

#### **Electronic structure methods meet machine learning**

Jan Hermann

FU Berlin, Department of Mathematics, Berlin, Germany https://jan.hermann.name

Theoretical Chemistry seminar, UCT & IOCB, 3 June 2022

# **Computational chemistry: Accuracy vs. efficiency**

• Electronic Schrödinger equation – exact in principle, impossible to solve exactly in practice



#### **Electronic Schrödinger equation**

$$\hat{H}\psi(\mathbf{r}_1, \dots, \mathbf{r}_N) = E\psi(\mathbf{r}_1, \dots, \mathbf{r}_N)$$
$$\hat{H} := \sum_{i=1}^N \left( -\frac{1}{2} \nabla_{\mathbf{r}_i}^2 - \sum_I \frac{Z_I}{|\mathbf{r}_i - \mathbf{R}_I|} \right) + \sum_{i < j} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|}$$

- Electron coordinates:  $(\mathbf{r}_1, ..., \mathbf{r}_N) \equiv \mathbf{r}$
- Molecule specified by atom charges  $Z_I$  and coordinates  $\mathbf{R}_I$
- Solution: eigenstate wave functions  $\psi_n$  and energies  $E_n$ , including the ground state  $\psi_0, E_0$
- Variational principle:  $E_0 = \min_{\psi} \int d\mathbf{r} \, \psi(\mathbf{r}) \hat{H} \psi(\mathbf{r})$

### **Electronic structure problem**



- Ab-initio: exact **1**, approximate **2** (*coupled clusters, quantum Monte Carlo*)
- Semi-empirical: approximate **1**, exact **2** (*density functional theory, tight-binding, Hubbard model*)

# **"First-principles" quantum chemistry frameworks**

	Variable	Unknown	Approximation	Universal	(Formal) cost
QMC	$\Psi(\mathfrak{r})$	$\Psi(\mathfrak{r})$	_	×	$\sim 10^3 \times O(N^{3+})$
CI/CC	$\Psi_n$	$\Psi_{\mathbf{n}}$	finite basis	×	O(N <sup>7</sup> )
KS-DFT	$\phi_i(\mathbf{r})$	$E_{\rm xc}[n]$	mean field		O(N <sup>3-4</sup> )
OF-DFT	$n(\mathbf{r})$	$T_{\text{Pauli}}[n] + E_{\text{xc}}[n]$	bosonic		O(N)
eFF	$\{\mathbf{r}_i\}$	$E(\{\mathbf{r}_i\})$	discrete	×	$\sim 10^{-3} \times O(N)$

## State of the art

	Tagline	Number of atoms	Accuracy
QMC/CC	Benchmark methods	20 🔽 80 🙏	predictive
KS-DFT	Routine computational chemistry	300 🔽 3000 🙏	explanatory, sometimes predictive
OF-DFT	For enthusiasts	104 🔽 106 🔔	rarely explanatory, mostly unusable
eFF	"What is this?"	10 <sup>5</sup> 🔽 107 🔔	force-field-like

## **Electronic structure problem**



- Ab-initio: exact **1**, approximate **2** (*coupled clusters, quantum Monte Carlo*)
- Semi-empirical: approximate **1**, exact **2** (*density functional theory, tight-binding, Hubbard model*)
- Machine learning (force fields, density functionals, wave function ansatzes)

# Neural networks in electronic structure methods

#### • All very recent (2020–2022)

QMC	<ul> <li>JH, Schätzle &amp; Noé, Nat. Chem. 12, 891</li> <li>(PauliNet) • Pfau et al., Phys. Rev. Research 2, 033429 (FermiNet)</li> </ul>	Outperforms traditional ansatzes in accuracy (~99% of correlation energy), but significantly slower, only molecules, up to ~40 electrons
CI/CC	• Choo et al., <i>Nat</i> . <i>Commun</i> . <b>11</b> , 2368	Essentially second-quantized QMC, essentially FCI accuracy, up to ~20 electrons
KS-DFT	• Kirkpatrick et al., <i>Science</i> <b>374</b> , 6573 (DM21)	Range-separated meta-GGA hybrid with state-of- the-art accuracy, homolytic dissociation
OF-DFT	<ul> <li>Ghasemi &amp; Khüne, J. Chem. Phys. 154, 074107</li> <li>Imoto et al., Phys. Rev. Research 3, 033198</li> </ul>	Either toy systems (1D, atoms) or just marginally better performance
eFF	• Cools-Ceuppens et al., <i>J.</i> <i>Chem. Theory Comput.</i> <b>18</b> , 1672	First attempt at small organic molecules

#### Variational quantum Monte Carlo

$$E_0 = \min_{\psi} E[\psi] \le \min_{\theta} E[\psi_{\theta}], \qquad E[\psi] := \int d\mathbf{r} \, \psi(\mathbf{r}) \hat{H} \psi(\mathbf{r})$$

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

$$E[\psi] = \mathbb{E}_{\mathbf{r} \sim |\psi|^2} \left[ E_{\text{loc}}[\psi](\mathbf{r}) \right], \qquad E_{\text{loc}}[\psi] = \hat{H}\psi/\psi$$

Traditional ansatz of the Slater–Jastrow–backflow type

 $\psi_{\text{SJBF}}(\mathbf{r}) := \det[\varphi_{\mu}(\mathbf{r}_{i}^{\uparrow})] \det[\varphi_{\mu}(\mathbf{r}_{i}^{\downarrow})] e^{J(\{\mathbf{r}_{i}^{\uparrow}\},\{\mathbf{r}_{i}^{\downarrow}\})}, \quad \mathbf{r}_{i} \to \mathbf{r}_{i} + \mathbf{q}_{i}(\mathbf{r})$ 

• Accuracy of variational QMC limited only by the parametric flexibility of the ansatz

## Intermezzo: Electron spin

• Electrons have both spatial and spin coordinates,  $\mathbf{r}_i \in \mathbb{R}^3, s_i \in \{\uparrow, \downarrow\}$ 

$$\psi(\mathbf{r}_1, s_1, \mathbf{r}_2, s_2, \ldots)$$

• Hamiltonian has no spin, so one consider the spins fixed

$$\psi(\mathbf{r}_{1}^{\uparrow},...,\mathbf{r}_{N_{up}}^{\uparrow},\mathbf{r}_{N_{up}+1}^{\downarrow},...,\mathbf{r}_{N}^{\downarrow}), \quad s_{i} = \begin{cases} \uparrow & i \leq N_{up} \\ \downarrow & i > N_{up} \end{cases}$$

- Antisymmetry required for same-spin electrons
- Opposite-spin exchange symmetry encodes spin state (singlet, triplet,...)
- Slater determinants det  $\varphi_k(\mathbf{r}_i) = \det \varphi_k^{\uparrow}(\mathbf{r}_i^{\uparrow}) \det \varphi_k^{\downarrow}(\mathbf{r}_i^{\downarrow})$

Pauncz, Spin eigenfunctions (Plenum Press, 1979)

# Deep neural networks for the electronic Schrödinger equation



### **Real-space neural-network electronic wave functions**

1. Slater determinant

 $\psi_{\boldsymbol{\theta}}(\mathbf{r}) := \det \varphi_{k,\boldsymbol{\theta}}(\mathbf{r}_i), \qquad k = 1, \dots, N$ 

#### 2. Generalized Slater determinant

$$\psi_{\theta}(\mathbf{r}) := \det f_{ki,\theta}(\mathbf{r}), \qquad \mathbf{f}_{\theta}(\mathscr{P}_{ij}\mathbf{r}) = \mathscr{P}_{ij}\mathbf{f}_{\theta}(\mathbf{r})$$

#### 3. Real-space baseline/envelope

$$\psi_{\boldsymbol{\theta}}(\mathbf{r}) := \det f_{ki,\boldsymbol{\theta}}(\mathbf{r}) \boldsymbol{\phi}_{k}(\mathbf{r}_{i})$$



#### 4. Multiple Slater determinants

$$\psi_{\theta}(\mathbf{r}) := \sum_{p} c_{p} \det f_{ki,\theta}^{(p)}(\mathbf{r}) \phi_{k}^{(p)}(\mathbf{r}_{i})$$

Pfau et al., *Phys. Rev. Research* **2**, 033429 (2020) **JH**, Schätzle & Noé, *Nat. Chem.* **12**, 891 (2020)

#### **Intermezzo: Determinant expansion curse**



Carleo & Troyer, *Science* **355**, 602 (2017) Luo & Clark, *Phys. Rev. Lett.* **122**, 226401 (2019) Choo et al., *Nat. Commun.* **11**, 2368 (2020)

## **PauliNet: Physics-inspired neural-network ansatz**

$$\psi_{\boldsymbol{\theta}}(\mathbf{r}) := \sum_{p} \det f_{ki,\boldsymbol{\theta}}^{(p)}(\mathbf{r}) \phi_{k}^{(p)}(\mathbf{r}_{i}) := e^{J_{\boldsymbol{\theta}}(\mathbf{r}) + \gamma(\mathbf{r})} \sum_{p} c_{p} \det f_{\kappa_{pk}i,\boldsymbol{\theta}}^{\prime}(\mathbf{r}) \varphi_{\kappa_{pk}}(\mathbf{r}_{i})$$

Molecular orbitals 🛛 🐨 😽 🌄 🔽

 $\phi_k(\mathbf{r}) := \varphi_k(\mathbf{r})$ 

Neural-network Jastrow and generalized backflow

\*\*\*

 $f_{ki,\boldsymbol{\theta}}(\mathbf{r}) := \mathrm{e}^{J_{\boldsymbol{\theta}}(\mathbf{r}) + \gamma(\mathbf{r})} f'_{ki,\boldsymbol{\theta}}(\mathbf{r})$ 

**Electron configurations** 

$$J^{(p)} := J, \quad f'_{ki}{}^{(p)} := f'_{\kappa_{pk}i}, \quad \phi^{(p)}_{k} := \phi_{\kappa_{pk}}$$

**Cusp** corrections



n–e in  $\varphi_k(\mathbf{r})$ , in via  $\gamma(\mathbf{r})$ 



JH, Schätzle & Noé, *Nat. Chem.* **12**, 891 (2020) Ceperley, *J. Stat. Phys.* **63**, 1237 (1991)

#### **Capturing strong correlation**



- Dissociation of linear H<sub>10</sub> as a test case of strong correlation<sup>1</sup>
- PauliNet with 16 determinants captures 98–99% of the correlation energy along the dissociation

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

### Targeting realistic 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

Lyakh et al., *Chem. Rev.* **112**, 182 (2012) Monino et al., arXiv:2204.05098 (2022)

#### **Excited states with deep variational QMC**

$$\min_{\boldsymbol{\theta}} \left[ \sum_{i} E[\psi_{\boldsymbol{\theta},i}] + \alpha \sum_{j>i} \frac{1}{1 - |\langle \psi_{\boldsymbol{\theta},i} | \psi_{\boldsymbol{\theta},j} \rangle|} \right]$$

• Extra term in the loss function ensures orthogonality between different eigenstates



Pathak et al., *J. Chem. Phys.* **154**, 034101 (2021) Entwistle, Schätzle, Erdman, **JH** & Noé, arXiv:2203.09472 (2022)



# t for excited states

• Accurate excitation energies and intensities of low-lying states in molecules as large as benzene



# **Summary & outlook**

- Machine learning can not only augment electronic-structure methods, but also integrate into them
- Deep variational QMC is a powerful framework to solve the electronic Schrödinger equation with high accuracy
- Active field: periodic systems, transferable wave functions, excited states, pseudopotentials, diffusion QMC, better architectures
- Expensive scaling up? Benchmarks? Failure modes? Size consistency? Optimization vs. expressivity?
- Aspiration: Accurate black-box method for systems with 100+ electrons and complex electronic structure