DeepMind

Better, Faster Fermionic Neural Networks

James S. Spencer[1], David Pfau[1], Aleksandar Botev[1], W. M. C. Foulkes[2]
[1]DeepMind, [2]Imperial College London

Ab initio neural network wavefunctions

The behaviour of a chemical system is encoded in the wavefunction, which depends upon all electron coordinates - a 3N dimensional space. The wavefunction can be found by solving the time-independent Schrödinger equation:

$$\hat{H}\Psi(\mathbf{r}_1, \mathbf{r}_2, \cdots \mathbf{r}_N) = E\Psi(\mathbf{r}_1, \mathbf{r}_2, \cdots \mathbf{r}_N)$$

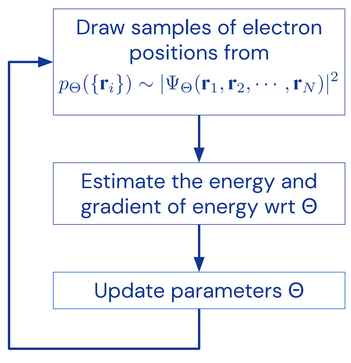
The Hamiltonian operator captures the kinetic energy and potential interactions between the electrons and atomic nuclei of the system:

$$\hat{H} = \underbrace{-\frac{1}{2}\sum_i \nabla_i^2}_{\text{kinetic}} + \overbrace{\sum_{i<j} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|}}^{\text{electron-electron}} - \underbrace{\sum_{iI} \frac{Z_I}{|\mathbf{r}_i - \mathbf{R}_I|}}_{\text{electron-nuclear}} + \overbrace{\sum_{I<J} \frac{Z_I Z_J}{|\mathbf{R}_I - \mathbf{R}_J|}}^{\text{nuclear-nuclear}}$$

Exact solutions are impossible for anything beyond the simplest systems.
An additional complication: electrons are fermions → The wavefunction must be antisymmetric with respect to exchange of pairs of electron coordinates.

Goal Directly learn a neural network that represents the wavefunction of a molecule given *only* the atomic positions and number of electrons without *any* external training data.

Solution Use the Variational Monte Carlo approach to optimise a parameterized function to give the lowest possible energy[1]:



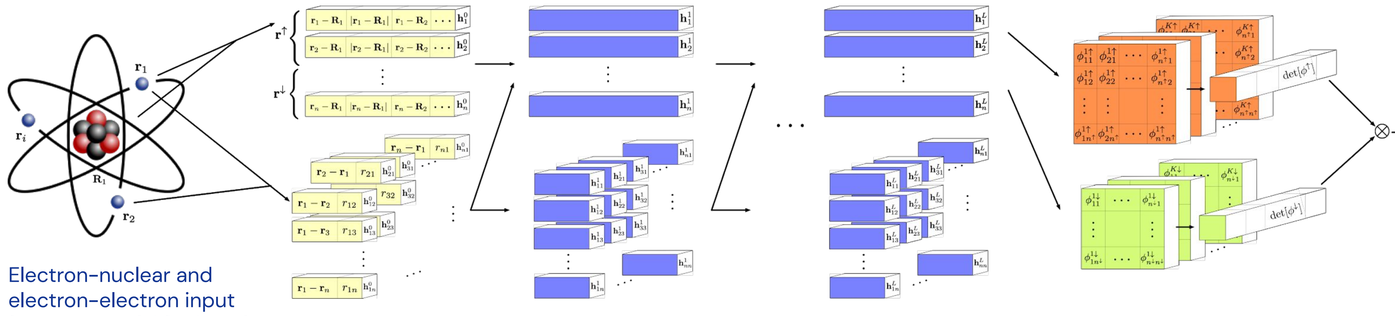
We use K-FAC[2], a scalable approximate second-order method, for optimization.

Simplifying and scaling FermiNet

1. JAX implementation of FermiNet[1]
   - Improved GPU utilization from 60% to 90%.
   - Reduced memory overhead - can run bicyclobutane (30 electrons) with a batch size of 4096 on just 4 V100 GPUs compared to 16 with TF1.
2. Efficient enforcing of boundary conditions
   - The wavefunction must go to zero at infinite distance from the nuclei.
   - Replacing an anisotropic decay function with an isotropic decay greatly improves performance, especially in gradient evaluation.

Overall: **10x reduction** in GPU hours required to train FermiNet on bicyclobutane.

FermiNet



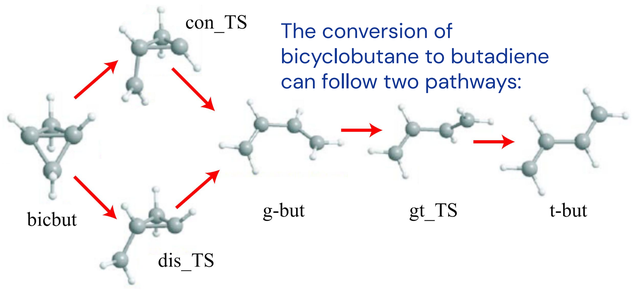
Electron-nuclear and electron-electron input features crucial for learning **cusps**

Permutation-equivariant intermediate activations mix one- and two-electron information

Determinants impose antisymmetry

Output in log domain for numerical stability

Bicyclobutane to 1,3-butadiene



The conversion of bicyclobutane to butadiene can follow two pathways:

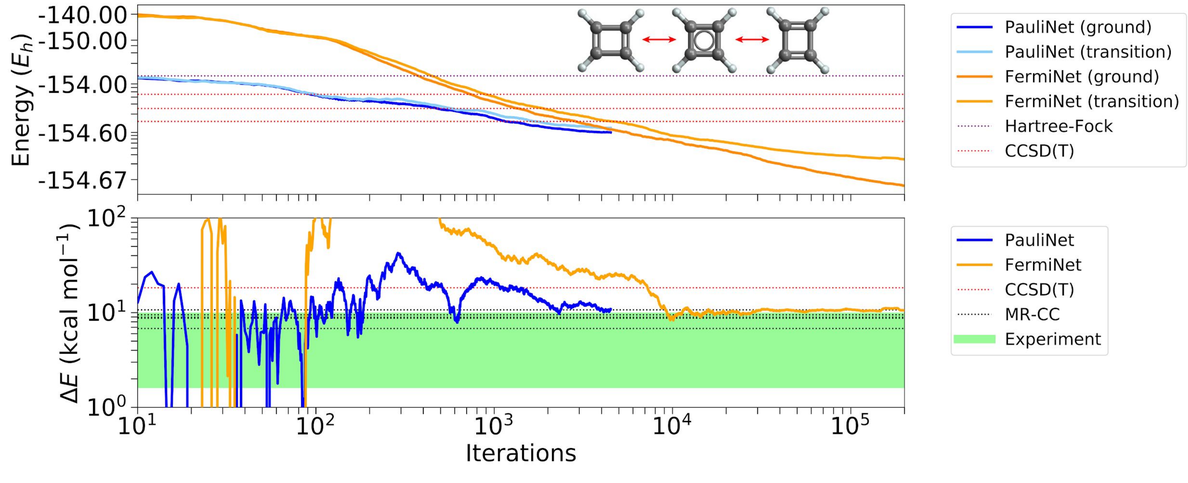
Conventional coupled cluster methods give a radically different energy for the disrotatory pathway.

FermiNet is in good agreement with other highly accurate methods and experiment

Energies in kcal/mol and relative to bicyclobutane.

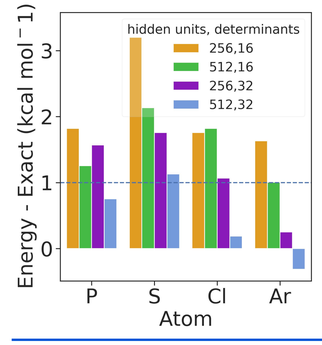
Method	con_TS	dis_TS	g-but	gt_TS	t-but
CCSD(T)	40.4	21.8	-25.1	-22.3	-28.0
CR-CC(2,3)	41.1	66.1	-24.9	-22.1	-27.9
CCSDt	40.1	59.0	-27.2	-25.3	-31.1
CC(t;3)	40.2	60.1	-25.3	-22.6	-28.3
DMC	40.4±0.5	58.6±0.5	-25.2±0.5	-22.2±0.5	-27.9±0.5
FermiNet	40.2±0.1	57.7±0.1	-25.3±0.1	-22.5±0.1	-28.4±0.1
Experiment	40.6±2.5				-25.9±0.4

Automerization of cyclobutadiene



FermiNet achieves lower energies than PauliNet[3] for the ground and transition states. The height of the transition barriers predicted by FermiNet and PauliNet are at the upper end of range of experimental values. FermiNet agrees well with the highest multireference coupled cluster (MR-CC) result.

Period 3 atoms



References

This work: Better, Faster Fermionic Neural Networks, JSS, DP, AB, WMCF. arXiv:2011.07125 (2020)

[1] Ab initio solution of the many-electron Schrödinger equation with deep neural networks, DP, JSS, Alexander G. D. G. Matthews, and WMCF. Phys. Rev. Research 2, 033429 (2020).

[2] Optimizing neural networks with Kronecker-factored approximate curvature, James Martens and Roger Grosse. ICML, 2408–2417. (2015).

[3] Deep-neural-network solution of the electronic Schrödinger equation, Jan Hermann, Zeno Schätzle, and Frank Noé. Nature Chemistry, 12, 891–897 (2020).