

Simulations of state-of-the-art fermionic neural network wave functions with diffusion Monte Carlo

Max Wilson, Nicholas Gao, Filip Wudarski, Eleanor Rieffel & Norm Tubman

Simulations of state-of-the-art fermionic neural network wave functions with diffusion Monte Carlo

Max Wilson*

*Quantum Engineering CDT, Bristol University, Bristol, BS8 1TH, UK
QuAIL, NASA Ames Research Center, Moffett Field, California 94035, USA and
KBR, Inc., 601 Jefferson St., Houston, TX 77002, USA*

Nicholas Gao

*QuAIL, NASA Ames Research Center, Moffett Field, California 94035, USA
KBR, Inc., 601 Jefferson St., Houston, TX 77002, USA
German Aerospace Center (DLR), Linder Höhe 51147 Köln, Germany and
Technical University of Munich, Boltzmann Str. 3 85748 Garching, Germany*

Filip Wudarski[†]

*QuAIL, NASA Ames Research Center, Moffett Field, California 94035, USA and
USRA Research Institute for Advanced Computer Science, Mountain View, California 94043, USA*

Eleanor Rieffel and Norm M. Tubman[‡]

QuAIL, NASA Ames Research Center, Moffett Field, California 94035, USA

(Dated: March 17, 2021)

Recently developed neural network-based *ab-initio* solutions (Pfau et. al arxiv:1909.02487v2) for finding ground states of fermionic systems can generate state-of-the-art results on a broad class of systems. In this work, we improve the results for this Ansatz with Diffusion Monte Carlo. Additionally, we introduce several modifications to the network (Fermi Net) and optimization method (Kronecker Factored Approximate Curvature) that reduce the number of required resources while maintaining or improving the modelling performance. In terms of the model, we remove redundant computations and alter the way data is handled in the permutation equivariant function. The Diffusion Monte Carlo results exceed or match state-of-the-art performance for all systems investigated: atomic systems Be-Ne, and the carbon cation C⁺.

Outline

The problem

- What are we trying to do?
- Why would we want to do that?

Related work

- What has been done before?
- What is related now?

The model

- What did we use to model the system?

The algorithms

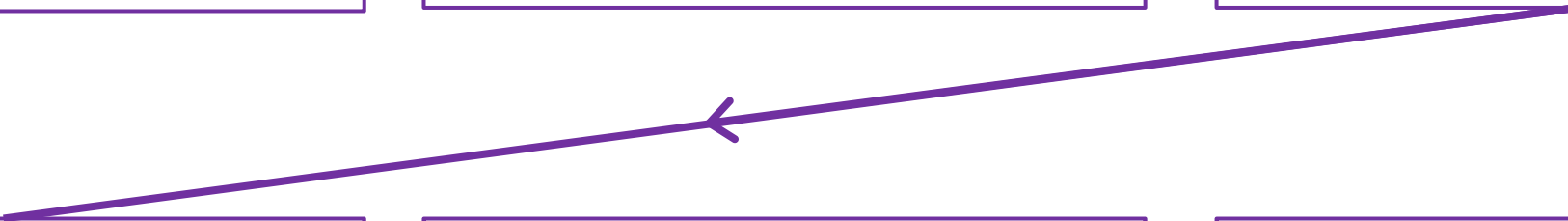
- What algorithms were used to solve the problem?

Some results

- What results did we get?

Next steps

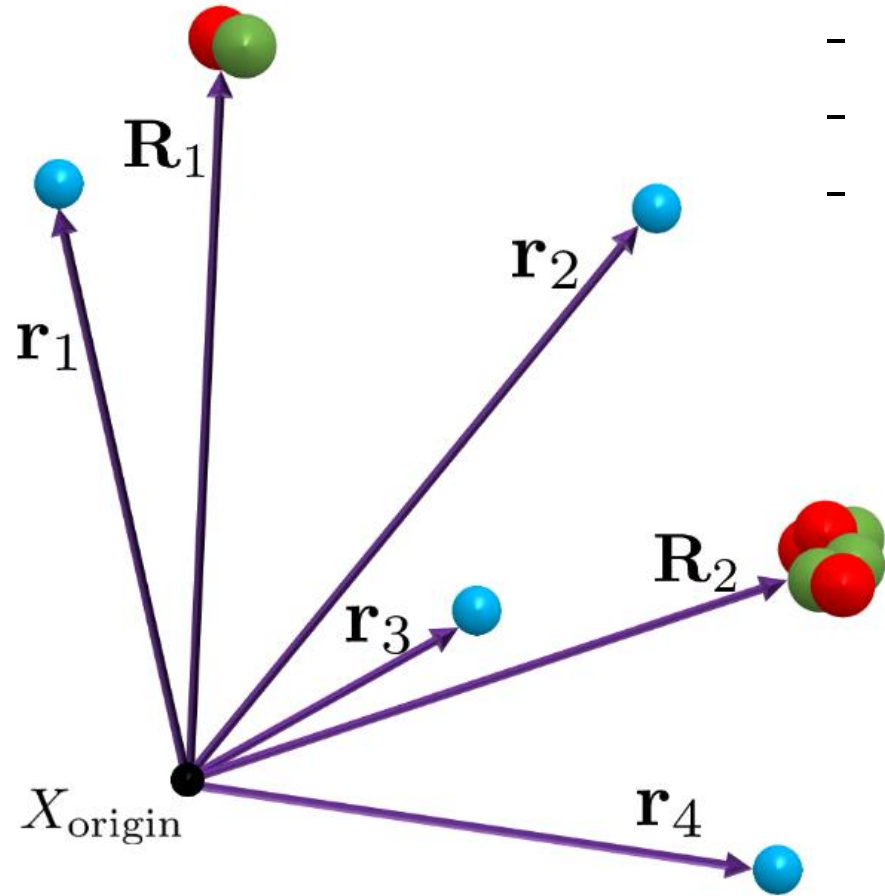
- Where is this going?



The problem

Quantum particles/systems

- **Superposition** (waves)
- **Entanglement** (non-classical correlations)
- **Exponential** amount of information to model



$$\hat{H}\psi(X) = E\psi(X)$$

Related work

Solving the QMB with ANNs

[Carleo & Troyer, Science 10, 2017](#)

- Spin system wave functions
- Boltzmann machine
- First demonstration

SchNet

[Schütt *et al*, J. Chem. Phys. 148, 2019](#)

- Atomic/Molecular properties predictions
- Graph NN

Fermi Net

[Pfau *et al*, Nature Chemistry 12, 2020](#)

- Atomic/molecular wave functions
- Multi-electron orbitals
- Graph NN
- Expensive, but very accurate

PauliNet

[Hermann *et al*, Nature Chemistry 12, 2020](#)

- Atomic/molecular wave functions
- Slater-Jastrow-backflow wave function
- Graph NN
- Fast, but less accurate

Related work

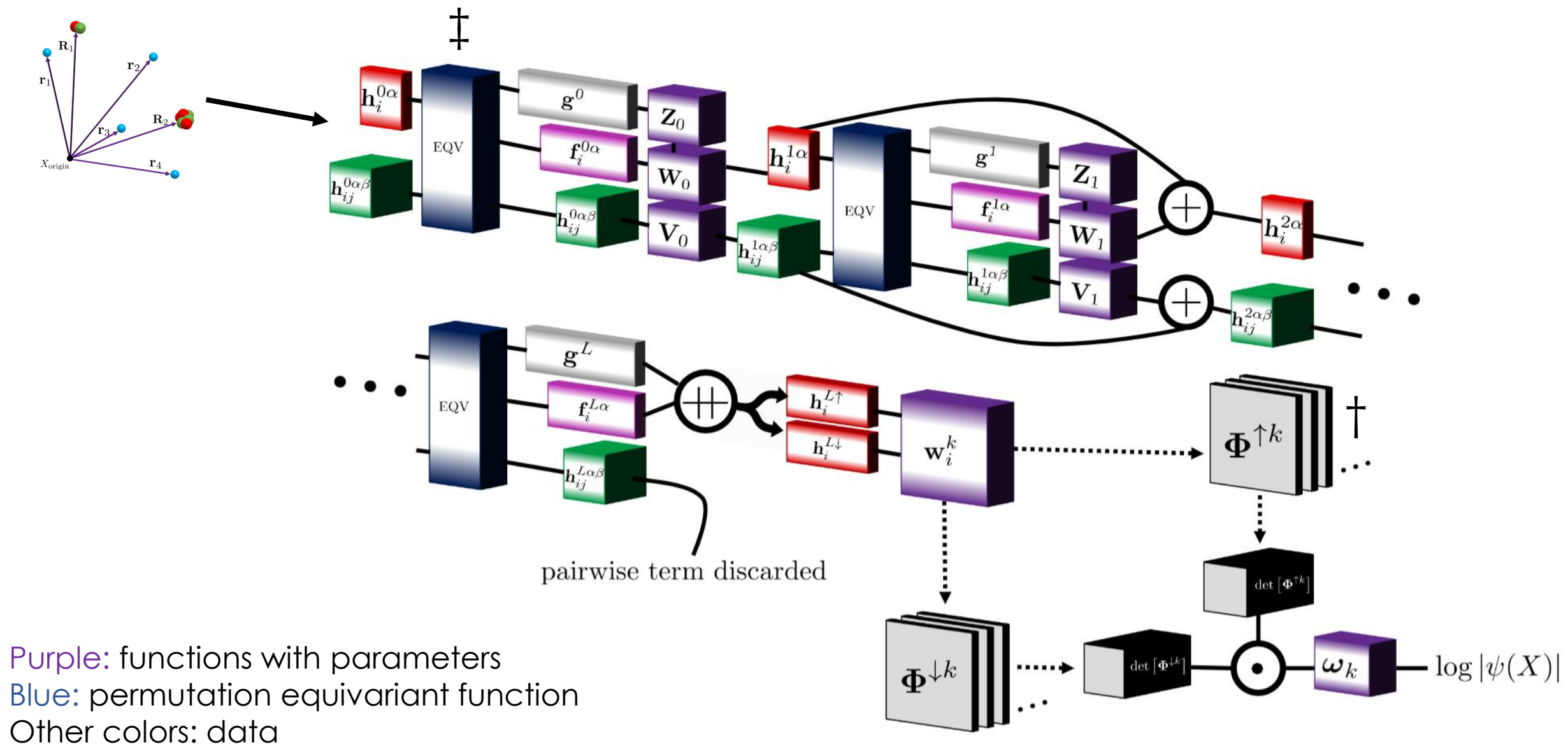
Contributions of this work

- **Improvements in efficiency** to Fermi Net
- **First (known to us) application of DMC** to neural network Ansatz
- **Matching or exceeding state-of-the-art results** on atomic systems Be-Ne and ion C^+

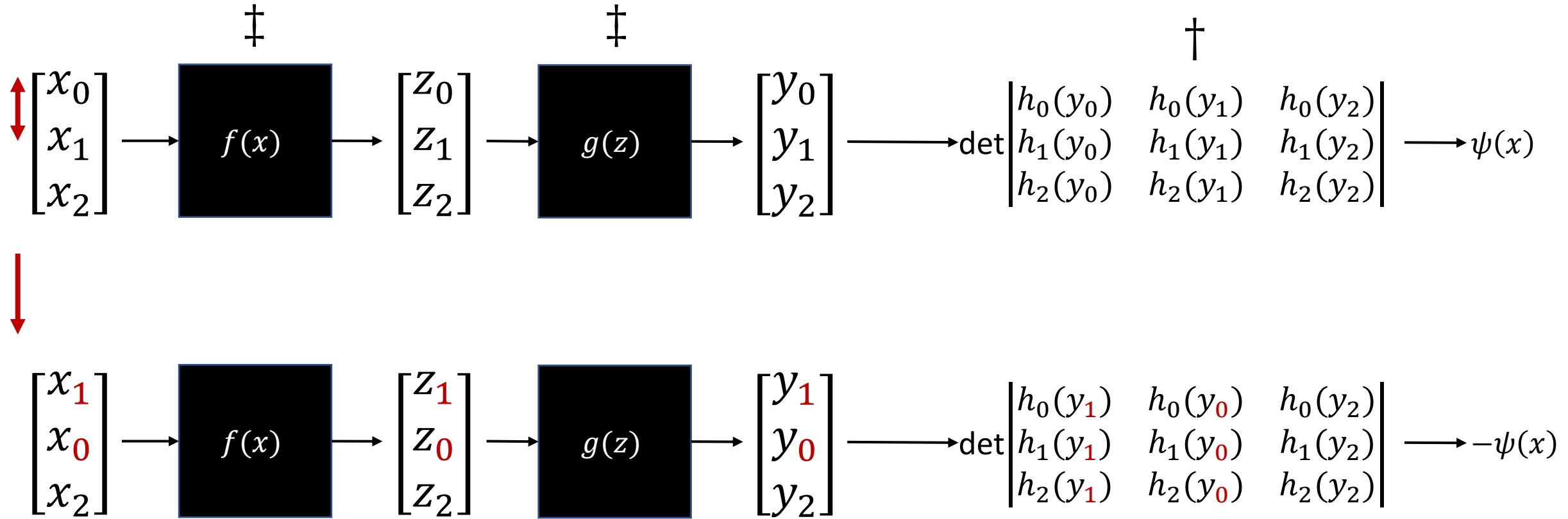
Other differences

- Changes to the optimisation (KFAC)

The model



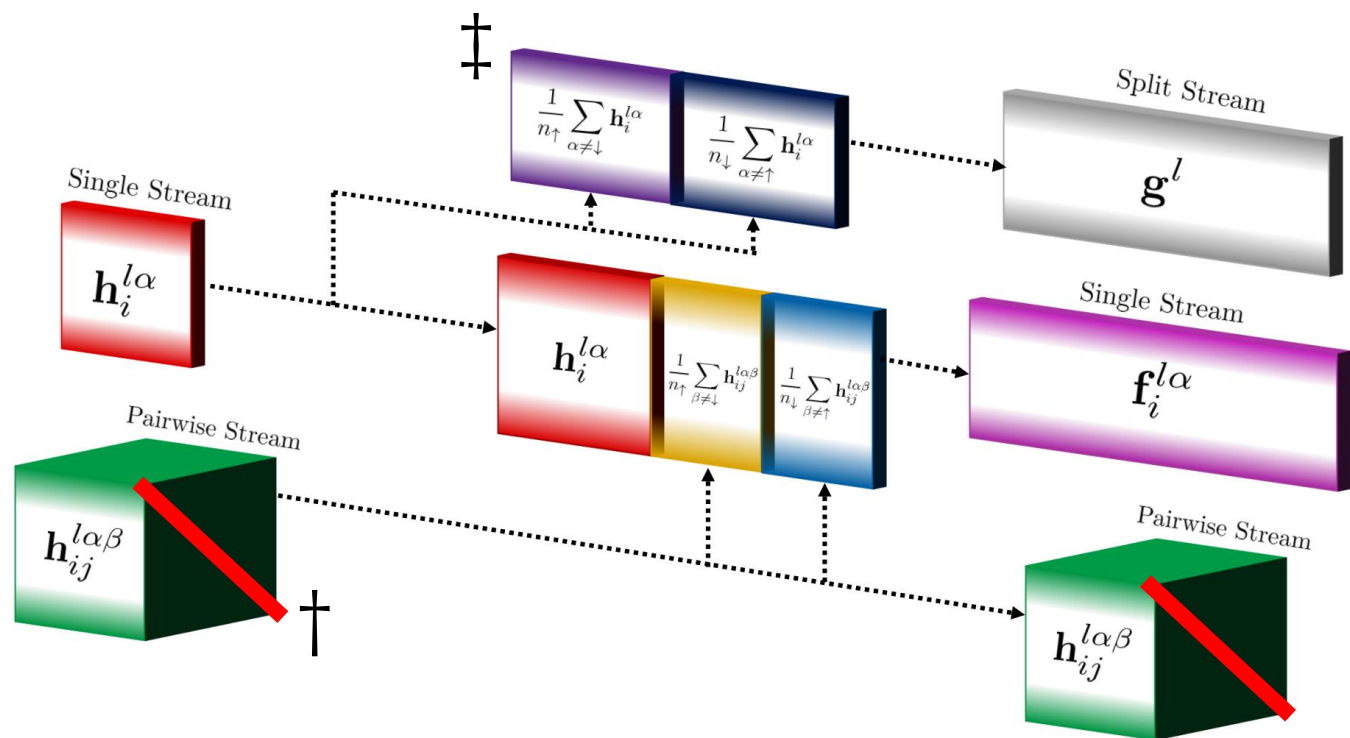
The model



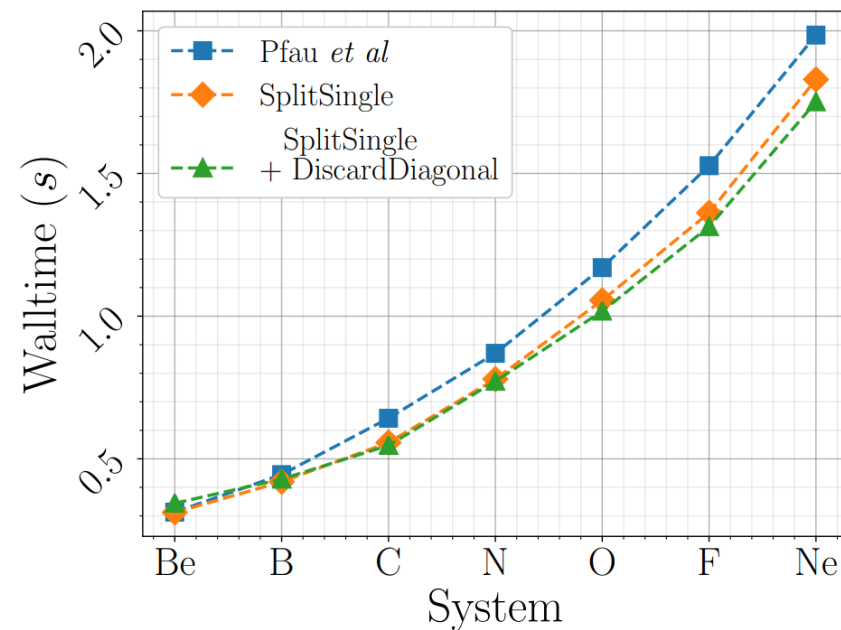
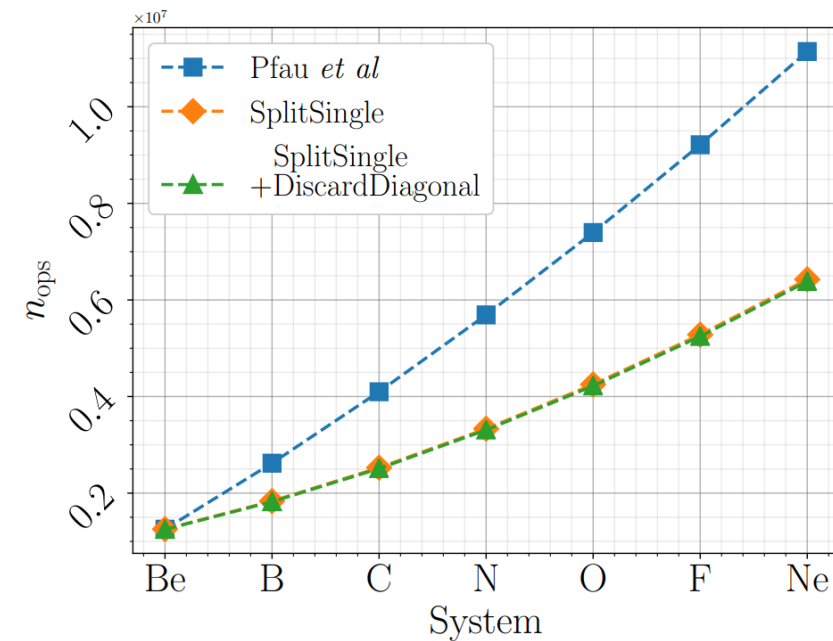
Swapping inputs results in change of sign of the wave function = antisymmetric = fermionic system

- Pfaffian wave function <https://pubmed.ncbi.nlm.nih.gov/16711968/>
- Vandermonde wave function <https://ieeexplore.ieee.org/document/9297114/>
- ...

The model



† Remove redundant diagonal elements $i = j$
 ‡ Split data that is the same in all indices



The algorithms

Variational Monte Carlo **VMC**

Iterative improvement of the wave function Ansatz by changing the parameters performed under the Born-Oppenheimer approximation (fixed nuclei)

Kronecker Factored Approximate Curvature **KFAC**

Approximate second order optimization related to stochastic reconfiguration

Diffusion Monte Carlo **DMC**

Projector method iteratively improving the wave function by reweighting the walkers performed under the fixed-node approximation

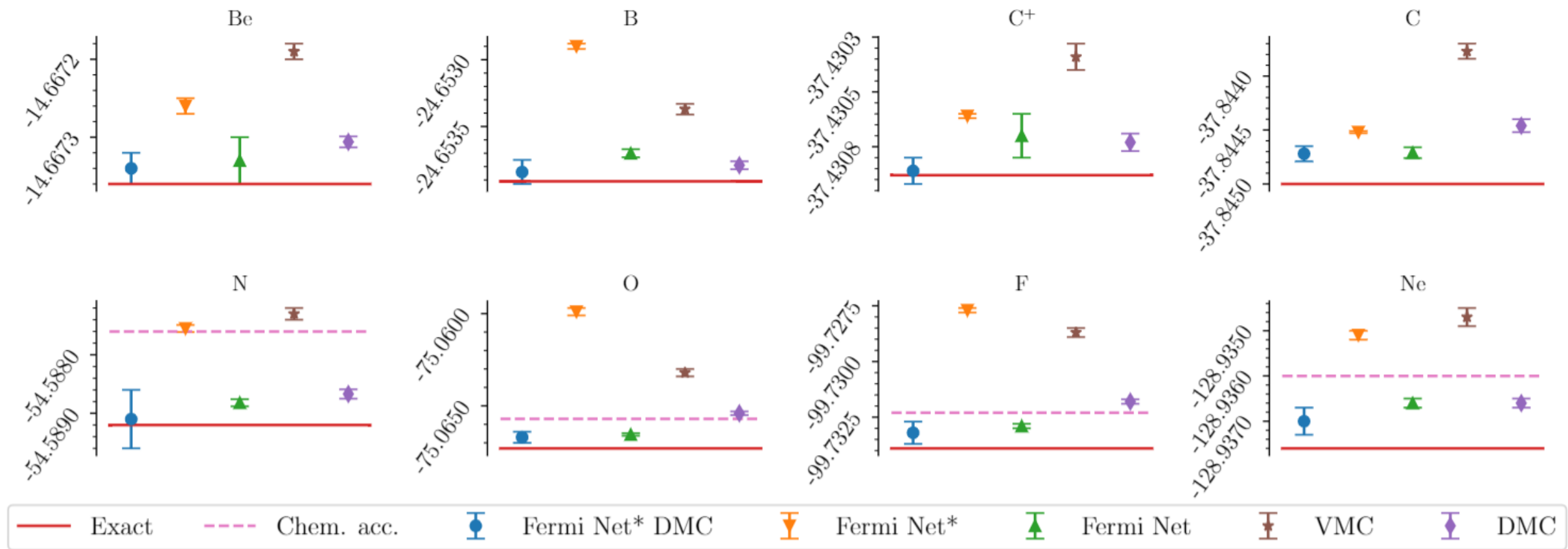


Approximate stochastic reconfiguration

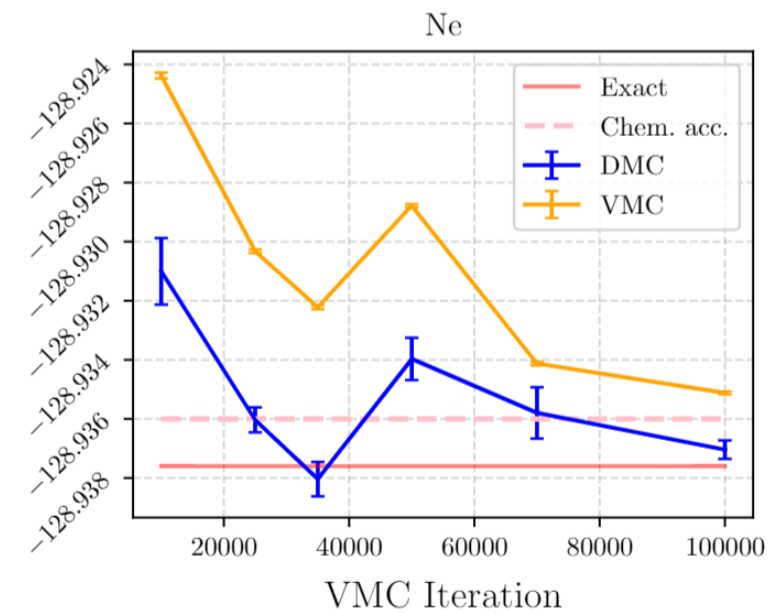
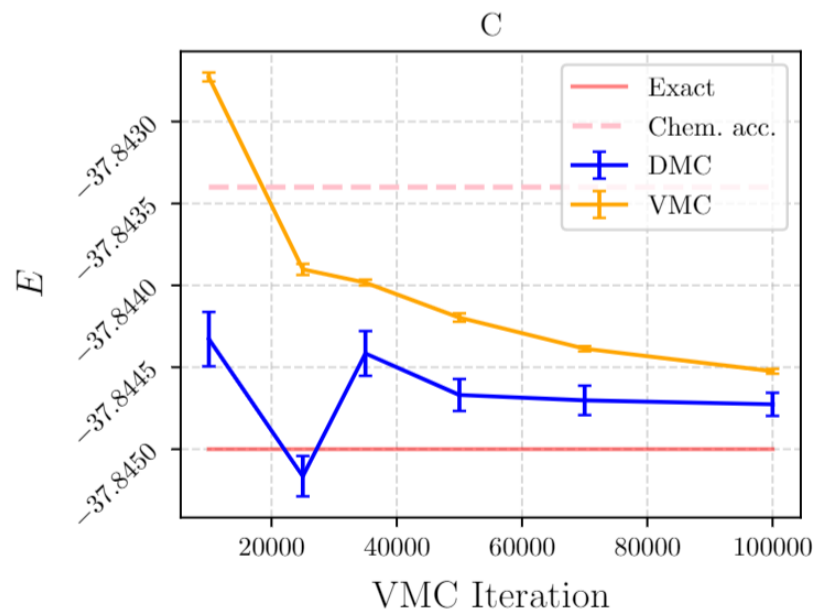
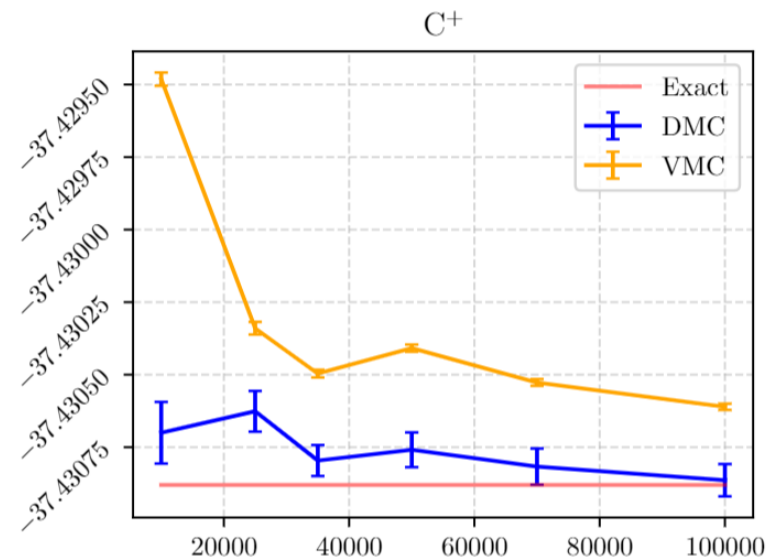
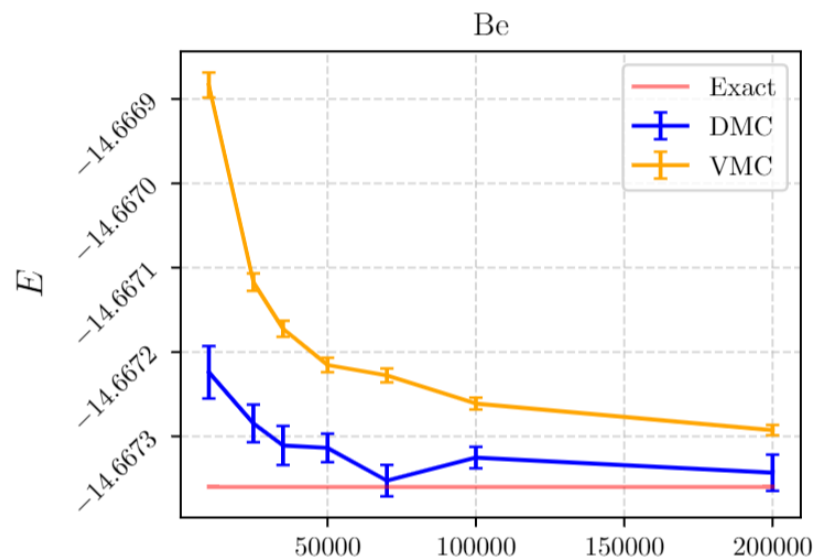
Some results

Atom	Fermi Net* + DMC (This work)	Fermi Net* (This work)	Fermi Net [7]	VMC [31]	DMC [31]	HF	Exact [32]
Be	-14.66734(2)	-14.66726(1)	-14.66733(3)	-14.66719(1)	-14.667306(7)	-14.35188	-14.66736
B	-24.65384(9)	-24.65290(2)	-24.65370(3)	-24.65337(4)	-24.65379(3)	-24.14899	-24.65391
C ⁺	-37.43086(6)	-37.43061(1)	-37.4307(1) [†]	-37.43034(6)	-37.43073(4)	-36.87037	-37.43088
C	-37.84472(7)	-37.84452(1)	-37.84471(5)	-37.84377(7)	-37.84446(6)	-37.08959	-37.8450
N	-54.5891(5)	-54.58755(6)	-54.58882(6)	-54.5873(1)	-54.58867(8)	-53.5545	-54.5892
O	-75.0667(3)	-75.0599(2)	-75.06655(7)	-75.0632(2)	-75.0654(1)	-73.6618	-75.0673
F	-99.7332(5)	-99.7277(1)	-99.7329(1)	-99.7287(2)	-99.7318(1)	-97.9865	-99.7339
Ne	-128.9370(3)	-128.9351(1)	-128.9366(1)	-128.9347(2)	-128.9366(1)	-126.6045	-128.9376

Some results



Some results



Next steps

Improved Ansatz design

- Can other changes be made to Fermi Net?
- Other wave functions?

Improved VMC

- CG SR?
- KFAC features?
- Mixed precision training?

Other systems

- Solids?
- Where else is VMC used?

Other VMC methods

- Pseudopotentials?

Scaling

- Most efficient way to model large systems?
- What are the technical challenges of scaling?

Final Comments

Long-term vision

- Infrastructure (GPUs and deep learning frameworks) will continue to improve
- Quantum chemistry neural network methods will adopt other standard quantum chemistry methods
- Neural network methods will supplement or replace standard quantum chemistry methods

Takeaways

- We improved the Fermi Net by changing how data was handled and extended with diffusion Monte Carlo
- There is a long-term case for investing in neural network method for modelling wave functions