Solving the electronic Schrödinger equation with deep learning

Jan Hermann[1,2], Zeno Schätzle[1], and Frank Noé[1,3,4]

[1]*FU Berlin, Department of Mathematics and Computer Science, Arnimallee 6, 14195 Berlin, Germany*
[2]*TU Berlin, Machine Learning Group, Marchstr. 23, 10587 Berlin, Germany*
[3]*FU Berlin, Department of Physics, Arnimallee 14, 14195 Berlin, Germany*
[4]*Rice University, Department of Chemistry, Houston, TX 77005, USA*

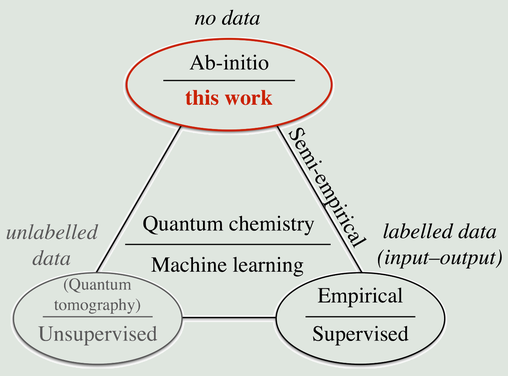
Background

- Electronic Schrödinger equation is a second-order differential eigenvalue equation which holds keys to the riches of quantum chemistry

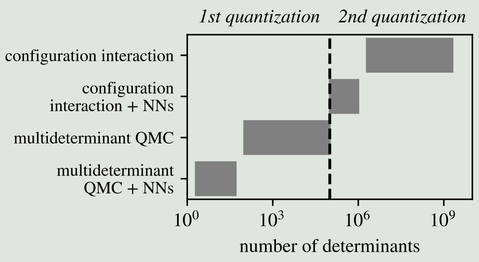
$$\hat{H}\psi(\mathbf{r}_1,\ldots,\mathbf{r}_N) = E\psi(\mathbf{r}_1,\ldots,\mathbf{r}_N)$$

$$\hat{H} := \sum_i \left(\tfrac{1}{2}\nabla^2_{\mathbf{r}_i} + \sum_I \frac{Z_I}{|\mathbf{r}_i-\mathbf{R}_I|}\right) + \sum_{i<j}\frac{1}{|\mathbf{r}_i-\mathbf{r}_j|}$$

- Machine learning for quantum chemistry has been so far mostly used to *circumvent* the Schrödinger equation, rather than to solve it



- Almost all quantum chemistry uses determinant expansions, which grow rapidly with system size in highly accurate methods, making them computationally inefficient



- Neural networks are known as universal function approximators and as efficient searchers of low-dimensional patterns in high-dimensional data—could they minimize the required number of determinants?
- Neural-network wave-function ansatzes were used successfully for discrete quantum systems[1]

Traditional wave function ansatzes

- Hartree–Fock (HF) method of quantum chemistry expresses uncorrelated electronic wave functions as a Slater determinant of molecular orbitals $\varphi_\mu(\mathbf{r}_i)$

$$\psi_{\mathrm{HF}}(\mathbf{r}) = \det[\varphi_\mu(\mathbf{r}_i^\uparrow)]\det[\varphi_\mu(\mathbf{r}_i^\downarrow)]$$

- Traditional QMC uses totally symmetric non-negative Jastrow factors to build correlations between electrons into uncorrelated wave functions, but these do not improve the *nodal surface*

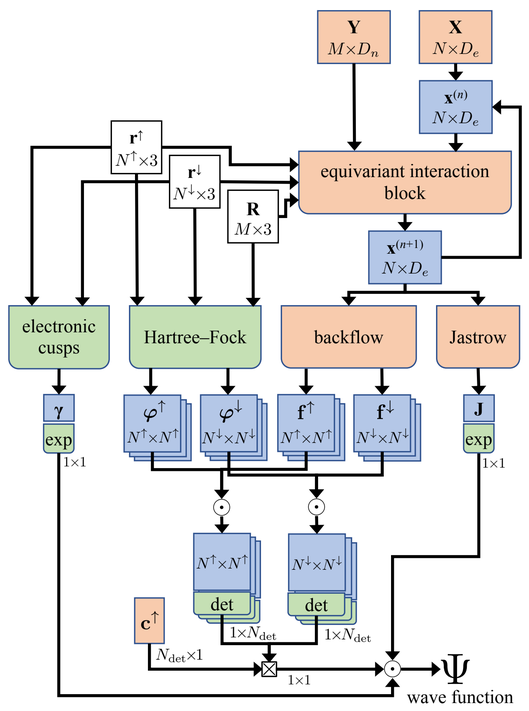
$$\psi(\mathbf{r}) := \psi_{\mathrm{HF}}(\mathbf{r})\mathrm{e}^{J(\mathbf{r})}$$

- Backflow transformation replaces one-electron coordinates with coupled coordinates of pseudoelectrons, which improves the nodal surface, but does not generalize well to larger systems

$$\mathbf{r}_i \mapsto \mathbf{r}_i + \mathbf{b}_i(\mathbf{r})$$

Summary & Outlook

- We developed PauliNet, a neural-network ansatz for electronic wave functions with multiple physical constraints built in, which is optimized with variational quantum Monte Carlo
- The flexibility of deep neural networks enables reaching much higher accuracy in the electronic energy than what is possible with traditional variational ansatzes of the same system-size scaling
- In future work we will incorporate additional established QMC techniques as well as ingredients from other neural-network ansatzes such as FermiNet[3]
- The final goal is to offer an electronic structure method able to treat complex electronic environments in systems with low hundreds of electrons



N number of electrons
$N^\uparrow$ number of spin-up electrons
$N^\downarrow$ number of spin-down electrons
M number of nuclei
D_e embedding dimension, electrons
D_n embedding dimension, nuclei

neural network (trainable function) | trainable array | fixed function | input array | hidden array

PauliNet: Deep-learning trial wave function

- Our proposed ansatz is of the multideterminant Slater–Jastrow–backflow type, where both the Jastrow factor J and the backflow $\mathbf{f}$ are deep neural networks

$$\psi_{\boldsymbol{\theta}}(\mathbf{r}) = \mathrm{e}^{\gamma(\mathbf{r})+J_{\boldsymbol{\theta}}(\mathbf{r})}\sum_p c_p \det[\tilde{\varphi}^\uparrow_{\boldsymbol{\theta},\mu_{p^i}}(\mathbf{r})]\det[\tilde{\varphi}^\downarrow_{\boldsymbol{\theta},\mu_{p^i}}(\mathbf{r})]$$

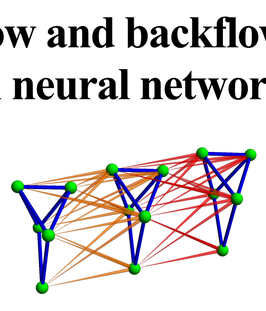
$$\tilde{\varphi}_{\boldsymbol{\theta},\mu i}(\mathbf{r}) = \varphi_\mu(\mathbf{r}_i) f_{\boldsymbol{\theta},\mu i}(\mathbf{r})$$

$\varphi_\mu(\mathbf{r}_i)$: One-electron molecular orbitals
$\gamma(\mathbf{r})$: Electron–electron cusp conditions
$J_{\boldsymbol{\theta}}(\mathbf{r})$: Jastrow-factor neural network
$\mathbf{f}_{\boldsymbol{\theta}}(\mathbf{r})$: Backflow neural network

Hartree–Fock baseline We build in multireference Hartree–Fock solution into PauliNet to ensure a good starting point for the optimization and to increase data efficiency

Asymptotic behavior Electronic eigenstates satisfy *cups conditions* when an electron approaches other electron or a nucleus, which are built into PauliNet through $\boldsymbol{\varphi}(\mathbf{r})$ and $\gamma(\mathbf{r})$

Jastrow and backflow with graph neural networks

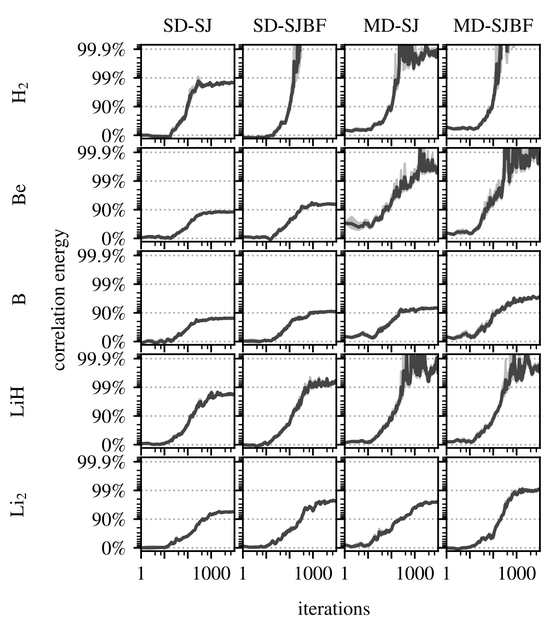


- To retain antisymmetry, the Jastrow-factor and backflow neural networks must be invariant and equivariant, respectively, with respect to the exchange of same-spin electrons

$$\varphi_\mu(\mathbf{r}_i) \to \tilde{\varphi}_{\mu i}(\mathbf{r}), \qquad \mathscr{P}\tilde{\varphi}_{\mu i}(\mathbf{r}) = \tilde{\varphi}_{\mu i}(\mathscr{P}\mathbf{r})$$

- PauliNet uses an adapted SchNet architecture,[2] which is a graph neural network designed for many-particle systems

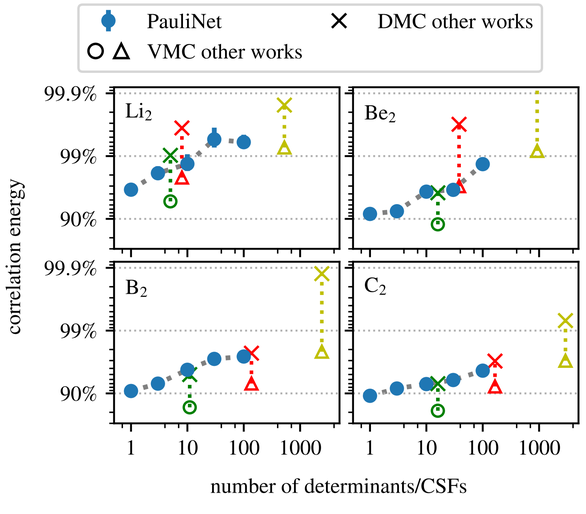
Approaching exact solution with few determinants



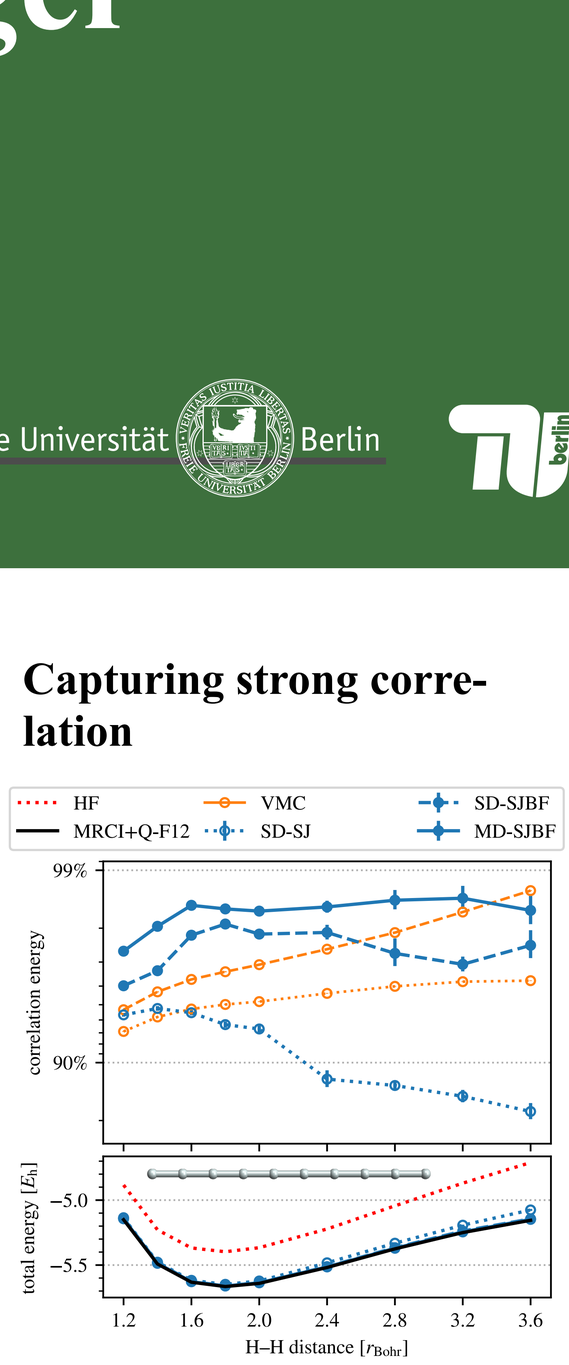
- PauliNet recovers 97% to 99.9% of the electron correlation energy after training for tens of minutes to a few hours on a single GTX 1080 Ti GPU
- Only 6 determinants are necessary to substantially reduce the correlation energy from a single-determinant ansatz

Scaling with determinants and system size

- Substantial improvement over existing VMC results and comparable to DMC results in the regime of a few number of determinants

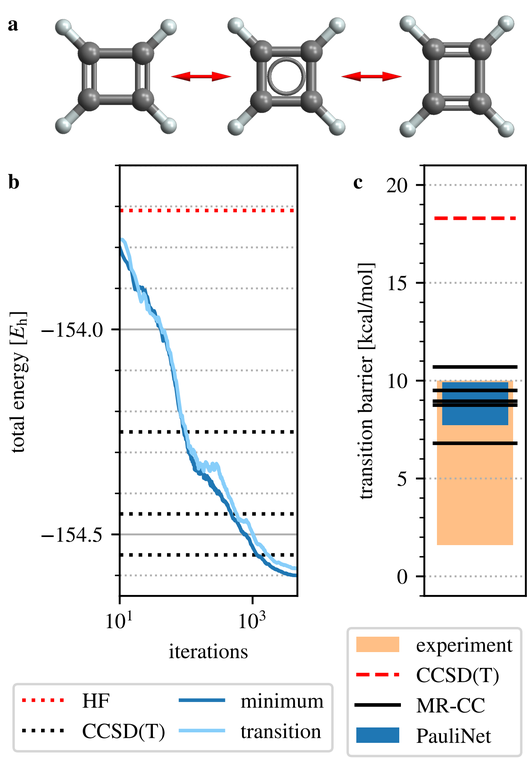


Capturing strong correlation



- Dissociation of linear H_{10} as a test case of strong correlation
- PauliNet with 16 determinants captures 98–99% of the correlation energy along the dissociation

Targeting state-of-the-art quantum chemistry



- Barrier of the automerization of cyclobutadiene (28 electrons) still not fully resolved
- Strong multireferential character: CCSD(T) overestimates two-fold
- Experiment: 1.6–10 kcal/mol, MR-CC 7–11 kcal/mol
- PauliNet with 10 determinants: 8–10 kcal/mol
- 3 days on GTX 1080 Ti GPU

Article

J. Hermann, Z. Schätzle & F. Noé. Deep-neural-network solution of the electronic Schrödinger equation. *Nat. Chem.* **12**, 891–897 (2020)

References

(1) Carleo & Troyer, *Science* **355**, 6325 (2017)
(2) Schütt et al., *J. Chem. Phys.* **148**, 241722 (2018)
(3) Pfau et al., *Phys. Rev. Research* **2**, 033429 (2020)