Compute less, discover more – towards faster electronic structure modeling

Efficient optimization of neural-network electronic wave functions

***T. Grutschus, Z. Schätzle, P.B. Szabó, M. Mezera, F. Noé**, Department of Mathematics and Computer Science, Freie Universität Berlin, Arnimallee 12, 14195 Berlin, Germany*

In Short

- Accurate simulation of electrons enables better materials and medicines, but exact quantum calculations are only feasible for the very smallest systems.
- Neural network variational Monte Carlo (VMC) reaches high accuracy with favorable scaling, yet remains costly to train.
- We attempt to speed up sampling with a surrogate sampler and reduce the cost of wave function evaluations by introducing approximate antisymmetrizers.
- Fewer and cheaper wave function evaluations, together with lower sample autocorrelation, are expected to cut wall-clock time and speed up convergence while maintaining accuracy.

The behavior of electrons in molecules on the quantum-mechanical level can explain a broad range of macroscopic phenomena. An accurate model of the electronic structure of molecules, therefore, would facilitate the prediction and optimization of material properties in simulations, benefiting many fields, ranging from drug discovery to material development. While we have a conceptually clear description of the electronic processes, computing solutions to the equations governing the electrons' behavior is difficult. Solutions to these equations are functions, called wave functions, that describe the state of a quantum system.

The primary reason is that the computational resources required to find the correct wave functions increase exponentially with the size of the systems under consideration. Thus, for any but the most trivial systems, such as a single hydrogen atom in a vacuum, we cannot compute the exact behavior of electrons. However, over the past decades, a variety of methods that approximate solutions have been developed. Typically, they trade off accuracy and the computational resources necessary.

One promising method for approximating solutions to the electronic structure problem is variational Monte Carlo (VMC). Here, the wave functions are explicitly modeled as parameterized functions and, in our case, as deep neural networks. To optimize the parameters of the neural network, we minimize the energy the system would have in the state given by the neural wave function. Computing the energy requires samples from the squared magnitude of the wave function, which are typically drawn using Markov chain Monte Carlo (MCMC) methods. These samples are then used to compute Monte Carlo estimates of the energy. This variant of VMC, employing neural networks as parameterized wave functions, has achieved extraordinary accuracy while scaling favorably with system sizes [1,3]. While this means that the computing resources required do not grow as rapidly with system size as those of other highly accurate methods, training neural networks to find wave functions remains computationally expensive.

One of the first programs to enable large-scale training of neural wave functions is the open-source DEEPQMC library [2], developed in our group. It delivers state-of-the-art performance both in computational efficiency by leveraging modern compiler technology to fully utilize multiple high-end graphics processing units in parallel, and in accuracy by incorporating many recent advances in the field. Building on this package, the computational resources provided by NHR enable us to apply neural network VMC to relevant systems.

One aspect that makes electronic wave function training expensive, even on modern accelerators, is the special structure the functions must obey. Specifically, they have to be antisymmetric in the exchange of electron coordinates, such that if two same-spin electrons swap places, the wave functions must change their sign. Traditionally, in quantum chemistry, this is achieved with Slater determinants, which are antisymmetric by construction. In the context of neural network VMC, this choice, however, introduces a computationally expensive operation in every evaluation of the wave function.

At the same time, each training and evaluation iteration requires new samples from the wave function. Since these samples are drawn using MCMC, an iterative method, many wave function evaluations are required at each step.

Here, our project aims to speed up the training and evaluation process of neural network wave functions. In particular, we propose departing from the Slater determinant as the primary method for antisymmetrization and approximating antisymmetry with efficient antisymmetrizers. Furthermore, we

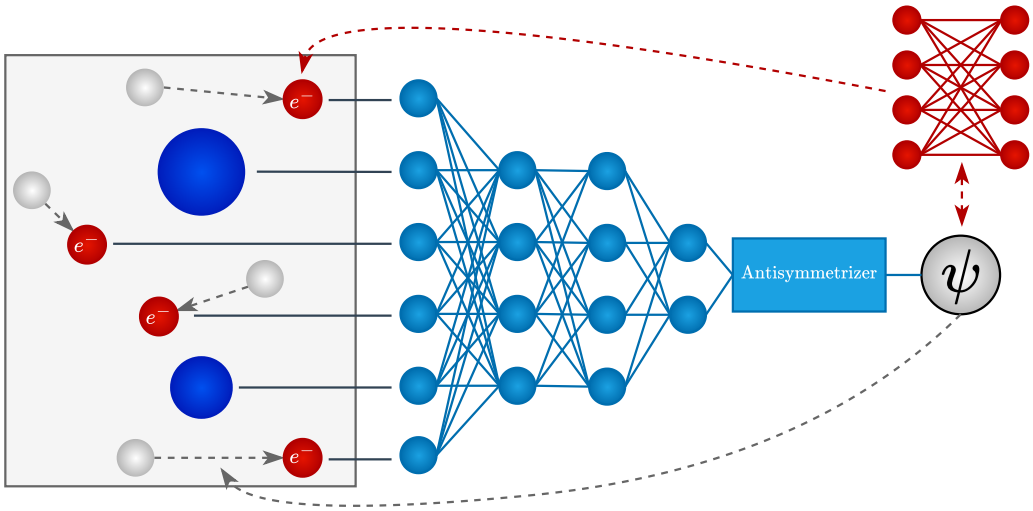

Figure 1: *Schematic illustration of a variational Monte Carlo calculation. Traditionally, the antisymmetrized wave function is used in a Markov chain to sample new electron positions (represented in grey). We propose a surrogate model for directly sampling new electron positions (represented in red). Figure adapted from [3].*

aim to reduce the computational burden during sampling: instead of running a Markov chain, we train a lightweight surrogate neural network in parallel to generate samples. Combined, these approaches promise to reduce the number of wave function evaluations required at each step, as well as the computational cost of these evaluations.

WWW

https://www.mi.fu-berlin.de/en/math/groups/ai4s/index.html

More Information

[1] J. Hermann, J. Spencer, K. Choo, A. Mezzacapo, W. M. C. Foulkes, D. Pfau, G. Carleo, and F. Noé, *Nat. Rev. Chem.* **7**(10), 692–709 (2023). doi:10.1038/s41570-023-00516-8

[2] Z. Schätzle, P. B. Szabó, M. Mezera, J. Hermann, and F. Noé, *J. Chem. Phys.* **159**, 094108 (2023). doi:10.1063/5.0157512

[3] J. Hermann, Z. Schätzle, and F. Noé, *Nat. Chem.* **12**, 891 (2020). doi:10.1038/s41557-020-0544-y

Funding

DFG Project EXC-2046/1

DFG Subject Area

327-01 Theoretical Chemistry: Electron Structure, Dynamics, Simulation