Data Analytics and Machine Learning Group
Department of Computer Science
Technical University of Munich

On Representing Electronic Wave Functions with Sign Equivariant Neural Networks

Nicholas Gao, Stephan Günnemann

TL;DR

- We investigate sign equivariant neural networks in quantum chemistry to parametrize electronic wave functions based on learnable non-linear combinations of simpler functions.
- We identify the theoretical and empirical limitations of such an approach.

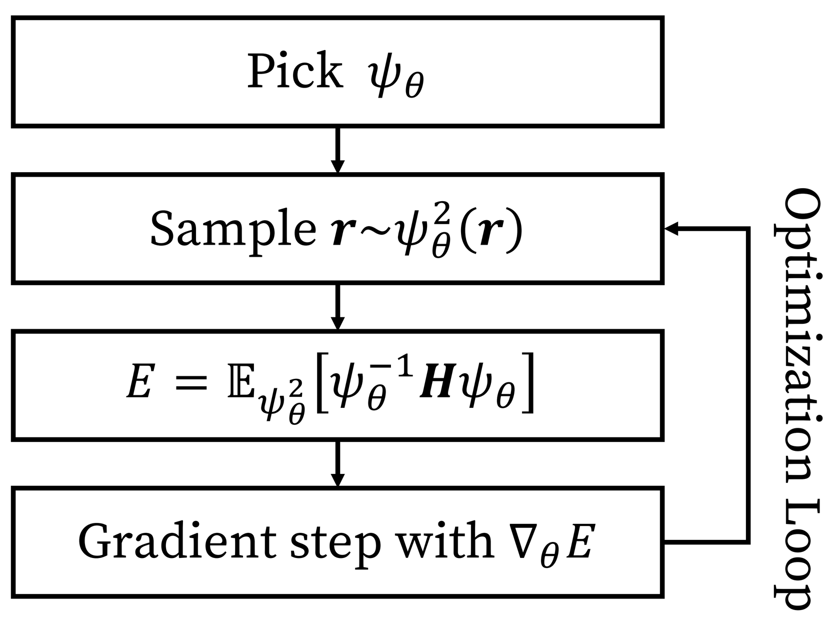
Stationary Schrödinger equation

$$H\psi_\theta(r_1,\dots,r_N) = E\psi_\theta(r_1,\dots,r_N)$$

$$H = \Delta + V(r_1,\dots,r_N), \Delta = \sum_{i=1}^{3N} \frac{\partial^2}{\partial r_i^2}$$

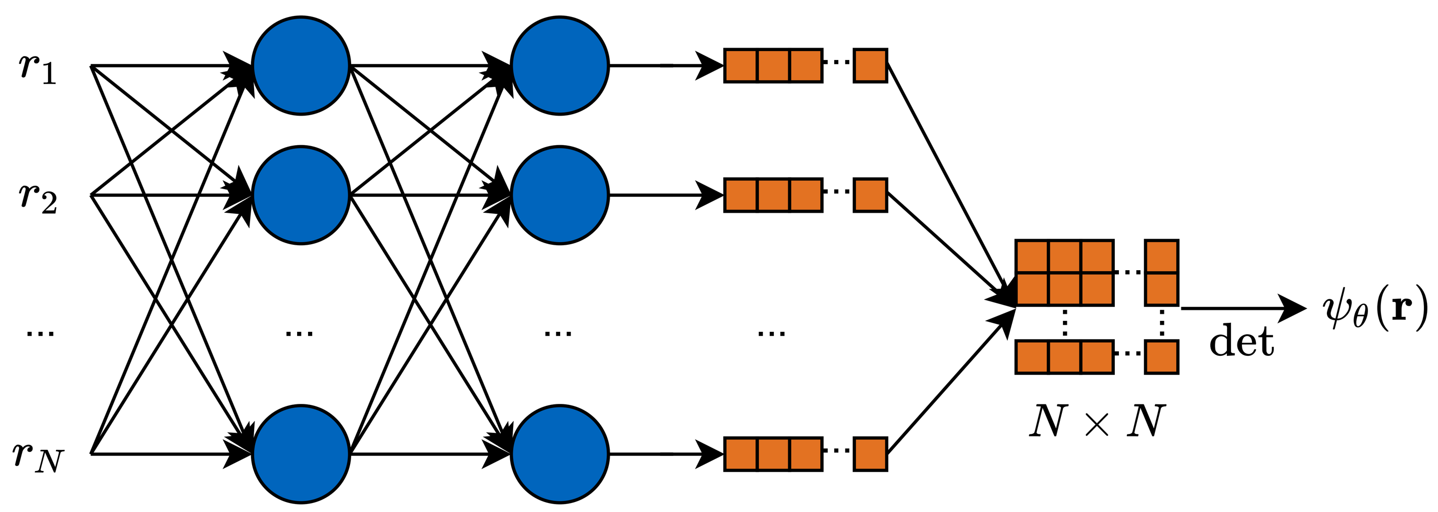
Hamiltonian	H	Wave function	$\psi_\theta: \mathbb{R}^{N\times 3} \to \mathbb{R}$
Electrons	$r \in \mathbb{R}^{N\times 3}$	Energy	$E \in \mathbb{R}$

Variational Monte Carlo



- Variational
  $\mathbb{E}_{\psi_\theta^2}[\psi_\theta^{-1}\boldsymbol{H}\psi_\theta] < E_0$
- Constraints
  - Antisymmetric: $\psi(\pi(\boldsymbol{r})) = \operatorname{sgn}(\pi)\,\psi(\boldsymbol{r})$
  - Finite integral: $\int \psi_\theta^2(\boldsymbol{r})d\boldsymbol{r} < \infty$

Neural Wave Functions



- Define ψ_θ as linear combination of determinants

$$\psi_\theta(r_1,\dots,r_N) = \sum_{k=1}^{K} \det\begin{bmatrix} \phi_1^k(r_1) & \dots & \phi_N^k(r_1) \\ \vdots & \ddots & \vdots \\ \phi_1^k(r_N) & \dots & \phi_N^k(r_N) \end{bmatrix} = \sum_{k=1}^{K} \psi_\theta^k(\dots)$$

- Parameterize ϕ_i via neural networks

⇒ Very accurate but expensive

Computational Complexity

Operation	Complexity (FermiNet)	Complexity (ours)
$\phi_i(r_j \mid \{r_1,\dots,r_N\})$	$O(N^2 D)$	$O(N^2)$
$\Delta\phi_i(r_j \mid \{r_1,\dots,r_N\})$	$O(N^3 D)$	$O(N^2)$
$\det[\phi_j^k(r_i)]_{i,j}$	$O(N^3 K)$	$O(N^3 K)$
$\Delta\det[\phi_j^k(r_i)]_{i,j}$	$O(N^4 K)$	$O(N^3 K)$
$f(\psi_\theta^1,\dots,\psi_\theta^K)$	$O(K)$	$O(KD)$
$\Delta f(\psi_\theta^1,\dots,\psi_\theta^K)$	$O(NK)$	$O(NKD)$

Can we improve accuracy and speed by using general non-linear combinations $f: \mathbb{R}^K \to \mathbb{R}$ of simple basis ψ_θ^k?

Symmetric Functions

Odd function:	$f(-\boldsymbol{x}) = -f(\boldsymbol{x})$	
Even function:	$g(-\boldsymbol{x}) = g(\boldsymbol{x})$	
Antisymmetric function:	$\psi(\pi(\boldsymbol{x})) = \operatorname{sgn}\pi\,\psi(\boldsymbol{x})$,	$\forall\pi \in S_N$
Symmetric function:	$J(\pi(\boldsymbol{x})) = J(\boldsymbol{x})$,	$\forall\pi \in S_N$

Implicit Odd Neural Networks

Composition of odd functions:

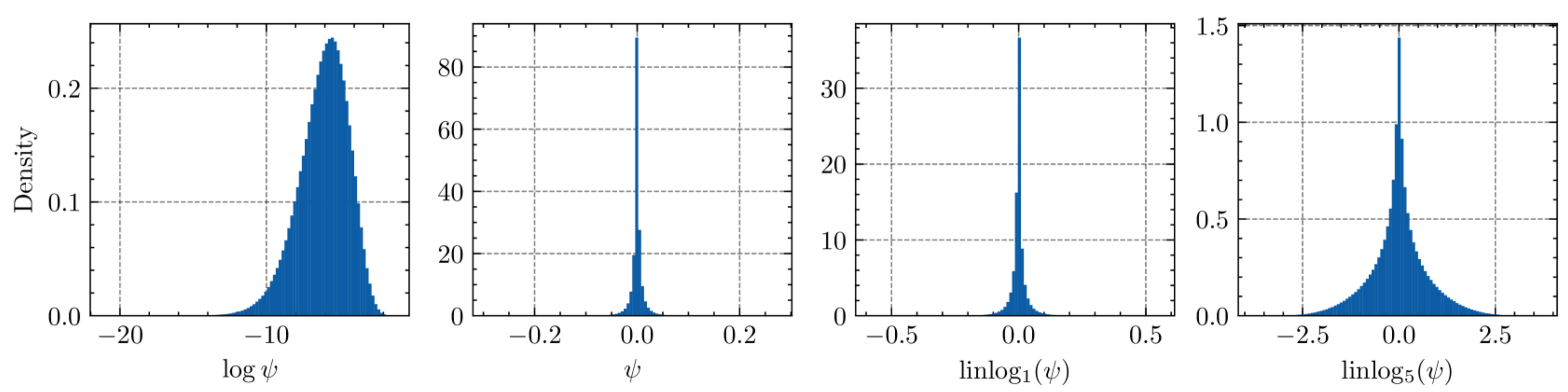
$$f^0(\boldsymbol{x}) = \boldsymbol{x}$$
$$f(\boldsymbol{x}) = f^T(\boldsymbol{x})$$
$$f^{t+1}(\boldsymbol{x}) = \tanh(f^t(\boldsymbol{x})W^t) \circ J^t(\boldsymbol{r})$$

Explicit Odd Neural Networks

Excplictly antisymmetrizing arbitrary functions:

$$f(\boldsymbol{x}) = g(\boldsymbol{x}) - g(-\boldsymbol{x})$$
$$g(\boldsymbol{x}) = \mathrm{MLP}(\boldsymbol{x})$$

Rescaling



$$\mathrm{Linlog}_\alpha(x) = \operatorname{sgn}(x)\log(|x|e^\alpha + 1)$$

Equivalence to Symmetric Jastrow

Theorem 1: Let R, X, Y, Z be real vector spaces, $\phi: R \to X$ an antisymmetric function, $J: R \to Y$ a symmetric function, and $f: X \times Y \to Z$ an in the first argument odd function, i.e., $f(-x, y) = -f(x, y)$. Any antisymmetric function $\psi: R \to Z; \psi(r) = f(\phi(r), J(r))$ can be expressed a.e. as $\psi(r) = (\phi(r)^T x)\hat{J}(r)$ with $x \in X\backslash\{0\}$ and a symmetric function $\hat{J}: R \to Z$.

➢ Non-linear combinations of $(\psi^1,\dots,\psi^K)$ are no more expressive in the presence of a symmetric function J.

Cusp Conditions

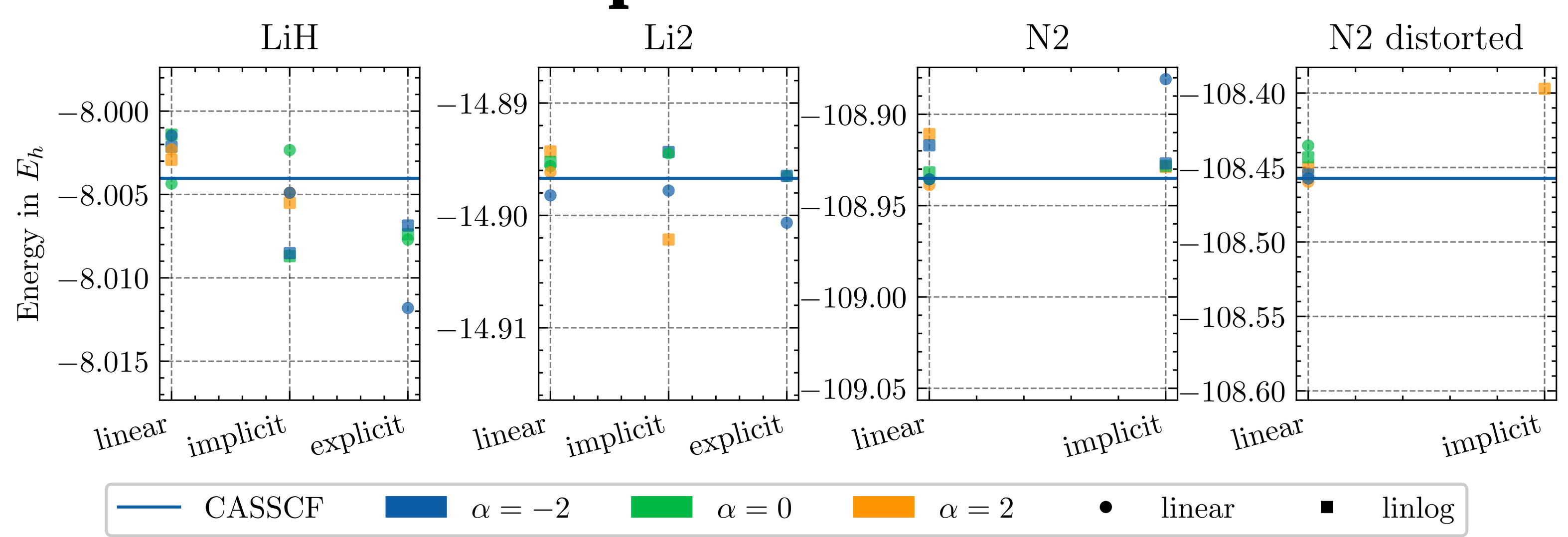
Kato's theorem states that the solutions must fulfill

$$\lim_{r\to R_m} -\frac{1}{\psi(\boldsymbol{r})}\frac{\partial\psi(\boldsymbol{r})}{\partial r} = Z_m.$$

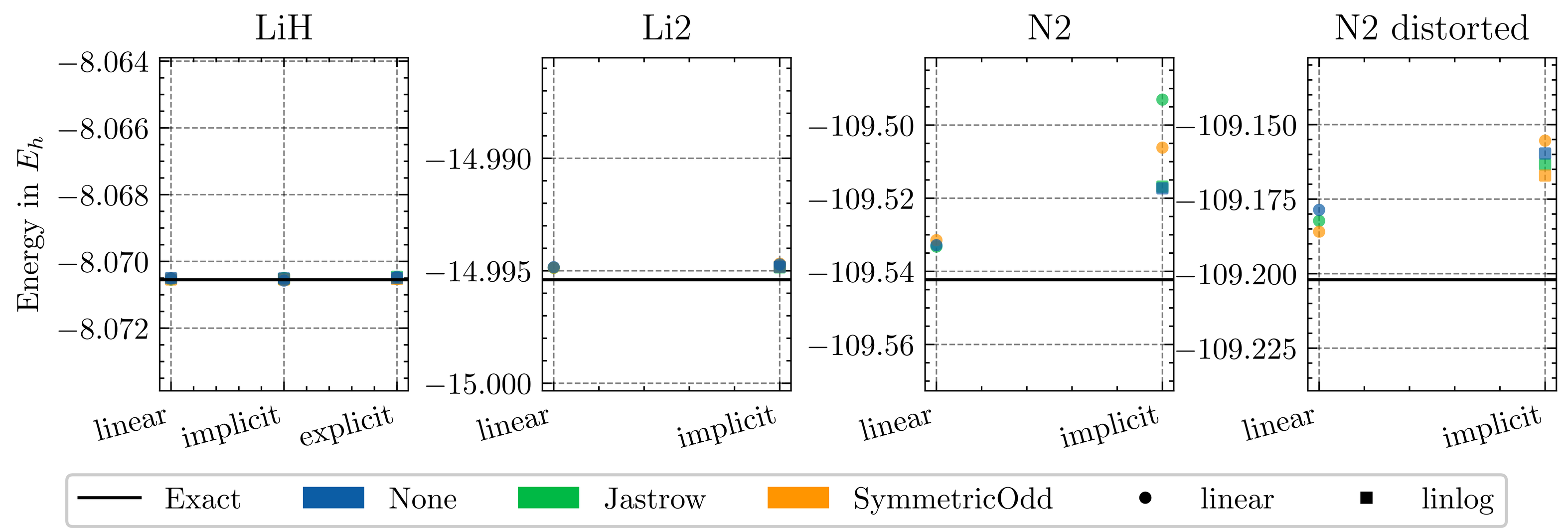
Gaussian-type orbitals (GTO)-based orbitals, cannot satisfy:

$$\frac{\partial\psi(\boldsymbol{r})}{\partial r} = \frac{\partial\psi(\boldsymbol{r})}{\partial\det[\phi_j^k(r_i)]_{i,j}} \underbrace{\frac{\partial\det[\phi_j^k(r_i)]_{i,j}}{\partial r}}_{0}.$$

Empirical Results



Complete Active Space Self Consistent Field (CASSCF) $\boldsymbol{\psi_\theta^k}$



FermiNet $\boldsymbol{\psi_\theta^k}$