Freie Universität Berlin

Approaching exact solutions of the electronic Schrödinger equation with deep quantum Monte Carlo

Jan Hermann Zeno Schätzle Frank Noé

FU Berlin, Department of Mathematics and Computer Science, Berlin, Germany

APS March Meeting, 16 March 2021

Electronic Schrödinger equation

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

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

- Fundamental law of chemistry and materials science
- Electron coordinates: $(\mathbf{r}_1, \ldots, \mathbf{r}_N) \equiv \mathbf{r}$
- Hamiltonian operator $\hat{H}$ specified by charges Z_I and coordinates $\mathbf{R}_I$ of M atomic nuclei
- Solution: eigenstates ψ_n and their energies E_n, including the ground state ψ_0, E_0
- Analytically solvable only for $M = N = 1$, but many approximate numerical methods in quantum chemistry

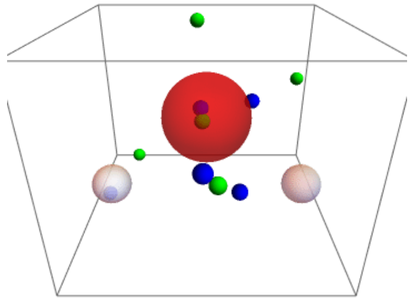
Electronic wave functions and antisymmetry

- Hamiltonian has no spin – electrons have spin coordinates, but those can be considered fixed

$$\psi(\mathbf{r}_1^{\uparrow}, \ldots, \mathbf{r}_{N_{\text{up}}}^{\uparrow}, \mathbf{r}_{N_{\text{up}}+1}^{\downarrow}, \ldots, \mathbf{r}_N^{\downarrow}), \quad s_i = \begin{cases} \uparrow & i \leq N_{\text{up}} \\ \downarrow & i > N_{\text{up}} \end{cases}$$

- Electrons are fermions – wave function must be *antisymmetric* w.r.t. exchange of same-spin electrons ($s_i = s_j$)

$$\psi(\ldots, \mathbf{r}_i, \ldots, \mathbf{r}_j, \ldots) = -\psi(\ldots, \mathbf{r}_j, \ldots, \mathbf{r}_i, \ldots)$$



- $\psi(\mathbf{r}) = 0$ defines a $(3N-1)$-dimensional *nodal surface* – *sign problem* in diffusion quantum Monte Carlo
- Expansion in Slater determinants – rapid scaling of computational cost with system size

Variational quantum Monte Carlo

$$E_0 = \min_{\psi} E[\psi] \leq \min_{\boldsymbol{\theta}} E[\psi_{\boldsymbol{\theta}}],$$

$$E[\psi] := \int \mathrm{d}\mathbf{r}\, \psi(\mathbf{r}) \hat{H} \psi(\mathbf{r})$$

- Standard quantum chemistry: $E[\psi]$ can be evaluated analytically
- Quantum Monte Carlo (QMC): quantum *expectation* value expressed as statistical *expected* value of local energy

$$E[\psi] = \mathbb{E}_{\mathbf{r}\sim|\psi|^2}\left[E_{\mathrm{loc}}[\psi](\mathbf{r})\right]$$

$$E_{\mathrm{loc}}[\psi] = \hat{H}\psi/\psi$$

- Traditional ansatz of the Slater–Jastrow–backflow type[1]

$$\psi_{\mathrm{SJBF}}(\mathbf{r}) := \mathrm{e}^{J(\{\mathbf{r}_i^{\uparrow}\},\{\mathbf{r}_i^{\downarrow}\})} \sum_p c_p \det[\varphi_{\mu_p}(\mathbf{r}_i^{\uparrow})] \det[\varphi_{\mu_p}(\mathbf{r}_i^{\downarrow})]$$

$$\mathbf{r}_i \to \mathbf{r}_i + \mathbf{q}_i(\mathbf{r})$$

- Accuracy of variational QMC limited only by the parametric flexibility of the ansatz – use neural networks[2]

[1]Brown et al., *J. Chem. Phys.* **126**, 224110 (2007) [2]Carleo & Troyer, *Science* **355**, 602 (2017)

PauliNet: Deep-learning trial wave function

$$\psi_{\boldsymbol{\theta}}(\mathbf{r}) = \mathrm{e}^{\gamma(\mathbf{r})+J_{\boldsymbol{\theta}}(\mathbf{r})} \sum_p c_p \det[\tilde{\varphi}^{\uparrow}_{\boldsymbol{\theta},\mu_p i}(\mathbf{r})] \det[\tilde{\varphi}^{\downarrow}_{\boldsymbol{\theta},\mu_p i}(\mathbf{r})], \qquad \tilde{\varphi}_{\boldsymbol{\theta},\mu i}(\mathbf{r}) = \varphi_{\mu}(\mathbf{r}_i) f_{\boldsymbol{\theta},\mu i}(\mathbf{r})$$

- Slater–Jastrow–backflow ansatz with Jastrow and backflow represented by neural networks
- Generalized backflow
- Built-in n–e and e–e cusp conditions
- Permutation-equivariant cusp-less graph neural networks to retain antisymmetry and cusp conditions

- $\varphi_{\mu}(\mathbf{r}_i)$: HF orbitals
- $J_{\boldsymbol{\theta}}(\mathbf{r})$: Jastrow-factor NN
- $\mathbf{f}_{\boldsymbol{\theta}}(\mathbf{r})$: backflow NN
- $\gamma(\mathbf{r})$: e–e cusp conditions

Hermann, Schätzle & Noé, *Nat. Chem.* **12**, 891 (2020)

PauliNet: Deep-learning trial wave function

Neural network

- Adapted SchNet architecture[1] as the underlying graph-neural network

$$\mathbf{x}_i^{(0)} := \mathbf{X}_{\boldsymbol{\theta},s_i}$$

$$\mathbf{z}_i^{(n,\pm)} := \sum\nolimits_{j\neq i}^{\pm} \mathbf{w}_{\boldsymbol{\theta}}^{(n,\pm)}\big(\mathbf{e}(|\mathbf{r}_i-\mathbf{r}_j|)\big)\odot\mathbf{h}_{\boldsymbol{\theta}}^{(n)}\big(\mathbf{x}_j^{(n)}\big)$$

$$\mathbf{z}_i^{(n,\mathrm{n})} := \sum\nolimits_{J} \mathbf{w}_{\boldsymbol{\theta}}^{(n,\mathrm{n})}\big(\mathbf{e}(|\mathbf{r}_i-\mathbf{R}_J|)\big)\odot\mathbf{Y}_{\boldsymbol{\theta},J}$$

$$\mathbf{x}_i^{(n+1)} := \mathbf{x}_i^{(n)} + \sum\nolimits_{\pm} \mathbf{g}_{\boldsymbol{\theta}}^{(n,\pm)}\big(\mathbf{z}_i^{(n,\pm)}\big)+\mathbf{g}_{\boldsymbol{\theta}}^{(n,\mathrm{n})}\big(\mathbf{z}_i^{(n,\mathrm{n})}\big)$$

$$J := \eta_{\boldsymbol{\theta}}\Big(\sum\nolimits_i \mathbf{x}_i^{(N)}\Big)$$

$$\mathbf{f}_i := \boldsymbol{\kappa}_{\boldsymbol{\theta}}\big(\mathbf{x}_i^{(N)}\big)$$

Algorithm

- Energy expectation value directly as a loss function
- Hermiticity of Hamiltonian enables efficient evaluation of the loss gradient

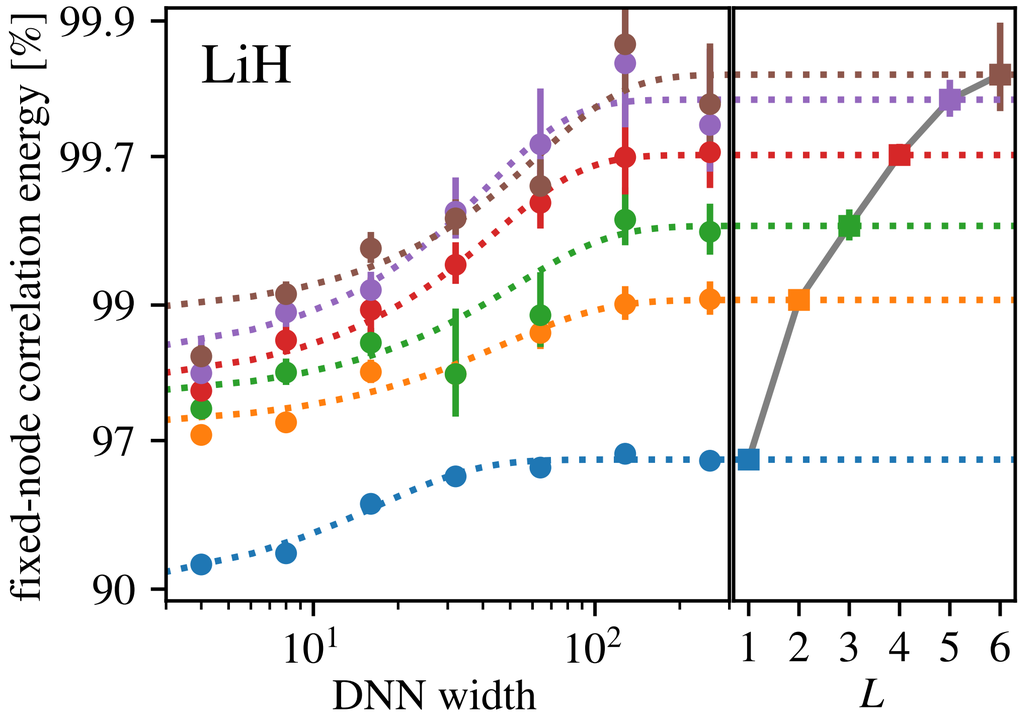
$$\mathscr{L}(\boldsymbol{\theta}) = \mathbb{E}_{\mathbf{r}\sim|\psi'^2|}\big[E_{\mathrm{loc}}[\psi_{\boldsymbol{\theta}}](\mathbf{r})\big] \quad \big(\equiv E[\psi_{\boldsymbol{\theta}}]\big)$$

$$\nabla_{\boldsymbol{\theta}}\mathscr{L}(\boldsymbol{\theta}) = 2\mathbb{E}_{\mathbf{r}\sim|\psi'^2|}\big[\big(E_{\mathrm{loc}}[\psi_{\boldsymbol{\theta}}](\mathbf{r})-E[\psi_{\boldsymbol{\theta}}]\big)\nabla_{\boldsymbol{\theta}}\ln|\psi_{\boldsymbol{\theta}}|\big]$$

- Alternating sampling (Markov-chain Monte Carlo) and optimization (Adam)

[1] Schütt et al., *J. Chem. Phys.* **148**, 241722 (2018)

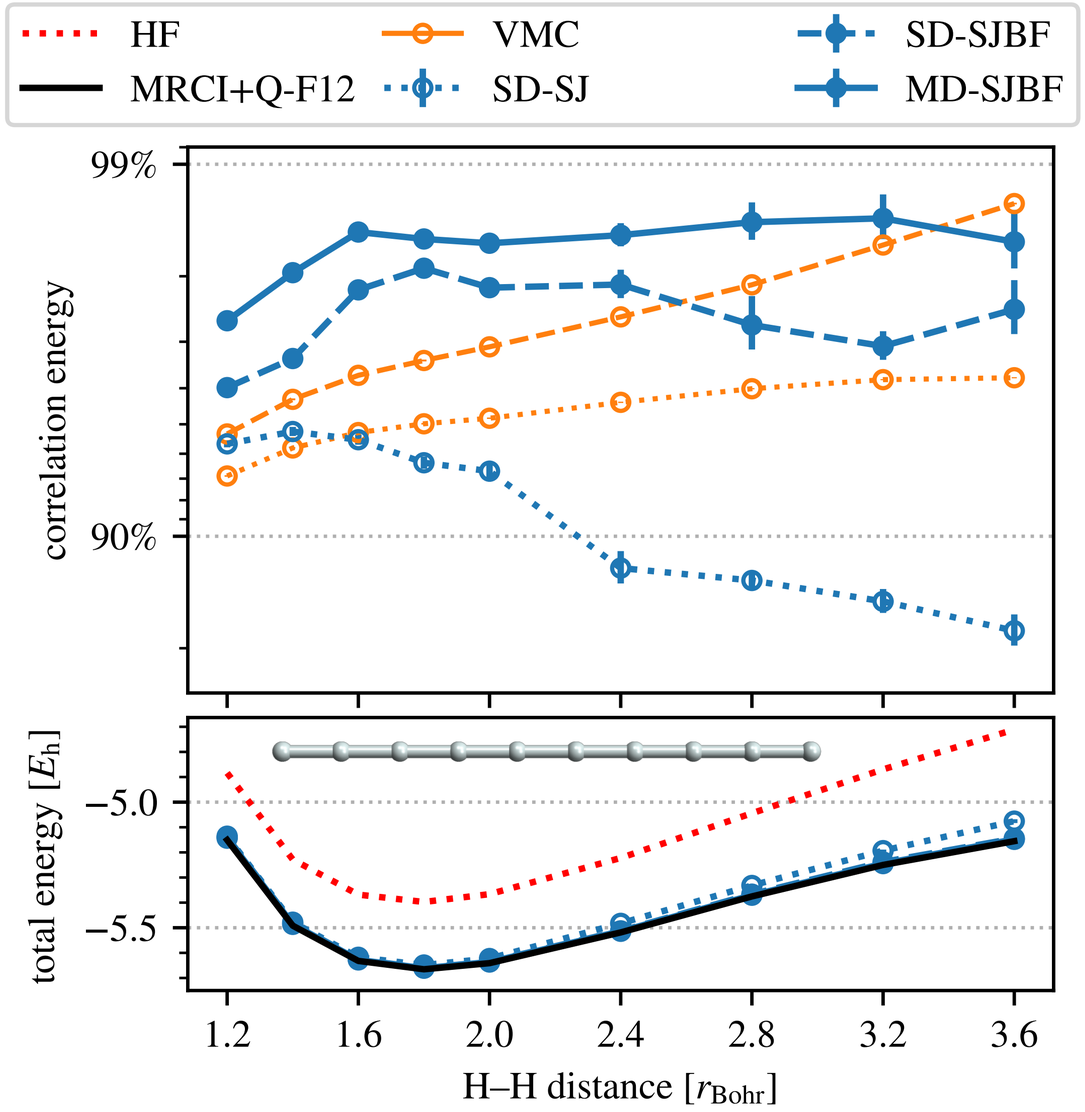
Approaching the exact solution



- Diffusion quantum Monte Carlo provides the exact variational energy for a given nodal surface
- The PauliNet Jastrow factor can be systematically converged to reach the exact fixed-node energy
- Beyond fixed nodes with backflow – how is sign problem manifested in deep variational QMC?

Schätzle, Hermann & Noé, *J. Phys. Chem.* (in print)

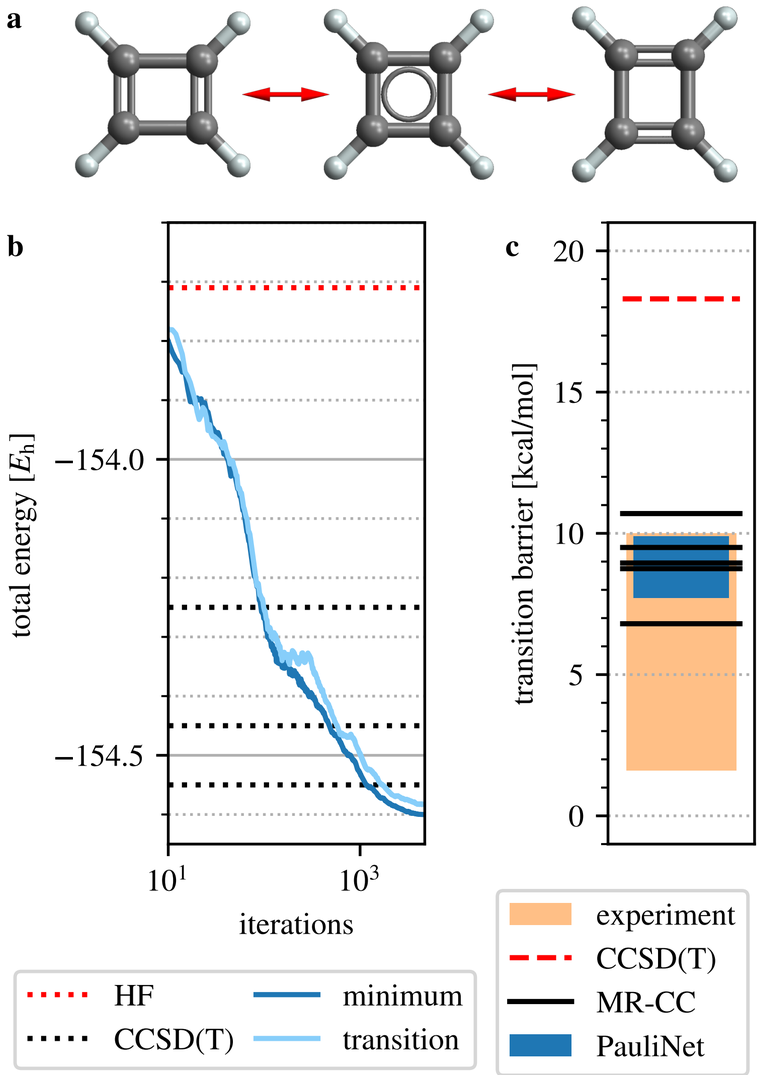
Capturing strong correlation



- Dissociation of linear H_{10} as a test case of *strong correlation*[1] – baseline HF gives unphysical dissociation energy
- PauliNet with 16 determinants captures 98–99% of the correlation energy along the dissociation
- PauliNet surpasses a traditional VMC approach specifically adapted for H_{10}

[1]Motta et al., *Phys. Rev. X* **7**, 031059 (2017)

Targeting state-of-the-art quantum chemistry



- Barrier of the automerization of cyclobutadiene (28 electrons) still not fully resolved[1]
- Strong multireferential character—CCSD(T) overestimates two-fold
- Experiment: 1.6–10 kcal/mol, multireference coupled clusters 7–11 kcal/mol
- PauliNet with 10 determinants: 8–10 kcal/mol
- 3 days on GTX 1080 Ti GPU, comparable to multireference methods, but better scaling

[1]Lyakh et al., *Chem. Rev.* **112**, 182 (2012)

Summary & outlook

- Deep QMC opens a new path to highly accurate electronic structure methods with favorable scaling with system size
- Reuse all standard techniques: diffusion QMC, pseudopotentials
- Improvements in neural network architecture
- Will the performance gain translate straightforwardly to solids?
- How quickly do we lose accuracy as we go to larger systems?

- Code at github.com/deepqmc/deepqmc

```python
from deepqmc import Molecule, train
from deepqmc.wf import PauliNet

mol = Molecule.from_name('LiH')
net = PauliNet.from_hf(mol).cuda()
train(net)
```