Backflow in Neural Quantum States

Slater–Jastrow–Backflow Wavefunctions and Links to Transformers

MHJ

Outline

1. Slater–Jastrow wavefunctions (baseline for fermions)
2. Backflow: explicit coordinate/orbital transformations
3. Slater–Jastrow–Backflow (SJB) equations (continuum and lattice)
4. Why backflow improves nodal structure and correlations
5. Relation to Transformer-style architectures

Slater–Jastrow (SJ) variational ansatz

For N fermions with coordinates $\boldsymbol{R} = (\boldsymbol{r}_1, \ldots, \boldsymbol{r}_N)$, a standard correlated ansatz is the Slater–Jastrow form

$$\Psi_{\mathrm{SJ}}(\boldsymbol{R}) = \det\big[\phi_k(\boldsymbol{r}_i)\big]\; e^{J(\boldsymbol{R})}.$$

- $\det[\phi_k(\boldsymbol{r}_i)]$ enforces antisymmetry and defines the nodal surface.
- $e^{J(\boldsymbol{R})}$ (Jastrow factor) captures symmetric correlations (often pairwise).

A common two-body Jastrow form:

$$J(\boldsymbol{R}) = \sum_{i<j} u(r_{ij}), \qquad r_{ij} = |\boldsymbol{r}_i - \boldsymbol{r}_j|.$$

Limitations of Slater–Jastrow

- The Jastrow factor improves correlation energy but **does not change the nodes**:

$$\Psi_{\mathrm{SJ}}(\boldsymbol{R}) = 0 \iff \det[\phi_k(\boldsymbol{r}_i)] = 0.$$

- For strongly correlated fermions, the **nodal surface** is often the main source of error.
- Backflow is designed specifically to introduce flexible, configuration-dependent nodal structure.

Backflow: configuration-dependent coordinates

Backflow replaces bare coordinates $\boldsymbol{r}_i$ by *effective* coordinates

$$\tilde{\boldsymbol{r}}_i = \boldsymbol{r}_i + \boldsymbol{\xi}_i(\boldsymbol{R}),$$

where the backflow displacement $\boldsymbol{\xi}_i$ depends on the full configuration. A classical (pairwise) backflow form:

$$\boldsymbol{\xi}_i(\boldsymbol{R}) = \sum_{j \neq i} \eta(r_{ij})\,(\boldsymbol{r}_i - \boldsymbol{r}_j),$$

with a scalar backflow function $\eta(\cdot)$ (often parameterized).

Slater–Jastrow–Backflow (SJB) ansatz

The Slater determinant is evaluated at backflow coordinates:

$$\Psi_{\mathrm{SJB}}(\boldsymbol{R}) = \det\left[\phi_k(\tilde{\boldsymbol{r}}_i)\right]\, e^{J(\boldsymbol{R})}.$$

- Backflow alters the nodes:

$$\Psi_{\mathrm{SJB}}(\boldsymbol{R}) = 0 \iff \det[\phi_k(\tilde{\boldsymbol{r}}_i(\boldsymbol{R}))] = 0,$$

  so the nodal surface becomes an implicit, many-body object.
- $J(\boldsymbol{R})$ still captures symmetric correlation and cusp conditions.

Backflow as orbital deformation (equivalent view)

Instead of moving coordinates, one may view backflow as making orbitals configuration-dependent:

$$\phi_k(\boldsymbol{r}_i) \quad \longrightarrow \quad \phi_k\big(\boldsymbol{r}_i; \boldsymbol{R}\big) := \phi_k\big(\tilde{\boldsymbol{r}}_i(\boldsymbol{R})\big).$$

Thus, each particle experiences an *effective environment-dependent orbital* set by the positions of all other particles.

Lattice analogue: backflow features for occupations

For lattice fermions with occupations $\boldsymbol{n} = (n_1, \dots, n_L)$, a backflow-like construction introduces *configuration-dependent features*

$$\tilde{h}_i(\boldsymbol{n}) = h_i + \sum_{j \neq i} f_{ij}\, \Phi(n_i, n_j, \boldsymbol{n}),$$

where $\tilde{h}_i$ plays the role of an effective site feature/embedding.
A determinant-based ansatz may then use effective orbitals

$$\Psi(\boldsymbol{n}) \propto \det\big[\varphi_k(\tilde{h}_i(\boldsymbol{n}))\big] \times e^{J(\boldsymbol{n})},$$

with a Jastrow-like factor $J(\boldsymbol{n})$ capturing density–density correlations.

Neural backflow: learned displacements or embeddings

In NQS practice, backflow is often **neural**:

$$\boldsymbol{\xi}_i(\boldsymbol{R}) = \mathcal{N}_\theta(i;\ \boldsymbol{R})\,,$$

with $\mathcal{N}_\theta$ a symmetry-respecting neural network (e.g. equivariant). Then

$$\tilde{\boldsymbol{r}}_i = \boldsymbol{r}_i + \mathcal{N}_\theta(i; \boldsymbol{R}), \qquad \Psi_{\mathrm{SJB}}(\boldsymbol{R}) = \det[\phi_k(\tilde{\boldsymbol{r}}_i)]\, e^{J(\boldsymbol{R})}.$$

This generalizes pairwise backflow to *collective* many-body backflow.

Transformer view: tokens, context, and backflow

A Transformer builds *context-dependent* representations:

$$\text{token embedding } \boldsymbol{e}_i \quad \longrightarrow \quad \text{contextual embedding } \boldsymbol{h}_i(\{\boldsymbol{e}_j\}_{j=1}^{N}).$$

Backflow is analogous:

$$\boldsymbol{r}_i \quad \longrightarrow \quad \tilde{\boldsymbol{r}}_i(\boldsymbol{R}) = \boldsymbol{r}_i + \boldsymbol{\xi}_i(\boldsymbol{R}),$$

i.e. each particle coordinate (or feature) becomes **contextual**, conditioned on the entire configuration.

Self-attention as a backflow-like map (schematic)

In a Transformer layer, one often has

$$\boldsymbol{h}_i = \boldsymbol{e}_i + \sum_j \alpha_{ij}\, \boldsymbol{V}\boldsymbol{e}_j, \qquad \alpha_{ij} = \mathrm{softmax}_j\left(\frac{(\boldsymbol{Q}\boldsymbol{e}_i)\cdot(\boldsymbol{K}\boldsymbol{e}_j)}{\sqrt{d}}\right).$$

Compare with pairwise backflow:

$$\tilde{\boldsymbol{r}}_i = \boldsymbol{r}_i + \sum_{j\neq i} \eta(r_{ij})\,(\boldsymbol{r}_i - \boldsymbol{r}_j).$$

Analogy:

- weights α_{ij} or $\eta(r_{ij})$ quantify *influence* of j on i,
- both produce nonlocal, permutation-symmetric updates of per-particle features.

Why Transformer-style backflow is powerful for NQS

- Long-range correlations are captured naturally via attention over all particles.
- Backflow/attention directly produces **configuration-dependent nodes** (fermions).
- Parameter-efficient: global structure can be learned with fewer parameters than very deep MLPs.
- Symmetry handling: permutation symmetry and (with equivariance) rotational symmetry can be enforced.

Takeaway: backflow is the physics analogue of contextual embeddings; Transformers are a natural modern architecture to implement neural backflow maps.

Summary

- Slater–Jastrow:
$$\Psi_{\mathrm{SJ}}(\boldsymbol{R}) = \det[\phi_k(\boldsymbol{r}_i)]\, e^{J(\boldsymbol{R})}$$
captures correlations but leaves nodes fixed.
- Backflow introduces effective coordinates
$$\tilde{\boldsymbol{r}}_i = \boldsymbol{r}_i + \boldsymbol{\xi}_i(\boldsymbol{R})$$
and yields Slater–Jastrow–Backflow:
$$\Psi_{\mathrm{SJB}}(\boldsymbol{R}) = \det[\phi_k(\tilde{\boldsymbol{r}}_i)]\, e^{J(\boldsymbol{R})}.$$
- Backflow $\leftrightarrow$ contextualization: conceptually aligned with Transformer self-attention.