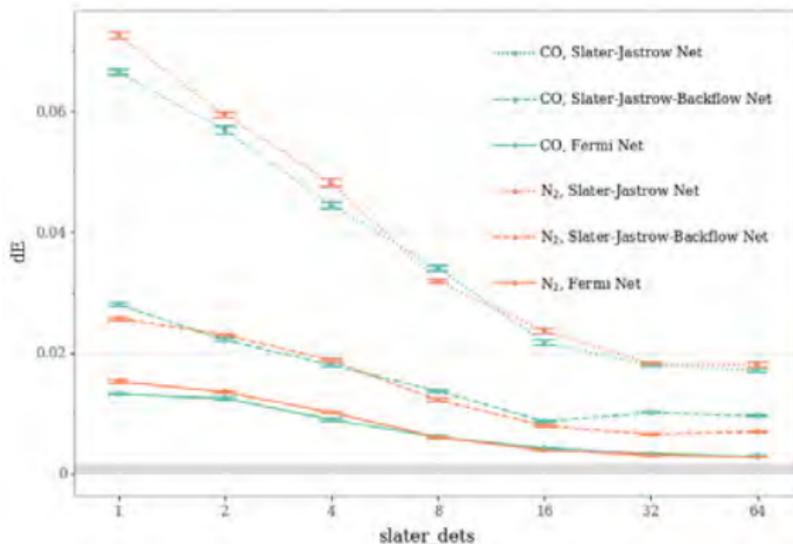


FermiNet versus Slater-Jastrow

Single Determinant



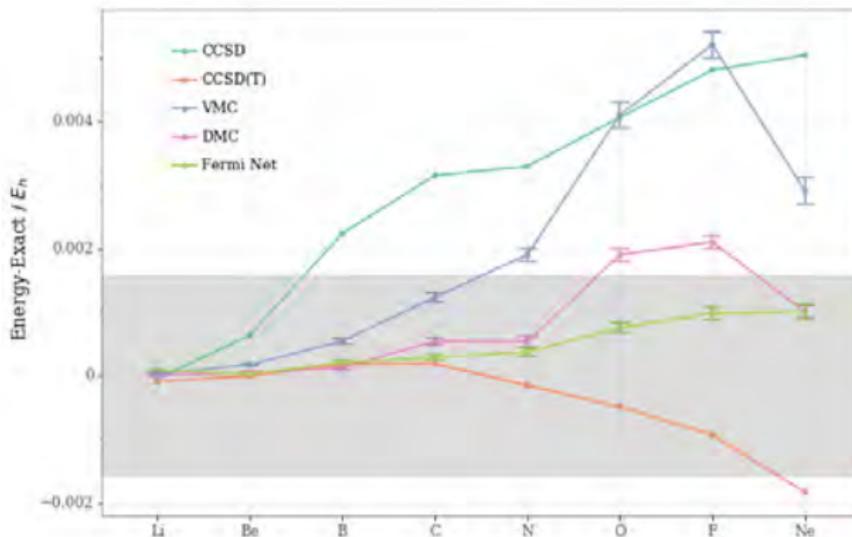
Adding Determinants



- Fermi Net converges quite rapidly with the number of determinants.
- Substantially easier to optimise than Slater-Jastrow and Slater-Jastrow-Backflow networks.

Atoms

Total Energies

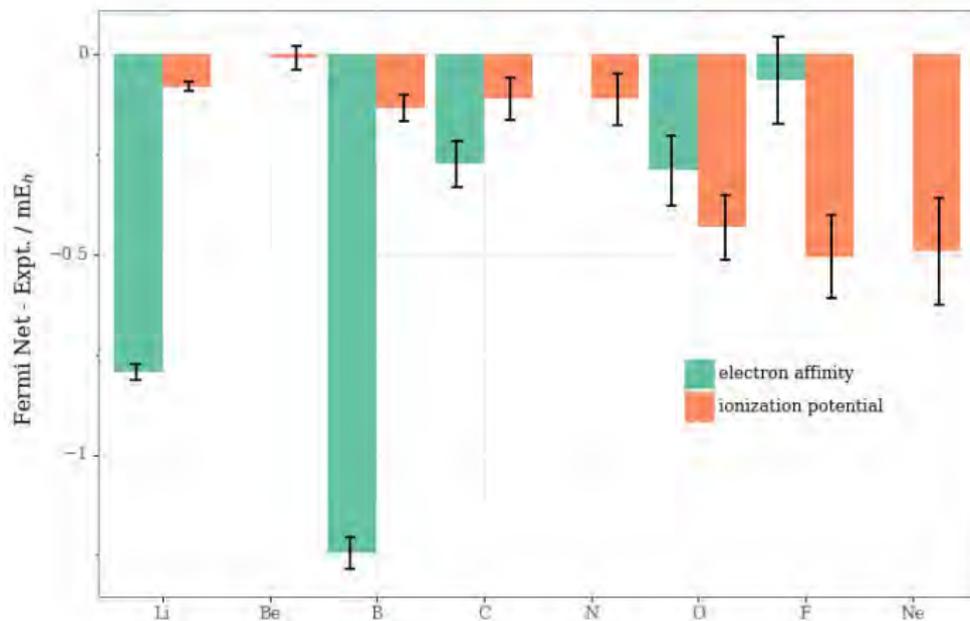


- 16 Fermi Net determinants. (SJB VMC & DMC used 50–100 SJB CSFs.)
- Fermi Net consistently captures 99.7% of correlation energy.

VMC, DMC: P. Seth, P.L. Ríos and R.J. Needs, *J. Chem. Phys.* **134**, 084105 (2011).

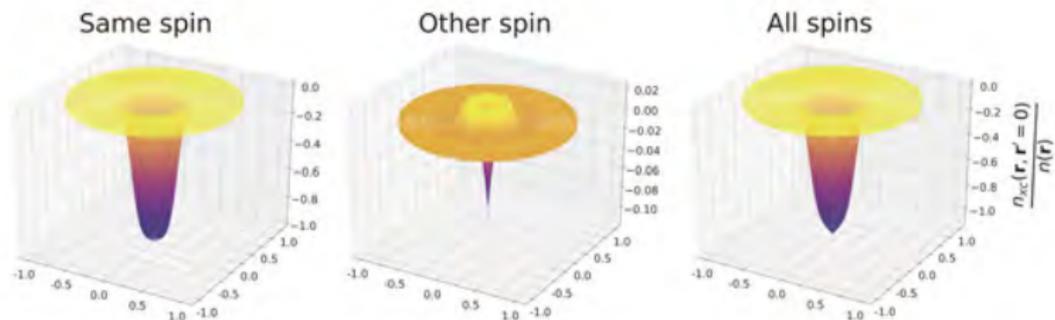
Exact: Chakravorty *et al.*, *Phys. Rev. B* **47**, 3649 (1993).

Ionization Potentials and Electron Affinities

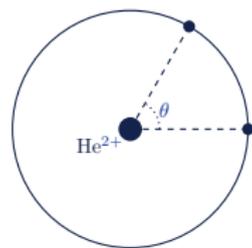
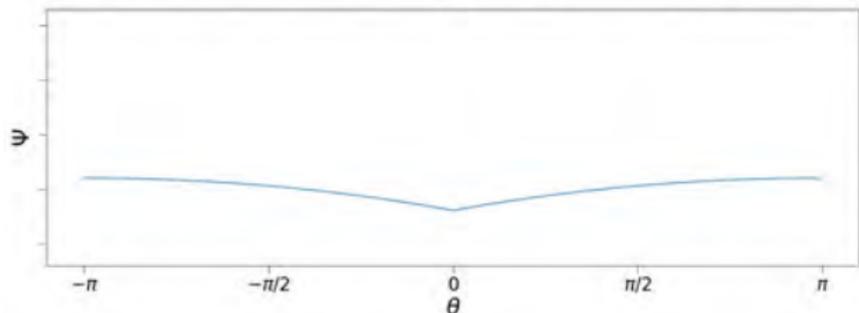
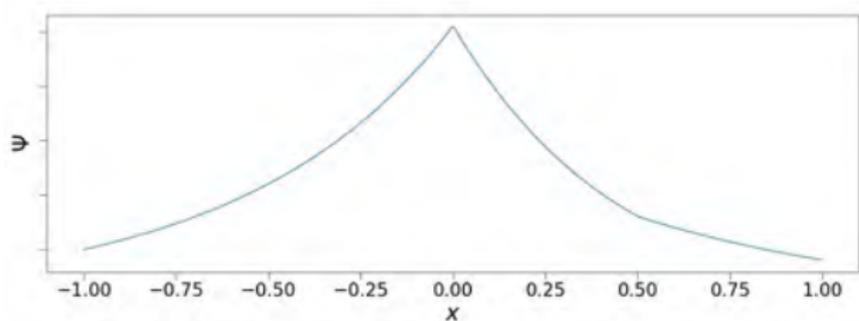


Experimental data corrected for relativistic effects from Klopper *et al.*, Phys. Rev. A 81, 022503 (2010)

Exchange-Correlation Hole



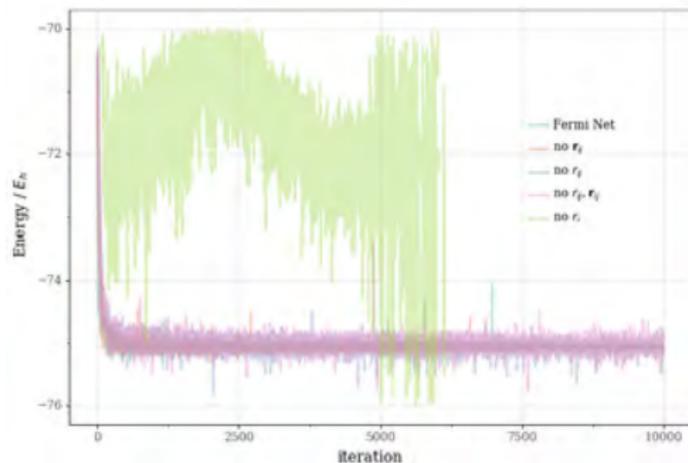
Learning the Cusps



$$\left. \frac{\partial \ln |\Psi|}{\partial r_1} \right|_{r_1=0} = -1.9979(4),$$

$$\left. \frac{\partial \ln |\Psi|}{\partial r_{12}} \right|_{r_{12}=0} = 0.4934(1).$$

Cusp Ablation

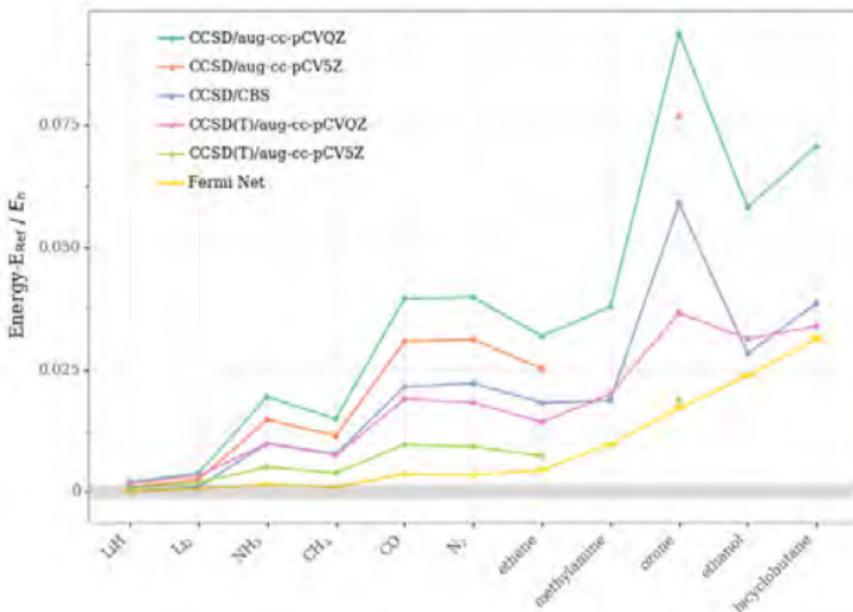


Error (mE_h) without feature inputs

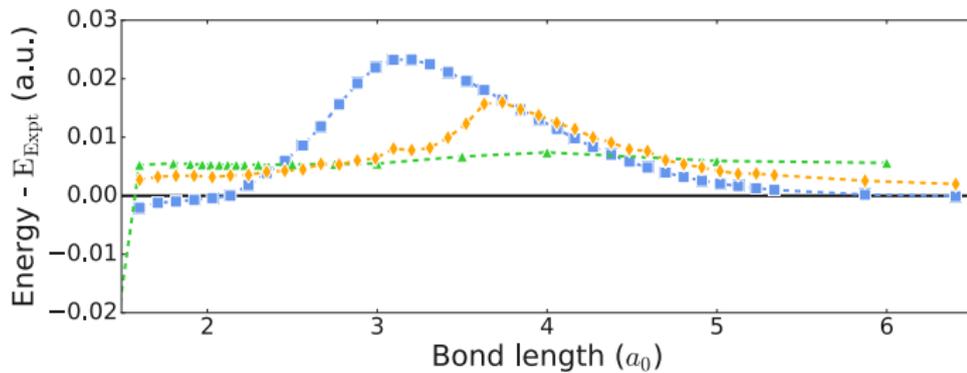
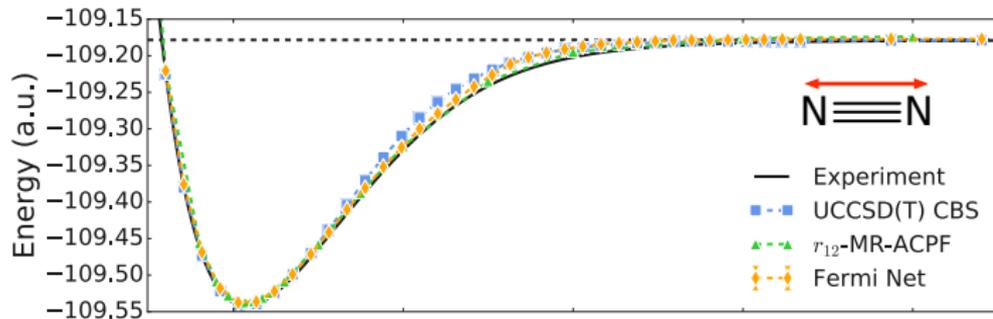
	without r_{ij}	with r_{ij}
without $ r_{ij} $	89.7	28.4
with $ r_{ij} $	1.2	0.8

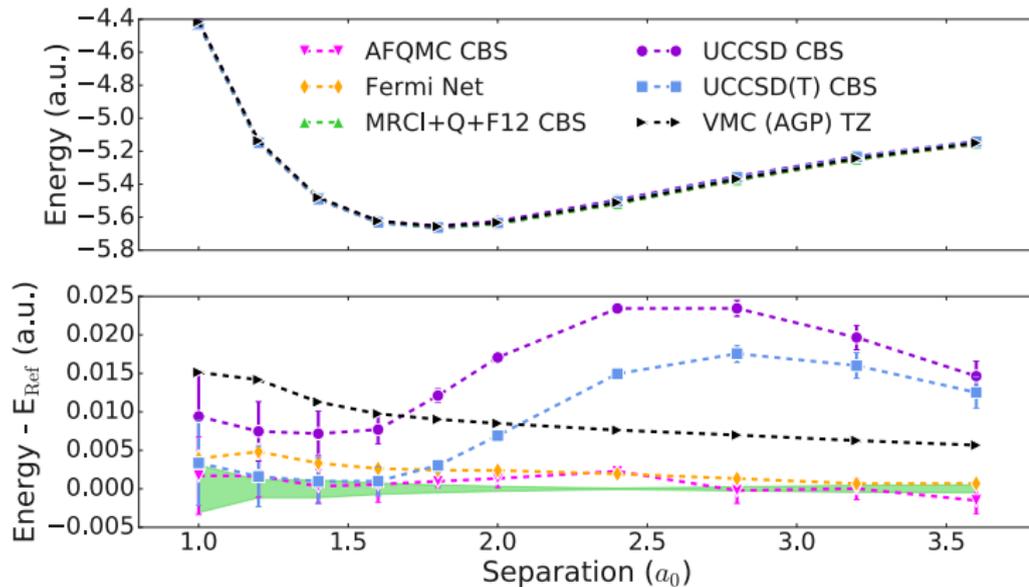
Molecules

Total Energies



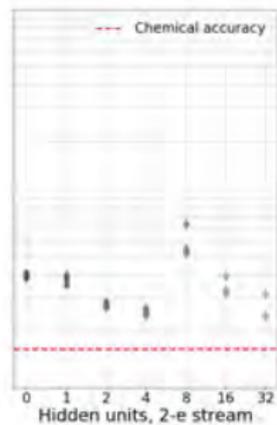
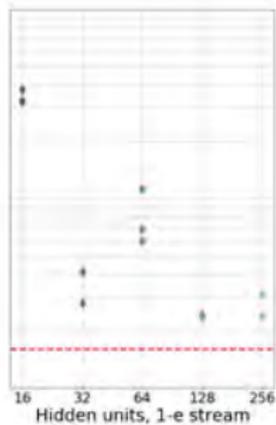
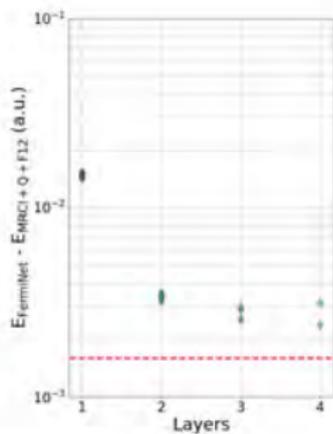
- Fermi Net outperforms CCSD(T) in QZ, 5Z basis sets.
- Accuracy degrades smoothly as number of electrons increases.

N_2 

H_{10} 

H_{10}

Network Ablation



Positives

- FermiNet is *way* better than any other form of VMC applicable to real molecules with more than a couple of atoms.
- Rivals coupled cluster for equilibrium geometries and outperforms it for molecules with a strong multi-determinant nature.
- Can serve as a trial function for DMC and other projector methods.
- We have only just begun. Coupled cluster and SJ VMC have a fifty-year start.

Questions

- How general is the FermiNet wavefunction?
- Can we really get away with only a few determinants?
- Limits to accuracy?
- Solids?
- Size extensivity? (SJ VMC is extensive)
- Locality of correlations?
- Scaling? (SJ VMC scales very well)