Machine learning many-electron wave functions via backflow transformations

1. **Backflow Transformations via Neural Networks for Quantum Many-Body Wave-Functions**
   Authors: D. Luo and B. K. Clark
   Phys. Rev. Lett. 122, 226401 (2019); arXiv:1807.10770

2. **Ab-Initio Solution of the Many-Electron Schrödinger Equation with Deep Neural Networks**
   Authors: D. Pfau, J. S. Spencer, A. G. de G. Matthews, and W. M. C. Foulkes
   arXiv:1909.02487

3. **Deep neural network solution of the electronic Schrödinger equation**
   Authors: J. Hermann, Z. Schätzle, and F. Noé
   arXiv:1909.08423

Recommended with a Commentary by Markus Holzmann, Univ. Grenoble Alpes, CNRS, LPMMC, 38000 Grenoble, France

The goal of determining the electronic structure of molecules and materials by solving the many-body Schrödinger equation has challenged theoretical physics and chemistry over the last century and driven the development of powerful approximations and computational methods. In the three papers above [1, 2, 3], the authors show how deep learning architectures can systematically improve many-body wave functions for Quantum Monte Carlo calculations, and benchmark their accuracy on the Hubbard model [1], and for light atoms and small molecules [2, 3]. Out of several recent approaches for machine learning quantum-many-body wave functions, e.g. [4, 5, 6, 7], I have chosen these three papers biased by the common underlying strategy taken there to explore, extend, and possibly exhaust existing structures of many-body trial wave functions based on backflow transformations employing neural networks.

The variational principle for the ground state energy, $E_0 \leq \int d\mathbf{R}\Psi_T^*(\mathbf{R})H\Psi_T(\mathbf{R})$, provides a simple, but powerful tool to obtain upper bounds for the ground state energy, E_0, of any Hamiltonian H. Combined with Monte Carlo methods to evaluate the highly dimensional integral over all particle coordinates, variational and quantum Monte Carlo calculations have provided most accurate values of the many-body Schrödinger equation [8, 9], only limited by the quality of the underlying trial wave function, $\Psi_T(\mathbf{R})$, for N fermions, $\mathbf{R} \equiv (\mathbf{r}_1, \ldots, \mathbf{r}_N)$. Can representations based on neural networks reduce this remaining bias similar successful

as for discrete spin lattice systems [10], and efficiently capture the sign structure of fermions as well as the continuous space description?

Requesting a manifest antisymmetric form essentially reduces the wave function to a determinant (or pfaffian) as the only known antisymmetric form with affordable computational cost for medium/large number of fermions. This yields the so-called Slater-Jastrow wave function $\Psi_T \sim \det_{ni} \phi_n(\mathbf{r}_i) \exp[-U(\mathbf{R})]$, where the Slater determinant is usually build out of orbitals $\phi_n(\mathbf{r})$ from independent particle/ self-consistent field type models and augmented by a (real) symmetric correlation factor $U(\mathbf{R})$. Eventually, the accuracy of quantum Monte Carlo calculations is limited by the sign structure of the trial wave function, which, within Slater-Jastrow, essentially reduces to that of an independent particle wave function.

How to improve the sign-structure of trial wave functions beyond Slater-Jastrow? Including linear superpositions of determinants, the ansatz eventually becomes exact, but the number of determinants needed to significantly reduce the bias turns out to grow exponentially in general [11]. Alternatively, backflow transformations from bare coordinates $\mathbf{r}_i$ to quasi-particle coordinates $\mathbf{q}_i(\mathbf{R}) = \mathbf{r}_i + \sum_j \mathbf{r}_{ij} \eta(r_{ij})$, symmetrically depending on the position of all particles, can be used to modify the structure of the determinant $\det_{ni} \phi_n(\mathbf{q}_i)$. Originally introduced by Feynman and Cohen [12] for the description of the excitation spectra in liquid helium, Pandharipande and Itoh [13] showed that the backflow structure naturally arises also in ground state wave functions from the momentum dependence of correlations. Introduction of backflow wave functions in quantum Monte Carlo calculations in various systems over the last decades [14, 15, 16, 17, 18, 19, 20] has systematically increased the accuracy of a single determinant description. Iterative applications of backflow transformations [21], yielding a non-linear network [22], indicate one possible pathway for further improvement of backflow functions.

In their paper [1], Di Luo and Bryan Clark have chosen a feed-forward neural network to represent the backflow wave functions for the ground state wave function of lattice (Hubbard) models, whereas David Pfau et al. [2] and Jan Hermann et al. [3] provide two different deep neural network implementations of backflow transformations for real-space electronic wave functions of first-row atoms and small molecules. Combined with machine learning algorithms, they were able to optimize these highly flexible networks and benchmark the quality obtained. The results presented in these papers show that neural network representations of backflow transformations can improve Slater-Jastrow results by order of magnitudes close to reaching chemical accuracy within a single determinant variational wave function.

The three articles [1, 2, 3] thus represent an important step forward towards fully abinitio electronic structure calculations where the precision of the results can be judged intrinsically, without making use to external (experimental) knowledge. Importantly, increasing the number of neurons/layers combined with variance extrapolations can help to quantify or even bound the residual bias. The variational description reached in these papers yields high quality wave functions with unpreceeded flexibility for electronic structure calculations, demonstrating once again the benefits from merging recent developments in the field of machine learning with those of quantum many-body calculations.

References

[1] Di Luo and B. K. Clark, Phys. Rev. Lett. 122, 226401 (2019).

[2] D. Pfau, J. S. Spencer, A. G. de G. Matthews, and W. M. C. Foulkes, arXiv:1909.02487.

[3] J. Hermann, Z. Schätzle, and Frank Noé, arXiv:1909.08423.

[4] Y. Nomura, A. S. Darmawan, Y. Yamaji, and M. Imada, Phys. Rev. B 96, 205152 (2017).

[5] L. R. Schwarz, A. Alavi, and G. H. Booth, Phys. Rev. Lett. **118**, 176403 (2017).

[6] J. Han, L. Zhang, and W. E, arXiv:1807.07014.

[7] K. Choo, A. Mezzacapo, and G. Carleo, arXiv:1909.12852.

[8] W.L. McMillan, Phys. Rev. **138**, A442 (1965).

[9] D. M. Ceperley and B. J. Alder, Phys. Rev. Lett. **45**, 566 (1980).

[10] G. Carleo and M. Troyer, Science **355**, 602 (2017).

[11] M. A. Morales, J. McMinis, B. K. Clark, J. Kim, and G. E. Scuseria, J. Chem. Theory Comput. **8**, 2181 (2012).

[12] R. P. Feynman and M. Cohen, Phys. Rev. **102**, 1189 (1956).

[13] V. R. Pandharipande and N. Itoh, Phys. Rev. A **8**, 2564 (1973).

[14] M.A. Lee, K.E. Schmidt, M.H. Kalos, and G.V. Chester, Phys. Rev. Lett. **46**, 728 (1981).

[15] Y. Kwon, D.M. Ceperley, and R.M. Martin, Phys. Rev. B **48**, 12 037 (1993).

[16] M. Holzmann, D. M. Ceperley, C. Pierleoni, and K. Esler, Phys. Rev. E **68**, 046707 (2003).

[17] P. López Ríos, A. Ma, N. D. Drummond, M. D. Towler, and R. J. Needs, Phys. Rev. E **74**, 066701 (2006).

[18] M. Holzmann, B. Bernu, and D.M. Ceperley, Phys. Rev. B **74**, 104510 (2006).

[19] L.F. Tocchio, F. Becca, A. Parola, and S. Sorella, Phys. Rev. B **78**, 041101 (2008).

[20] M. Holzmann and S. Moroni, Phys. Rev. B **99**, 085121 (2019).

[21] M. Taddei, M. Ruggeri, S. Moroni, and M. Holzmann, Phys. Rev. B **91**, 115106 (2015).

[22] M. Ruggeri, S. Moroni, and M. Holzmann, Phys. Rev. Lett. **120**, 205302 (2018).