COMMUN. MATH. SCI.
Vol. 21, No. 5, pp. 1447–1453
© 2023 International Press

*FAST COMMUNICATION*

# GEOMETRY OF BACKFLOW TRANSFORMATION ANSATZE FOR QUANTUM MANY-BODY FERMIONIC WAVEFUNCTIONS*

HANG HUANG†, JOSEPH M. LANDSBERG‡, AND JIANFENG LU§

**Abstract.** Wave function ansatze based on the backflow transformation are widely used to parametrize anti-symmetric multivariable functions for many-body quantum problems. We study the geometric aspects of such ansatze, in particular we show that in general totally antisymmetric polynomials cannot be efficiently represented by backflow transformation ansatze at least in the category of polynomials. In fact, if there are $N$ particles in the system, one needs a linear combination of at least $O(N^{3N-3})$ determinants to represent a generic totally antisymmetric polynomial. Our proof is based on bounding the dimension of the source of the ansatze from above and bounding the dimension of the target from below.

**Keywords.** Backflow transformation; fermionic wavefunction; anti-symmetry; secant variety.

**AMS subject classifications.** 81V74; 70G75; 14N07.

## 1. Introduction

Finding efficient numerical methods for quantum many-body systems has been a long-standing challenge, due to the notorious curse of dimensionality and the Fermionic sign problems. Many approaches have been proposed over the years; a popular class of methods is known as the variational quantum Monte Carlo methods. The basic idea is as follows: Let $H$ denote the Hamiltonian operator of a quantum system, choose a class of functions $\mathcal{F}$ as a variational ansatz and solve

$$E_{\mathcal{F}} := \inf_{\Psi\in\mathcal{F}} \frac{\langle\Psi|H|\Psi\rangle}{\langle\Psi|\Psi\rangle} \tag{1.1}$$

for $\Psi$ and $E_{\mathcal{F}}$. For Fermionic systems, due to the Pauli exclusion principle, the wave function has to be totally antisymmetric. Thus, for a system with $N$ particles one takes $\mathcal{F}\subset\bigwedge^N L^2(\mathbb{R}^3)$, so that the above $E_{\mathcal{F}}$ gives a variational upper bound for the true ground state energy. Here $\bigwedge^N L^2(\mathbb{R}^3)$ denotes the anti-symmetric tensor product of $n$ copies of $L^2(\mathbb{R}^3)$; the single particle Hilbert space is $L^2(\mathbb{R}^3)$, i.e., we assume that each particle lives in $\mathbb{R}^3$ and have neglected the spin degree of freedom for simplicity.

The question now becomes the choice of $\mathcal{F}$. The most straightforward approach is to take $\mathcal{F}=\bigwedge^N L^2(\mathbb{R}^3)$, or more precisely, for numerical purposes, one chooses a finite dimensional subspace $V\subset L^2(\mathbb{R}^3)$ and takes $\mathcal{F}=\bigwedge^N V$, in the spirit of Galerkin's method in numerical analysis. This is however not practical for actual computations as

*Received: May 27, 2022; Accepted (in revised form): October 09, 2022. Communicated by Lexing Ying.
Landsberg was supported by NSF grant AF-1814254. Lu was supported in part by NSF grants DMS-2012286 and CHE-2737263. This project is an outcome of the IPAM program: *Tensor Methods and Emerging Applications to the Physical and Data Sciences* March 8, 2021 – June 11, 2021.

†Department of Mathematics, Texas A&M University, College Station, TX 77843-3368, USA (huanghang1109@gmail.com).
‡Department of Mathematics, Texas A&M University, College Station, TX 77843-3368, USA (jml@math.tamu.edu).
§Departments of Mathematics, Physics, and Chemistry, Duke University, Durham, NC 27708, USA (jianfeng@math.duke.edu).

the dimension of $\mathcal{F}$ grows exponentially with $N$, known as the curse of dimensionality. Therefore, a smaller class of functions needs to be fixed for the variational search.

The most well-known choice of $\mathcal{F}$, which is essentially the starting point of quantum chemistry, is the collection of Slater determinants, i.e.,

$$\mathcal{F}=\left\{\varphi_1\wedge\varphi_2\wedge\cdots\wedge\varphi_N \,|\, \varphi_i\in L^2(\mathbb{R}^3), \langle\varphi_i,\varphi_j\rangle_{L^2(\mathbb{R}^3)}=\delta_{ij}, i,j=1,\cdots,n\right\}. \tag{1.2}$$

This leads to the celebrated Hartree-Fock method. In practice one also uses a finite dimensional approximation to $L^2(\mathbb{R}^3)$. While useful as a first approximation, the class of Slater determinants is often too small to capture a good approximation to the ground state energy. The difference between the result of the Hartree-Fock method and the true ground state energy is called the correlation energy; more complicated variational ansatze have been proposed to reduce the error.

One of the first approaches to go beyond the Hartree-Fock ansatz is known as the Slater-Jastrow wave function, for which one considers the product of a Slater determinant with a totally symmetric function $g$ and hence the product is anti-symmetric. The function $g$ is often parametrized as

$$g(x_1,\cdots,x_N)=\exp\Big(\sum_{i\le j}U(|x_i-x_j|)\Big), \tag{1.3}$$

where $U$ is some function on $\mathbb{R}$. The $g$ given above is obviously totally symmetric, while more general ansatze for $g$ have also been proposed and studied. Unlike the Hartree-Fock method, it is no longer possible to explicitly evaluate the Rayleigh quotient, and thus the ansatz is optimized by Monte Carlo approaches in practice, i.e., variational quantum Monte Carlo methods. One generalization uses linear combinations of Slater-Jastrow wave functions, which is often referred to as the multi-configurational approach in the variational quantum Monte Carlo methods literature.

Another direction is to change the Slater determinants to some other anti-symmetric functions; for example, Pfaffians (when $N$ is even), Vandemonde determinants, and determinants with backflow transformations as defined below. This has become a very active field in recent years thanks to the rise of neural networks as a versatile ansatz for high dimensional functions, after the influential work [2] of parameterizing many-body wave functions using neural networks. Several variational classes have been proposed by replacing components in the anti-symmetric function ansatz by neural networks, see e.g., [1, 3, 6–11, 13]. While the details of these ansatze differ, the general framework is based on the backflow transformation originally proposed by Feynman and Cohen [4]. To introduce the ansatz, define the function class

$$\mathcal{S}=\left\{\varphi\in L^2\big(\mathbb{R}^3\times(\mathbb{R}^3\otimes\mathbb{R}^{N-1})\big) \,|\, \varphi(\mathbf{x};\overline{\boldsymbol{y}})=\varphi(\mathbf{x};\sigma\overline{\boldsymbol{y}}),\quad \forall\sigma\in\mathfrak{S}_{N-1}\right\}.$$

Thus, functions in $\mathcal{S}$ are totally symmetric with respect to the second argument in $\mathbb{R}^3\otimes\mathbb{R}^{N-1}$ (the permutation group $\mathfrak{S}_{N-1}$ acts on $\mathbb{R}^{N-1}$). It is easy to check then for $\varphi_i\in\mathcal{S}$, $i=1,\cdots,N$, the following function on $\mathbb{R}^3\otimes\mathbb{R}^N$ is totally anti-symmetric:

$$\Phi_{\mathrm{BF}}[\varphi_1,\ldots,\varphi_N](\mathbf{x}_1,\cdots,\mathbf{x}_N)=\det\begin{pmatrix}\varphi_1(\mathbf{x}_1;\overline{\mathbf{x}_{-1}}) & \cdots & \varphi_N(\mathbf{x}_1;\overline{\mathbf{x}_{-1}})\\ \vdots & \ddots & \vdots\\ \varphi_1(\mathbf{x}_N;\overline{\mathbf{x}_{-N}}) & \cdots & \varphi_N(\mathbf{x}_N;\overline{\mathbf{x}_{-N}})\end{pmatrix}$$

where $\overline{\mathbf{x}_{-i}}:=(\mathbf{x}_1,\cdots,\mathbf{x}_{i-1},\mathbf{x}_{i+1},\cdots,\mathbf{x}_N)$. This generalizes the Slater determinants, which correspond to the case that $\varphi_i$ only depends on the first variable, the Slater-Jastrow wave

functions, which correspond to absorbing a $g^{1/N}$ factor into $\varphi_i$'s, and the Vandermonde determinants. We call $\Phi_{\mathrm{BF}}$ the *ansatz map*.

In this work, we focus on the study of the ansatz given by the backflow transformation:

$$\mathcal{F}_{\mathrm{BF}}^N = \{\Phi_{\mathrm{BF}}[\varphi_1,\ldots,\varphi_N] \,|\, \varphi_i \in \mathcal{S}, i=1,\ldots,N\}.$$

We are interested in the representation power of the class of functions $\mathcal{F}_{\mathrm{BF}}^N$. In particular, we ask for a given totally antisymmetric function $\Psi:\mathbb{R}^3\otimes\mathbb{R}^N\to\mathbb{R}$, whether it belongs to $\mathcal{F}_{\mathrm{BF}}^N$, i.e., whether it is possible to find $\{\varphi_i\}$, such that $\Psi=\Phi_{\mathrm{BF}}[\varphi_1,\ldots,\varphi_N]$. This question may be asked in several different flavors, depending on the assumed function classes of $\Psi$ and the $\{\varphi_i\}$. In this work, we approach the question from the perspective of algebra. Thus, we consider $\Psi$ and hence the $\{\varphi_i\}$ to be polynomials.

Let $S^\delta\mathbb{R}^{3*}$ denote the space of homogeneous degree $\delta$ polynomials on $\mathbb{R}^3$. Consider

$$\bigwedge^N (L^2(\mathbb{R}^3))_{alg} := \bigwedge^N (\bigoplus_{\delta=0}^{\infty} S^\delta \mathbb{R}^{3*}),$$

where for any given element of $L^2(\mathbb{R}^3)_{alg}$, we only allow a finite number of $\delta$ to be used. Fix a total degree $D$ and consider $\bigwedge^N (L^2(\mathbb{R}^3))_{alg}\cap\mathbb{R}^{3*\otimes D}$ and call this the $(N,D)$-*space*. Taking the ratio of the estimates in Sections 3 and 4 we will see that for a given $N$, one must have $D$ at least on the order of $N^{\frac{4}{3}}$.

Our main result states that in general totally antisymmetric polynomials do not belong to $\mathcal{F}_{\mathrm{BF}}^N$. We stratify the subset $\mathcal{F}_{\mathrm{BF},poly}^N$ of $\mathcal{F}_{\mathrm{BF}}^N$ consisting of polynomials by total degree $D$ and write $\mathcal{F}_{\mathrm{BF},poly}^N = \bigoplus_D \mathcal{F}_{\mathrm{BF}}^{D,N}$. Then $\mathcal{F}_{\mathrm{BF}}^{D,N}$ is a semi-algebraic subset of the $(N,D)$-space $\bigwedge^N (L^2(\mathbb{R}^3))_{alg}\cap\mathbb{R}^{3*\otimes D}$. Note that the elements mapping to $\mathcal{F}_{\mathrm{BF}}^{D,N}$ are the $(\varphi_1,\ldots,\varphi_N)$ such that $\sum_{j=1}^N \deg(\varphi_j)=D$ and each $\varphi_j$ is homogeneous.

THEOREM 1.1. *For each fixed $N$, for all $D$ sufficiently large, the algebraic ansatz map is not surjective. The dimension of the target is at least $O(N^{3N-3})$ times larger than the dimension of the source.*

*In particular, when $D$ is sufficiently large, for all $r<N^{3N-3}$, the set of sums of $r$ elements of $\mathcal{F}_{BF}^{D,N}$ still lies in a proper subvariety of $\bigwedge^N (L^2(\mathbb{R}^3))_{alg}\cap\mathbb{R}^{3*\otimes D}$. In particular, it is a set of measure zero.*

Theorem 1.1 will follow immediately from the upper bound on the dimension of the source in §4 and the lower bound on the dimension of the target in §3.

REMARK 1.1. Our estimates are coarse, but they only assume $D>N^3$. The map will still fail to be surjective for smaller $D$ for all but very few admissible values of $D$.

REMARK 1.2. The ansatz map has positive dimensional fibers. It would be interesting to determine their dimensions.

REMARK 1.3. The closure of the image of the ansatz map is some variety invariant under the action of $GL_3$. It would be interesting to obtain geometric information about this variety.

Our theorem, to some extent, is a negative result for the representation power of the backflow transformation ansatz, at least in the category of polynomials, since the number of elements needed grows exponentially in $N$. Note that any analytic function is a limit of a sequence of polynomials. Restriction to homogeneous polynomials is not

a restriction as we may always project to homogeneous components, which will be the images of homogeneous $\phi_i$.

We remark that recent work [8] argues that any totally anti-symmetric function can be represented by the backflow ansatz (see [8, Theorem 7]), however the $\varphi_i$ used in the construction involve a sorting of coordinates $\mathbf{x}_i$ in "lexicographical" order, and are hence discontinuous and also impractical for actual computations. The work [5] establishes universal approximation of totally anti-symmetric funciton using backflow ansatz, however the approximation requires exponential number of degrees of freedom. The efficient approximation of totally anti-symmetric functions remains a challenging question.

**Organization.** In §2 we review standard results needed for the proof. In §3 we bound the dimension of the target from below and in §4 we bound the dimension of the source from above. The two estimates together prove Theorem 1.1. We conclude in §5 with geometric remarks.

## 2. Preliminaries

Let $\overline{p}_k(m)$ denote the number of partitions of $m$ with at most $k$ parts. The following estimates are standard. We include proofs for the convenience of the reader.

PROPOSITION 2.1. $\overline{p}_k(m) = \frac{1}{k!(k-1)!}m^{k-1} + O(m^{k-2})$.

*Proof.* Recall Faulhaber's formula: $1^k + 2^k + \cdots + m^k = \frac{m^{k+1}}{k+1} + O(m^k)$.
We have the induction formula

$$\overline{p}_k(m) = \sum_{i=0}^{\lfloor \frac{m}{k} \rfloor} \overline{p}_{k-1}(m - ik).$$

When $k=1$, we have $\overline{p}_k(m) = 1$. Assume by induction that $\overline{p}_u(m) = c_u m^{u-1} + O(m^{u-2})$ for all $u < k$. We prove it holds for $k$ and $c_u = \frac{1}{u!(u-1)!}$.

$$\begin{aligned}
\overline{p}_k(m) &= \sum_{i=0}^{\lfloor \frac{m}{k} \rfloor} \overline{p}_{k-1}(m - ik) \\
&= \sum_{i=0}^{\lfloor \frac{m}{k} \rfloor} c_{k-1}(m - ik)^{k-2} + O((m-ik)^{k-3}) \\
&= k^{k-2} c_{k-1} \sum_{i=0}^{\lfloor \frac{m}{k} \rfloor} (\frac{m}{k} - i)^{k-2} + O((m-ik)^{k-3}) \\
&= k^{k-2} c_{k-1} \frac{1}{k-1} (\frac{m}{k})^{k-1} + O(m^{k-2}) \\
&= \frac{1}{k(k-1)} c_{k-1} m^{k-1} + O(m^{k-2}).
\end{aligned}$$

We conclude by induction. ∎

PROPOSITION 2.2. *Let $\overline{q}_k(m)$ denote the number of partitions of $m$ with either $k$ or $k-1$ parts that are strictly decreasing. Then $\overline{q}_k(m + \binom{k}{2}) = \overline{p}_k(m)$.*

*Proof.* Given a partition of $m$ with at most $k$ parts, we may obtain a new partition that is strictly decreasing by adding one to the $(k-1)$-st part, two to the $(k-2)$-nd, up

to $(k-1)$ to the first. Moreover all strictly decreasing partitions with $k$ or $k-1$ parts arise in this way so we have established a bijection. ∎

Recall from Stirling's formula that $n! = \sqrt{2\pi n}(\frac{n}{e})^n(1+O(\frac{1}{n}))$ so in particular $n!(n-1)! = 2\pi e^{-2n+1}n^{2n}(1+O(\frac{1}{n}))$.

In what follows we will be concerned with estimates so we suppress round ups and round downs to integers from the notation.

The dimension of an algebraic variety is the dimension of the largest irreducible component, see, e.g., [12, §I.6.1]. A map between algebraic varieties given by polynomials will be surjective if and only if the restriction to some irreducible component of the source is surjective. Thus a necessary condition for surjectivity of a map is that the source contains a component of dimension at least that of the target. Because of this we may restrict to homogeneous polynomials and a single source.

**3. Target space**

We fix a total degree $D$, set $W = \mathbb{R}^3$, and lower bound the dimension of

$$\bigoplus_{\substack{0\le p_1\le p_2\le\cdots\le p_N\\ p_1+\cdots+p_N=D}} S^{p_1}W^*\otimes\cdots\otimes S^{p_N}W^*.$$

We only consider terms where $p_1 \ge D/2N$ and $p_1 < p_2 < \cdots < p_N$. The sum becomes

$$\sum_{\substack{q_1+\cdots+q_N=D/2\\ q_1<q_2<\cdots<q_N}} S^{q_1+\frac{D}{2N}}W^*\otimes\cdots\otimes S^{q_N+\frac{D}{2N}}W^*.$$

Recall that $\dim S^dW^* = \binom{d+2}{2}$. The smallest term in the summation is when $q_j = j-1$ for $j<N$ and $q_N = \frac{D}{2} - \binom{N-1}{2}$. This is because for integers $a_1 < a_2 < \cdots < a_N$ with $a_1 + \cdots + a_N$ fixed, the product $a_1^2\cdots a_N^2$ is minimized when all factors but the last are as small as possible. Assume $D > N^3$. This term has dimension

$$\binom{\frac{N+1}{2N}D - \binom{N-1}{2}+2}{2}\prod_{j=0}^{N-2}\binom{j+\frac{D}{2N}+2}{2} = \frac{D^{2N}}{2^{3N}N^{2(N-1)}} + O(D^{2N-1}).$$

Here $O(D^{2N-1})$ is considered as a function of $N$.

The number of terms is

$$\begin{aligned}\overline{q}_N(\frac{D}{2}) &= \frac{1}{N!(N-1)!}(\frac{D}{2} - \frac{N(N-1)}{2})^{N-1} + O((\frac{D}{2} - \frac{N(N-1)}{2}))^{N-2}\\ &= \frac{D^{N-1}e^{2N-1}}{\pi 2^N N^{2N}} + O(D^{N-2}).\end{aligned}$$

Thus the dimension of the target is bounded below by

$$\left[\frac{D^{N-1}e^{2N-1}}{\pi 2^N N^{2N}} + O(D^{N-2})\right]\left[\frac{D^{2N}}{2^{3N}N^{2(N-1)}} + O(D^{2N-1})\right] = \frac{D^{3N-1}e^{2N-1}}{\pi 2^{4N}N^{4N-2}} + O(D^{3N-2}).$$

**4. Source space**

Let $d_j = \deg\varphi_j$, then $\mathcal{F}_{\mathrm{BF}}^{D,N}$ is the union of the images of the ansatz maps over all $(\varphi_1,\ldots,\varphi_N)$ with $d_1 \le \cdots \le d_N$ and $\sum d_j = D$. By the discussion regarding dimension in algebraic geometry mentioned above, the ansatz is dominant onto the $(N,D)$ space if and only if one of the $(d_1,\ldots,d_N)$-ansatz maps is dominant.

Fix $d=d_j$ and suppress the $j$ index. Write

$$\varphi=\varphi(\mathbf{x}_1;\overline{\mathbf{x}_{-1}})=\sum_{z=0}^{d}\sum_{\substack{\delta_2+\cdots+\delta_N=d-z\\0\le\delta_2\le\delta_3\le\cdots\le\delta_N}} f_z^{z,\overline{\delta}}(\mathbf{x}_1)\sum_{\sigma\in\mathfrak{S}_{N-1}} h_2^{z,\overline{\delta}}(\mathbf{x}_{\sigma(2)})\cdots h_N^{z,\overline{\delta}}(\mathbf{x}_{\sigma(N)}).$$

Here $\overline{\delta}=(\delta_2,\ldots,\delta_N)$, $f_z^{z,\overline{\delta}}$ has degree $z$, and $h_j^{z,\overline{\delta}}$ has degree $\delta_j$. We are assuming without loss of generality that the $h$'s are non-decreasing in degree with $j$ because they appear symmetrically. On the other hand, we have to allow the degree of the $f$'s to be any value from 0 to $d$.

The dimension of the source space of $\varphi=\varphi_j$ with $j$ fixed and $d=d_j$ is thus

$$\sum_{z=0}^{d}\sum_{\substack{\delta_2+\cdots+\delta_N=d-z\\0\le\delta_2\le\delta_3\le\cdots\le\delta_N}}\left[\binom{z+2}{2}\binom{\delta_2+2}{2}\cdots\binom{\delta_N+2}{2}\right]$$

where $\binom{z+2}{2}$ is the dimension of the space of $f_z$'s of degree $z$ and $\binom{\delta_j+2}{2}$ is the dimension of the space of $h_{\delta_j}$'s of degree $\delta_j$.

The largest component of the source will be when $d_j=\frac{D}{N}=:d$ for all $j$. This is because for integers $a_1,\ldots,a_N$, with $a_1+\cdots+a_N$ fixed, the product $a_1^2\cdots a_N^2$ is maximized when all factors but the last are equal. We have the following upper bound for the dimension of the source:

$$\begin{aligned}
&N\sum_{z=0}^{d}\sum_{\substack{\delta_2+\cdots+\delta_N=d-z\\0\le\delta_2\le\delta_3\le\cdots\le\delta_N}}\left[\binom{z+2}{2}\binom{\delta_2+2}{2}\cdots\binom{\delta_N+2}{2}\right]\\
&<N\sum_{z=0}^{d}\binom{z+2}{2}\binom{\frac{d-z}{N-1}+2}{2}^{N-1}\overline{p}_{N-1}(d-z)\\
&<Nd\binom{\frac{d}{N}+2}{2}^{N}\overline{p}_{N-1}(d)\\
&=Nd\frac{d^{2N}}{2^N N^{2N}}\frac{d^{N-2}e^{2N-3}}{2\pi N^{2N-2}}+O(d^{3N-2})\\
&=\frac{D^{3N-1}e^{2N-3}}{N^{7N-4}2^{N+1}\pi}+O(D^{3N-2}).
\end{aligned}$$

The second line holds because the product of the binomial coefficients with the $\delta_i$'s is largest when they are all equal. The third holds because the largest term in the second line occurs when $z=\frac{d}{N}$, and the last by using approximations to the terms.

## 5. Geometric discussion

Although we did not make use of geometry in our proofs, our intuition was developed by computing small $D$ examples explicitly with the aid of algebraic geometry. We also believe such geometry will be useful in further study of ansatze along the lines of this paper so we provide a brief discussion here.

Given a vector space $V$ and an algebraic subset $X\subset V$, one can define the variety of points on $r$-secant planes of $X$:

$$\mathrm{Sec}_r(X):=\overline{\{v\in V\mid\exists x_1,\ldots,x_r\in X, v\in\mathrm{span}\{x_1,\ldots,x_r\}\}}.$$

For any $X$, a naïve dimension count gives

$$\dim \mathrm{Sec}_r(X) \leq r \dim(X) + r, \tag{5.1}$$

as one chooses $r$ points on $X$ and a point in their span. These secant varieties (or more precisely, their cousins in projective space) are intensely studied in algebraic geometry.

In our case a natural question is to take $X$ to be the image of an ansatz map and to ask when its secant varieties fill the ambient space.

**Acknowledgement.** This project is an outcome of the IPAM program *Tensor Methods and Emerging Applications to the Physical and Data Sciences* March 8, 2021 to June 11, 2021. The authors thank IPAM for bringing us (virtually) together and Landsberg thanks the Clay foundation for the opportunity to serve as a senior scholar during the program.

## REFERENCES

[1] A. Acevedo, M. Curry, S.H. Joshi, B. Leroux, and N. Malaya, *Vandermonde wave function ansatz for improved variational Monte Carlo*, 2020 IEEE/ACM Fourth Workshop on Deep Learning on Supercomputers (DLS), IEEE, 40–47, 2020. 1

[2] G. Carleo and M. Troyer, *Solving the quantum many-body problem with artificial neural networks*, Science, 355(6325):602–606, 2017. 1

[3] K. Choo, A. Mezzacapo, and G. Carleo, *Fermionic neural-network states for ab-initio electronic structure*, Nature Commun., 11(1):1–7, 2020. 1

[4] R.P. Feynman and M. Cohen, *Energy spectrum of the excitations in liquid helium*, Phys. Rev., 102(5):1189, 1956. 1

[5] J. Han, Y. Li, L. Lin, J. Lu, J. Zhang, and L. Zhang, *Universal approximation of symmetric and anti-symmetric functions*, Commun. Math. Sci., 20(5):1397–1408, 2022. 1

[6] J. Han, L. Zhang, and W. E, *Solving many-electron Schrödinger equation using deep neural networks*, J. Comput. Phys., 399:108929, 2019. 1

[7] J. Hermann, Z. Schätzle, and F. Noé, *Deep-neural-network solution of the electronic Schrödinger equation*, Nat. Chem., 12(10):891–897, 2020. 1

[8] M. Hutter, *On representing (anti) symmetric functions*, arXiv preprint, arXiv:2007.15298, 2020. 1, 1

[9] D. Luo and B.K. Clark, *Backflow transformations via neural networks for quantum many-body wave functions*, Phys. Rev. Lett., 122(22):226401, 2019. 1

[10] D. Pfau, J.S. Spencer, A.G.D.G. Matthews, and W.M.C. Foulkes, *Ab initio solution of the many-electron Schrödinger equation with deep neural networks*, Phys. Rev. Res., 2(3):033429, 2020. 1

[11] Z. Schätzle, J. Hermann, and F. Noé, *Convergence to the fixed-node limit in deep variational Monte Carlo*, J. Chem. Phys., 154(12):124108, 2021. 1

[12] I.R. Shafarevich, *Basic Algebraic Geometry 1: Varieties in Projective Space*, Springer, Heidelberg, Third Edition, 2013. 2

[13] J.S. Spencer, D. Pfau, A. Botev, and W.M.C. Foulkes, *Better, faster fermionic neural networks*, arXiv preprint, arXiv:2011.07125, 2020. 1