Title: Neural-network quantum states for ultra-cold Fermi gases

Speakers: Jane Kim

Series: Machine Learning Initiative

Date: March 15, 2024 - 2:30 PM

URL: https://pirsa.org/24030116

Abstract: Ultra-cold Fermi gases exhibit a rich array of quantum mechanical properties, including the transition from a fermionic superfluid Bardeen-Cooper-Schrieffer (BCS) state to a bosonic superfluid Bose-Einstein condensate (BEC), which can be precisely probed experimentally. However, accurately describing these properties poses significant theoretical challenges due to strong pairing correlations and non-perturbative interactions. In this talk, I will discuss our recent development--a Pfaffian-Jastrow neural-network quantum state equipped with a message-passing architecture, designed to efficiently capture pairing and backflow correlations. We benchmark our approach against existing Slater-Jastrow frameworks and state-of-the-art diffusion Monte Carlo methods. Analysis of pair distribution functions and pairing gaps reveals the emergence of strong pairing correlations around unitarity. We demonstrate that transfer learning stabilizes the training process in the presence of strong, short-ranged interactions, allowing for an effective exploration of the BCS-BEC crossover region. Our findings highlight the potential of neural-network quantum states as a promising strategy for investigating ultra-cold Fermi gases.

Zoom link

NEURAL-NETWORK QUANTUM STATES FOR ULTRA-COLD FERMI GASES

JANE KIM

Ohio University (Previously MSU/FRIB)

Perimeter Institute for Theoretical Physics Seminar

15 March 2024

arXiv:2305:08831
accepted Comm. Phys.

In collaboration with:

M. Hjorth-Jensen, A. Lovato, B. Fore, G. Carleo, G. Pescia, J. Nys, and S. Gandolfi

MSU/FRIB/UiO ANL EPFL LANL

ULTRA-COLD FERMI GASES

Dilute → mostly s-wave

- Characterized by strong, short-range attraction between fermions
- Can be created in the lab by tuning an external magnetic field near a Feshbach resonance
    - $a < 0$ BCS regime of long-range Cooper pairs
    - $a > 0$ BEC regime of tightly-bound dimers
    - $|a| \to \infty$ Unitary limit (universal)
- The unitary Fermi gas is an extreme example of a strongly-interacting quantum system

ULTRA-COLD FERMI GASES

$T = 0$

Two components
(Spin-up & spin-down, heavy & light, etc.)

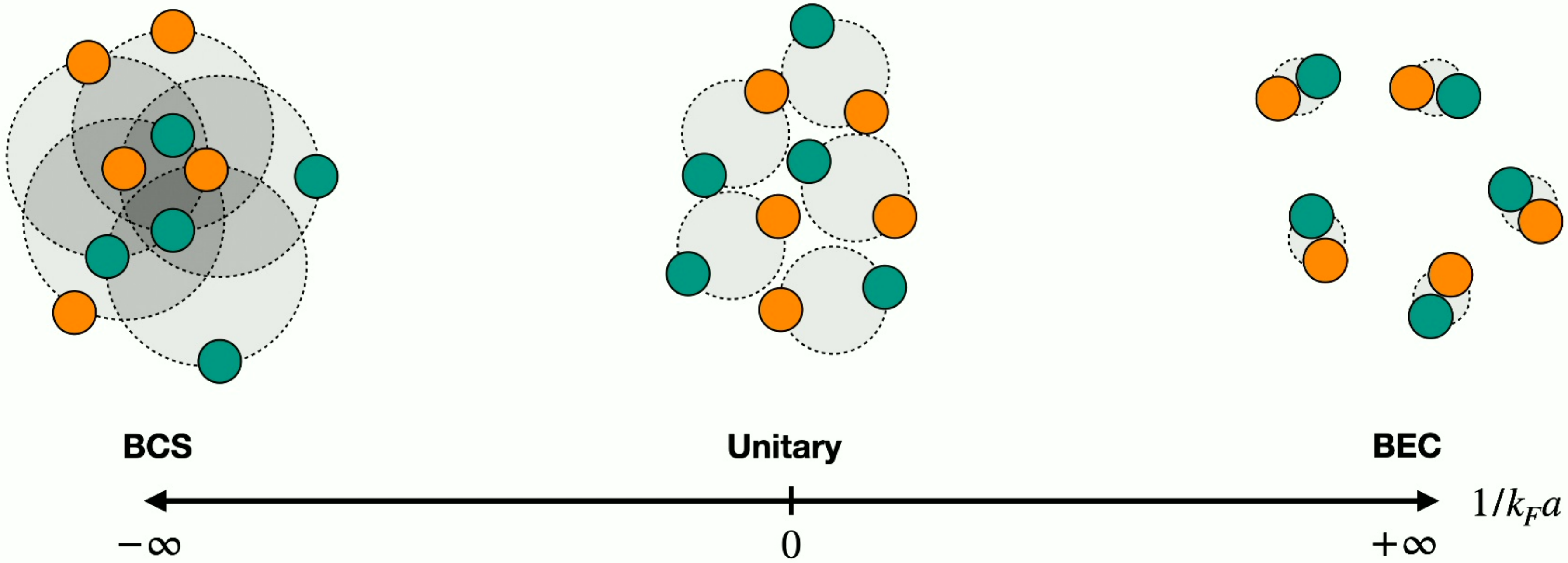
Dilute → mostly s-wave

- Characterized by strong, short-range attraction between fermions
- Can be created in the lab by tuning an external magnetic field near a Feshbach resonance

- $a < 0$	BCS regime of long-range Cooper pairs
- $a > 0$	BEC regime of tightly-bound dimers
- $\lvert a \rvert \to \infty$	Unitary limit (universal)

- The unitary Fermi gas is an extreme example of a strongly-interacting quantum system

THE BCS-BEC CROSSOVER



BCS

Unitary

BEC

$1/k_F a$

$-\infty$

0

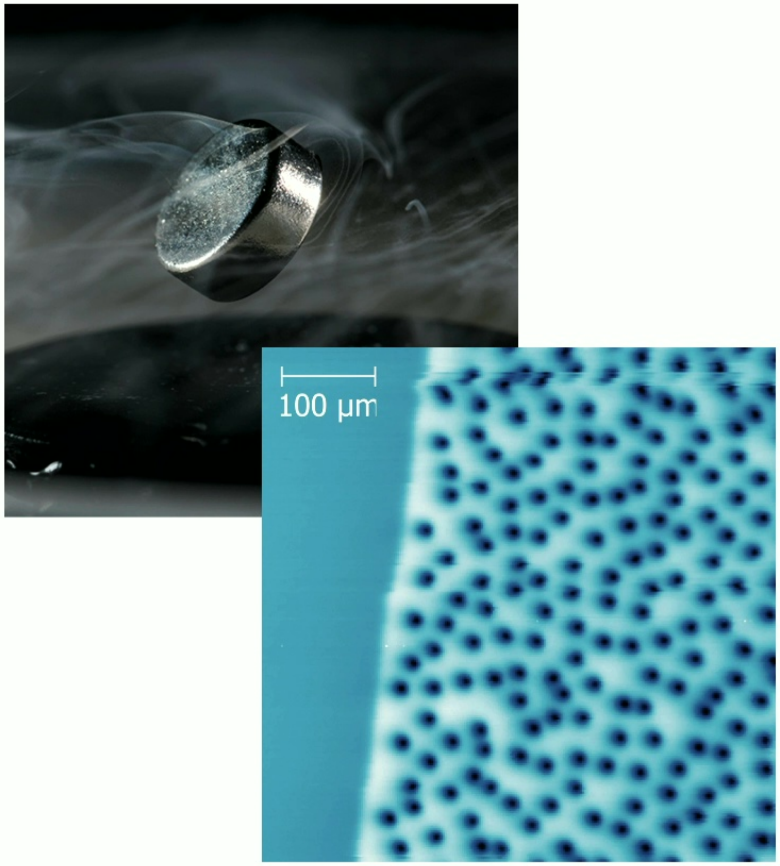
$+\infty$

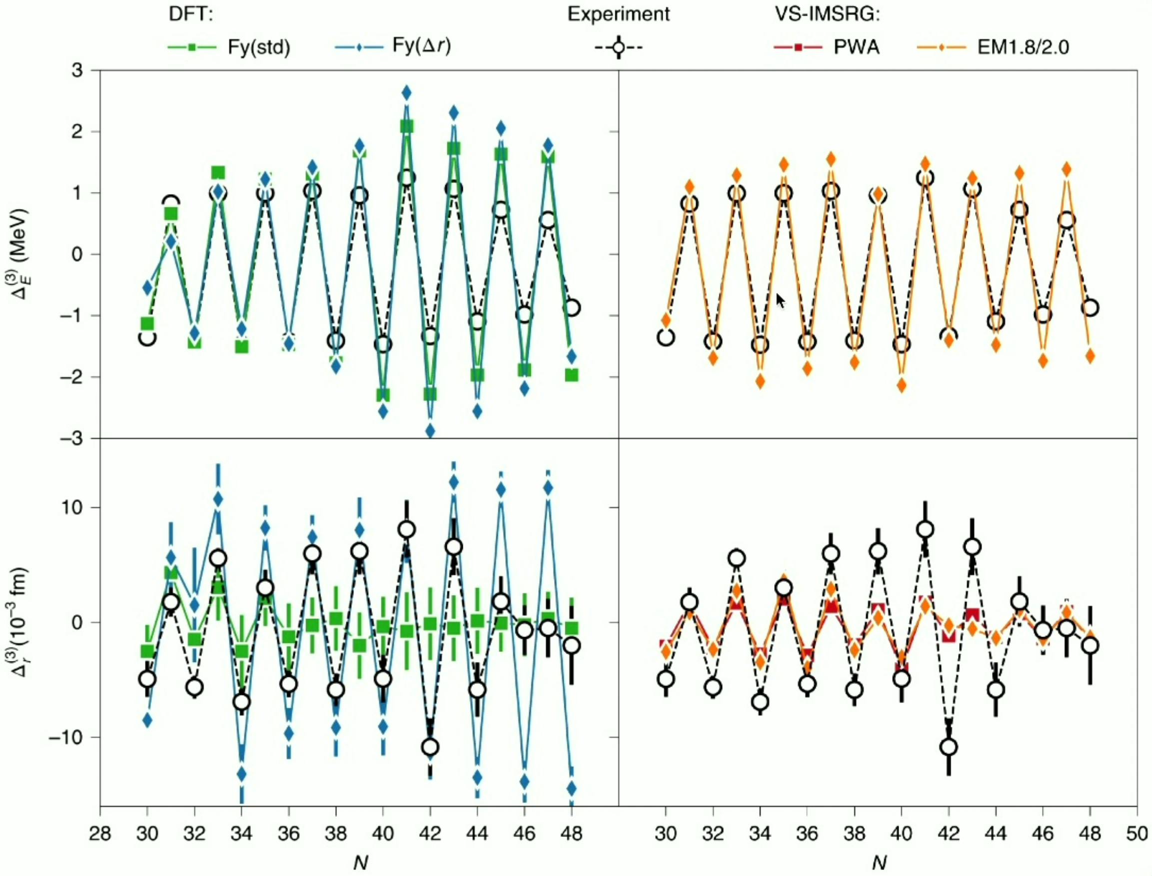
Weakly attractive superfluid

Strongly interacting superfluid

Weakly repulsive superfluid

PAIRING PHENOMENA





J. Adam Fenster; Wells et al. Scientific Reports 5, 8677 (2015); de Groote et al., Nature Physics **16**, 620-624 (2020)

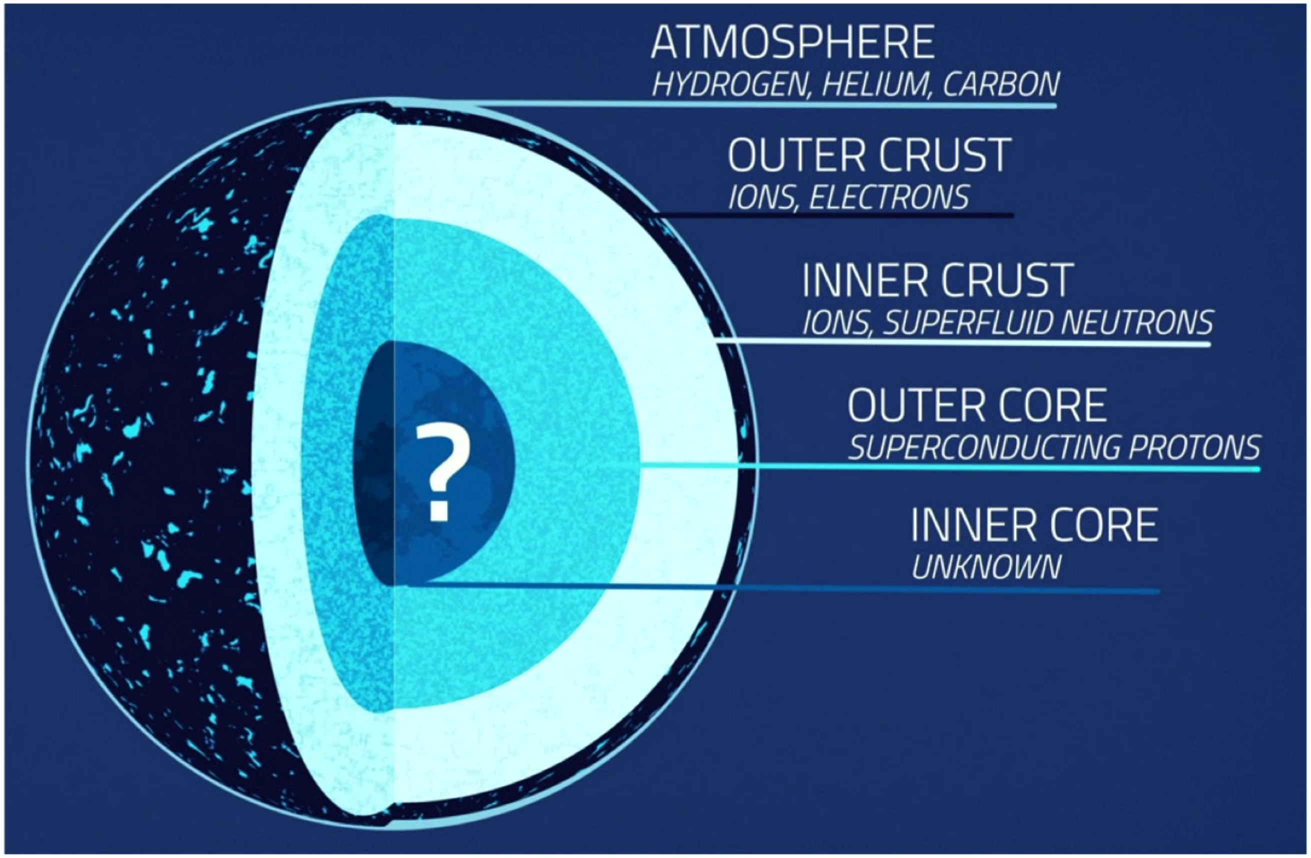
NEUTRON STAR CRUSTS

- Neutron-neutron scattering length:

$$a_{nn} = -\,18.63(84)\ \mathrm{fm}$$

- Dilute neutron matter is similar to unitary Fermi gas

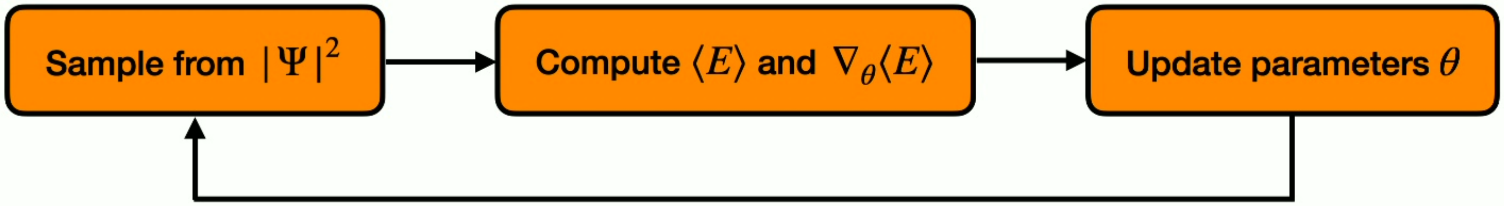
$$r_e \ll n^{1/3}$$



Chen et al., *Phys. Rev. C* **77**, 054002 (2008); NASA's Goddard Space Flight Center / Conceptual Image Lab

OUR APPROACH

- Simulate unpolarized gas of N fermions in a periodic box of side length L
- Design a neural-network quantum state (NQS) that efficiently encodes pairing and backflow correlations, while enforcing symmetries and boundary conditions
- Train NQS using variational Monte Carlo (VMC) method with stochastic reconfiguration



- Keep NQS applicable for nuclear Hamiltonians…

$$\hat{O} = 1,\ \sigma_i \cdot \sigma_j,\ \tau_i \cdot \tau_j,\ (\sigma_i \cdot \sigma_j)(\tau_i \cdot \tau_j),\ S_{ij},\ S_{ij}(\tau_i \cdot \tau_j),\ \mathbf{L} \cdot \mathbf{S},\ (\mathbf{L} \cdot \mathbf{S})(\tau_i \cdot \tau_j),\ \mathbf{L}^2,\ \mathbf{L}^2(\tau_i \cdot \tau_j), \ldots$$

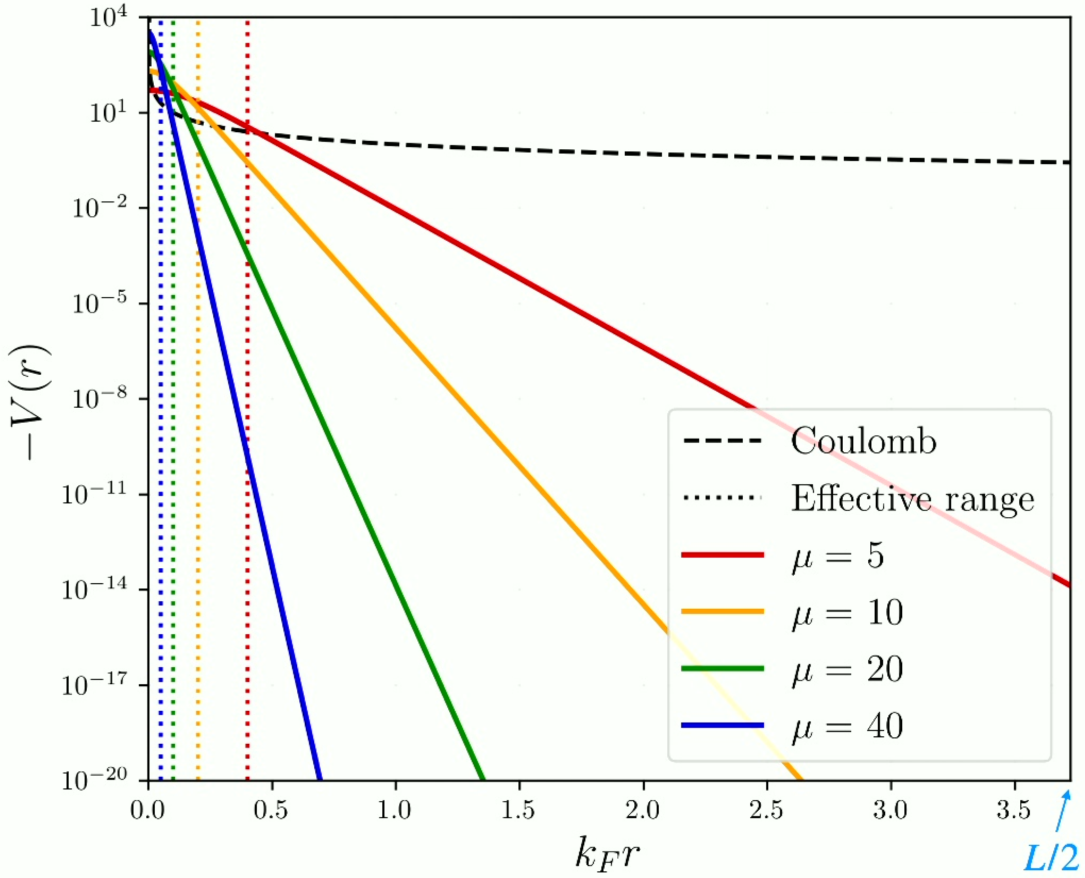
Ordóñez, Ray, and van Kolck, Phys. Rev. C **53**, 2086 (1996)

PÖSCHL-TELLER POTENTIAL

- Regularized, short-range attraction

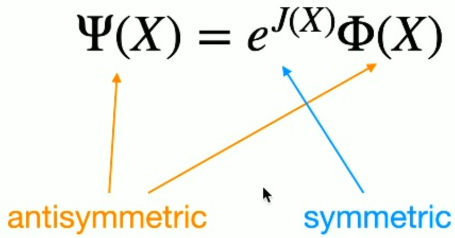
$$V_{ij} = (\delta_{s_i,s_j} - 1)v_0 \frac{2\hbar^2}{m} \frac{\mu^2}{\cosh^2(\mu r_{ij})}$$

- Exact solutions of two-body problem
- At unitarity: $v_0 = 1,\ r_e = 2/\mu$



FERMIONIC WAVE FUNCTIONS

- Antisymmetric w.r.t. particle exchange:



$$\Psi(X) = e^{J(X)}\Phi(X)$$

antisymmetric

symmetric

$$X \equiv \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N\}$$

$$\mathbf{x}_i = (\mathbf{r}_i, s_i)$$

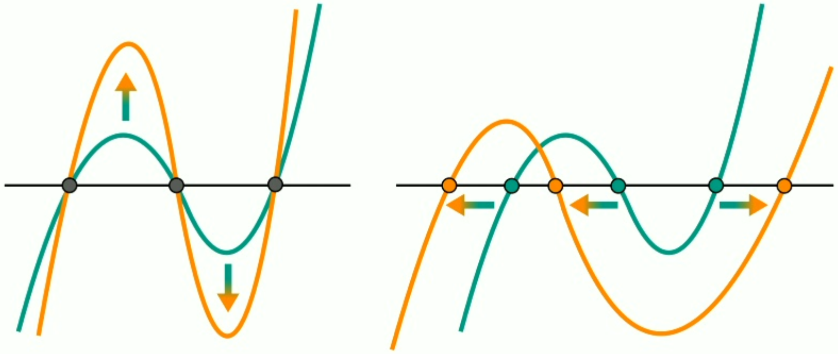
- Slater determinant:

$$\Phi(X) = \det \begin{bmatrix} \phi_1(\mathbf{x}_1) & \phi_1(\mathbf{x}_2) & \cdots & \phi_1(\mathbf{x}_N) \\ \phi_2(\mathbf{x}_1) & \phi_2(\mathbf{x}_2) & \cdots & \phi_2(\mathbf{x}_N) \\ \vdots & \vdots & \ddots & \vdots \\ \phi_N(\mathbf{x}_1) & \phi_N(\mathbf{x}_2) & \cdots & \phi_N(\mathbf{x}_N) \end{bmatrix}$$

- Fixed-node approximation with plane wave orbitals (SJ-PW): $\phi_\alpha(\mathbf{x}_i) = e^{i\mathbf{k}_\alpha \cdot \mathbf{r}_i}\delta_{s_\alpha, s_i}$

BACKFLOW TRANSFORMATIONS

- Symmetric, positive-definite Jastrow factor cannot change the nodes of the wave function



- Transform to a new basis that incorporates the influences from surrounding particles (must be permutation equivariant)

$$\mathbf{r}_i \mapsto \mathbf{r}_i + \mathbf{u}_i(X) + i\mathbf{v}_i(X)$$

$$|s_i\rangle \mapsto |\chi_i\rangle = \cos\left(\frac{\theta_i(X)}{2}\right)|s_i\rangle + \sin\left(\frac{\theta_i(X)}{2}\right)\sigma_i^x|s_i\rangle$$

BCS WAVE FUNCTION

- Used for unpolarized systems with strong singlet pairing correlations
- Relies on separating spin-up and spin-down particles (cannot be used for nuclear systems)

$$\Phi(X) = \det \begin{bmatrix} \phi(\mathbf{x}_{1\uparrow}, \mathbf{x}_{1\downarrow}) & \phi(\mathbf{x}_{1\uparrow}, \mathbf{x}_{2\downarrow}) & \cdots & \phi(\mathbf{x}_{1\uparrow}, \mathbf{x}_{N/2\downarrow}) \\ \phi(\mathbf{x}_{2\uparrow}, \mathbf{x}_{1\downarrow}) & \phi(\mathbf{x}_{2\uparrow}, \mathbf{x}_{2\downarrow}) & \cdots & \phi(\mathbf{x}_{2\uparrow}, \mathbf{x}_{N/2\downarrow}) \\ \vdots & \vdots & \ddots & \vdots \\ \phi(\mathbf{x}_{N/2\uparrow}, \mathbf{x}_{1\downarrow}) & \phi(\mathbf{x}_{N/2\uparrow}, \mathbf{x}_{2\downarrow}) & \cdots & \phi(\mathbf{x}_{N/2\uparrow}, \mathbf{x}_{N/2\downarrow}) \end{bmatrix}$$

- Can expand matrix to include unpaired orbitals for spin-polarized systems

PFAFFIAN WAVE FUNCTION

- Simplest and most general way to build an antisymmetrized product of pairing orbitals rather than single-particle orbitals

$$\Phi(X) = \mathrm{pf}\begin{bmatrix} 0 & \phi(\mathbf{x}_1, \mathbf{x}_2) & \cdots & \phi(\mathbf{x}_1, \mathbf{x}_N) \\ \phi(\mathbf{x}_2, \mathbf{x}_1) & 0 & \cdots & \phi(\mathbf{x}_2, \mathbf{x}_N) \\ \vdots & \vdots & \ddots & \vdots \\ \phi(\mathbf{x}_N, \mathbf{x}_1) & \phi(\mathbf{x}_N, \mathbf{x}_2) & \cdots & 0 \end{bmatrix}$$

PFAFFIAN WAVE FUNCTION

- Simplest and most general way to build an antisymmetrized product of pairing orbitals rather than single-particle orbitals

$$\Phi(X) = \mathrm{pf} \begin{bmatrix} 0 & \phi(\mathbf{x}_1, \mathbf{x}_2) & \cdots & \phi(\mathbf{x}_1, \mathbf{x}_N) \\ \phi(\mathbf{x}_2, \mathbf{x}_1) & 0 & \cdots & \phi(\mathbf{x}_2, \mathbf{x}_N) \\ \vdots & \vdots & \ddots & \vdots \\ \phi(\mathbf{x}_N, \mathbf{x}_1) & \phi(\mathbf{x}_N, \mathbf{x}_2) & \cdots & 0 \end{bmatrix}$$

What is the “Pfaffian”?

DETERMINANT VS. PFAFFIAN

Defined for $n \times n$ matrices

$$\det(A) = \sum_{\sigma \in S_n} \mathrm{sgn}(\sigma) \prod_{i=1}^{n} a_{i,\sigma(i)}$$

$$\det(A^T) = \det(A)$$

$$\det(A)\det(B) = \det(AB)$$

Defined for $2n \times 2n$ skew-symmetric matrices

$$\mathrm{pf}(A) = \frac{1}{2^n n!} \sum_{\sigma \in S_{2n}} \mathrm{sgn}(\sigma) \prod_{i=1}^{n} a_{\sigma(2i-1),\sigma(2i)}$$

$$\mathrm{pf}(A^T) = (-1)^n \mathrm{pf}(A)$$

$$\mathrm{pf}(A)\mathrm{pf}(B) = \exp\left(\frac{1}{2}\mathrm{tr}\log(A^T B)\right)$$

$$\det(A) = \mathrm{pf}(A)^2$$

Pfaffian implementation: Wimmer, ACM Trans. Math. Softw. **38**, 4 (2012)

PFAFFIAN WAVE FUNCTION

- Simplest and most general way to build an antisymmetrized product of pairing orbitals rather than single-particle orbitals

$$\Phi_{PJ}(X) = \mathrm{pf}\begin{bmatrix} 0 & \phi(\mathbf{x}_1, \mathbf{x}_2) & \cdots & \phi(\mathbf{x}_1, \mathbf{x}_N) \\ \phi(\mathbf{x}_2, \mathbf{x}_1) & 0 & \cdots & \phi(\mathbf{x}_2, \mathbf{x}_N) \\ \vdots & \vdots & \ddots & \vdots \\ \phi(\mathbf{x}_N, \mathbf{x}_1) & \phi(\mathbf{x}_N, \mathbf{x}_2) & \cdots & 0 \end{bmatrix}$$

- Pairing orbital into commonly decomposed into explicit singlet and triplet contributions
- We take advantage of universal approximation theorem: $\phi(\mathbf{x}_i, \mathbf{x}_j) \equiv \nu(\mathbf{x}_i, \mathbf{x}_j) - \nu(\mathbf{x}_j, \mathbf{x}_i)$, where ν is a neural network.
- Naturally encodes singlet and triplet pairing because ν takes spins as input

STU Pfaffian: Bajdich et al., Phys. Rev. Lett. **96**, 130201 (2006)

MESSAGE-PASSING NEURAL NETWORK

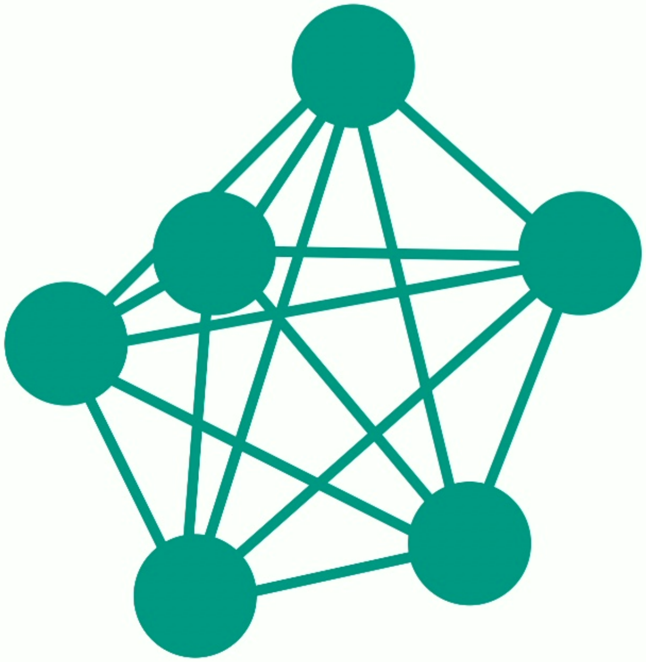
- Permutation-equivariant graph neural network
- Represent quantum system by a fully-connected graph
- Iteratively build backflow correlations into new one- and two-body features
- Skip connections help avoid vanishing gradient problem
- Visible nodes/one-body features: $\mathbf{v}_i = (s_i)$
- Visible edges/two-body features: $\mathbf{v}_{ij} = (r_{ij}, \mathbf{r}_{ij}, s_i \cdot s_j)$
- Preprocessing step: $\mathbf{h}_i^{(0)} = (\mathbf{v}_i, A\mathbf{v}_i)$

$$\mathbf{h}_{ij}^{(0)} = (\mathbf{v}_{ij}, B\mathbf{v}_{ij})$$

The set-up

Homogeneous electron gas: Pescia et al., arXiv:2305.07240 (2023)

MESSAGE-PASSING NEURAL NETWORK



for $t = 1, \ldots, T$:

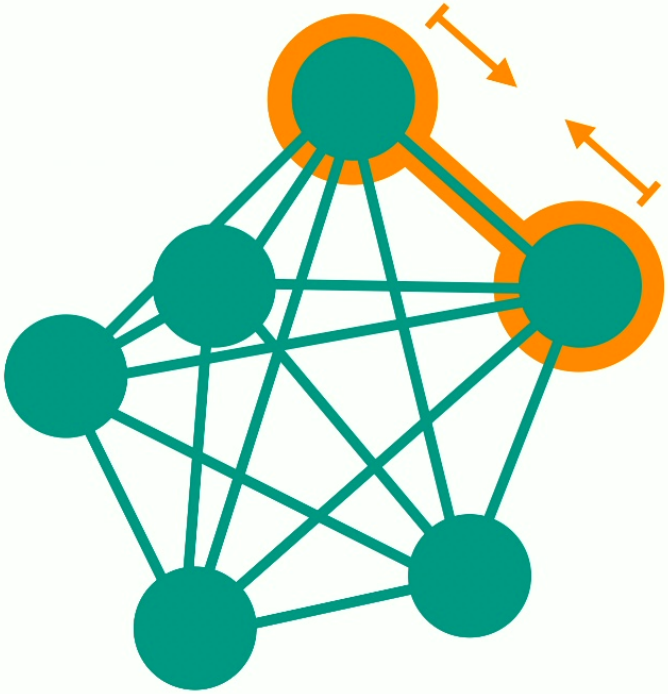
$$\mathbf{m}_{ij}^{(t)} = \mathbf{M}_t\left(\mathbf{h}_i^{(t-1)}, \mathbf{h}_j^{(t-1)}, \mathbf{h}_{ij}^{(t-1)}\right)$$

$$\mathbf{h}_{ij}^{(t)} = \left(\mathbf{v}_{ij}, \mathbf{G}_t\left(\mathbf{h}_{ij}^{(t-1)}, \mathbf{m}_{ij}^{(t)}\right)\right)$$

$$\mathbf{m}_i^{(t)} = \mathrm{Pool}\left(\{\mathbf{m}_{ij}^{(t)} \,|\, j \neq i\}\right)$$

$$\mathbf{h}_i^{(t)} = \left(\mathbf{v}_i, \mathbf{F}_t\left(\mathbf{h}_i^{(t-1)}, \mathbf{m}_i^{(t)}\right)\right)$$

MESSAGE-PASSING NEURAL NETWORK



for $t = 1, \ldots, T$:

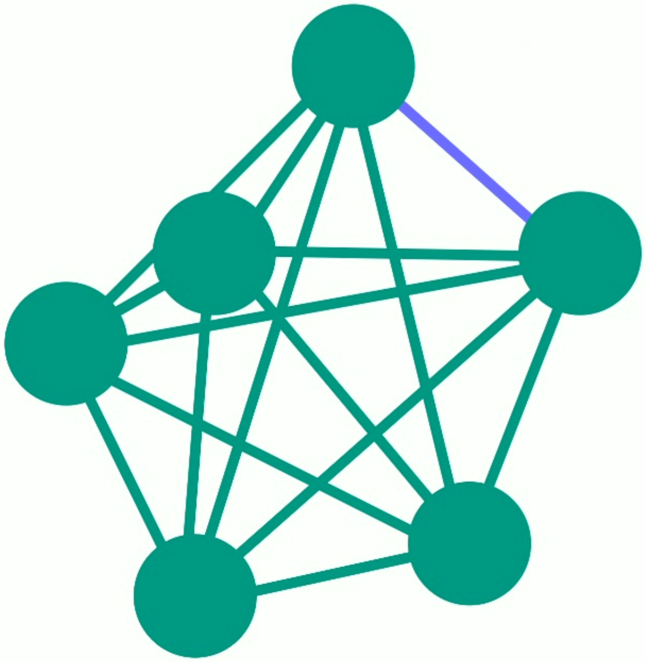
$$\mathbf{m}_{ij}^{(t)} = \mathbf{M}_t\left(\mathbf{h}_i^{(t-1)}, \mathbf{h}_j^{(t-1)}, \mathbf{h}_{ij}^{(t-1)}\right)$$

$$\mathbf{h}_{ij}^{(t)} = \left(\mathbf{v}_{ij}, \mathbf{G}_t\left(\mathbf{h}_{ij}^{(t-1)}, \mathbf{m}_{ij}^{(t)}\right)\right)$$

$$\mathbf{m}_i^{(t)} = \mathrm{Pool}\left(\{\mathbf{m}_{ij}^{(t)} \mid j \neq i\}\right)$$

$$\mathbf{h}_i^{(t)} = \left(\mathbf{v}_i, \mathbf{F}_t\left(\mathbf{h}_i^{(t-1)}, \mathbf{m}_i^{(t)}\right)\right)$$

MESSAGE-PASSING NEURAL NETWORK



for $t = 1, \ldots, T$:

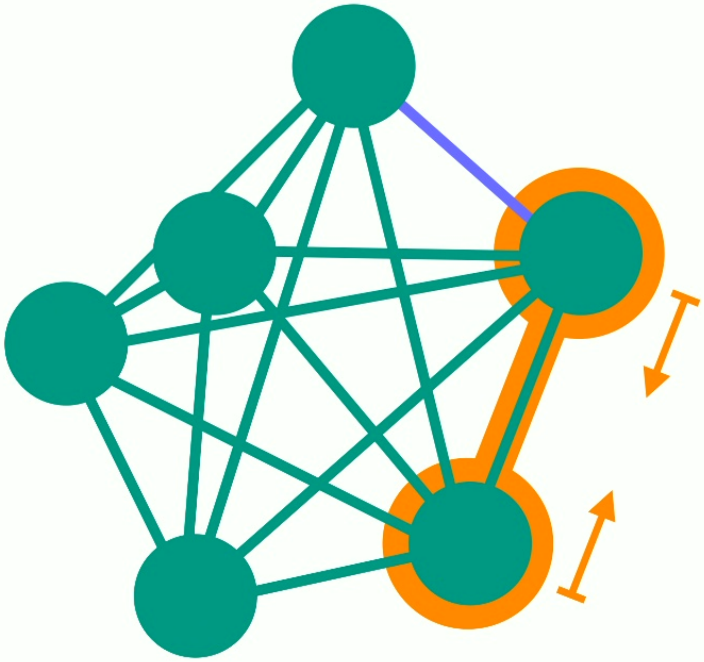
$$\mathbf{m}_{ij}^{(t)} = \mathbf{M}_t\left(\mathbf{h}_i^{(t-1)}, \mathbf{h}_j^{(t-1)}, \mathbf{h}_{ij}^{(t-1)}\right)$$

$$\mathbf{h}_{ij}^{(t)} = \left(\mathbf{v}_{ij}, \mathbf{G}_t\left(\mathbf{h}_{ij}^{(t-1)}, \mathbf{m}_{ij}^{(t)}\right)\right)$$

$$\mathbf{m}_i^{(t)} = \mathrm{Pool}\left(\{\mathbf{m}_{ij}^{(t)} \,|\, j \neq i\}\right)$$

$$\mathbf{h}_i^{(t)} = \left(\mathbf{v}_i, \mathbf{F}_t\left(\mathbf{h}_i^{(t-1)}, \mathbf{m}_i^{(t)}\right)\right)$$

MESSAGE-PASSING NEURAL NETWORK



for $t = 1, \ldots, T$:

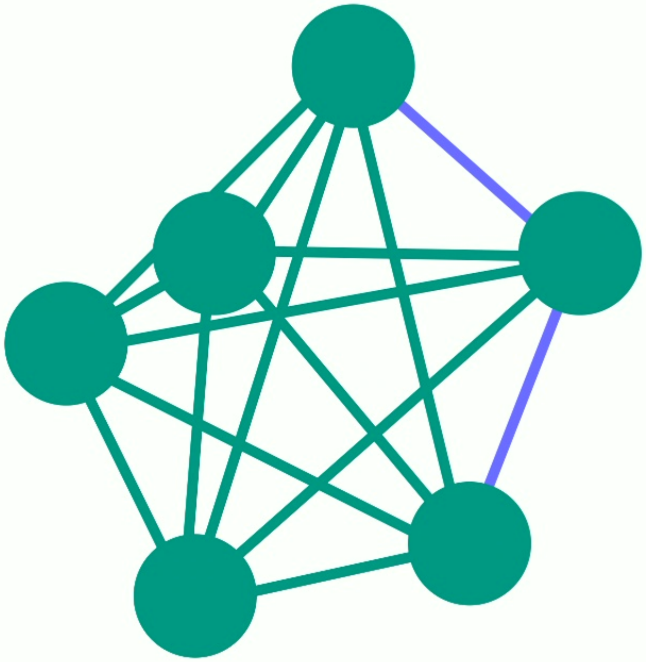
$$\mathbf{m}_{ij}^{(t)} = \mathbf{M}_t\left(\mathbf{h}_i^{(t-1)}, \mathbf{h}_j^{(t-1)}, \mathbf{h}_{ij}^{(t-1)}\right)$$

$$\mathbf{h}_{ij}^{(t)} = \left(\mathbf{v}_{ij}, \mathbf{G}_t\left(\mathbf{h}_{ij}^{(t-1)}, \mathbf{m}_{ij}^{(t)}\right)\right)$$

$$\mathbf{m}_i^{(t)} = \mathrm{Pool}\left(\{\mathbf{m}_{ij}^{(t)} \,|\, j \neq i\}\right)$$

$$\mathbf{h}_i^{(t)} = \left(\mathbf{v}_i, \mathbf{F}_t\left(\mathbf{h}_i^{(t-1)}, \mathbf{m}_i^{(t)}\right)\right)$$

MESSAGE-PASSING NEURAL NETWORK



for $t = 1, \ldots, T$:

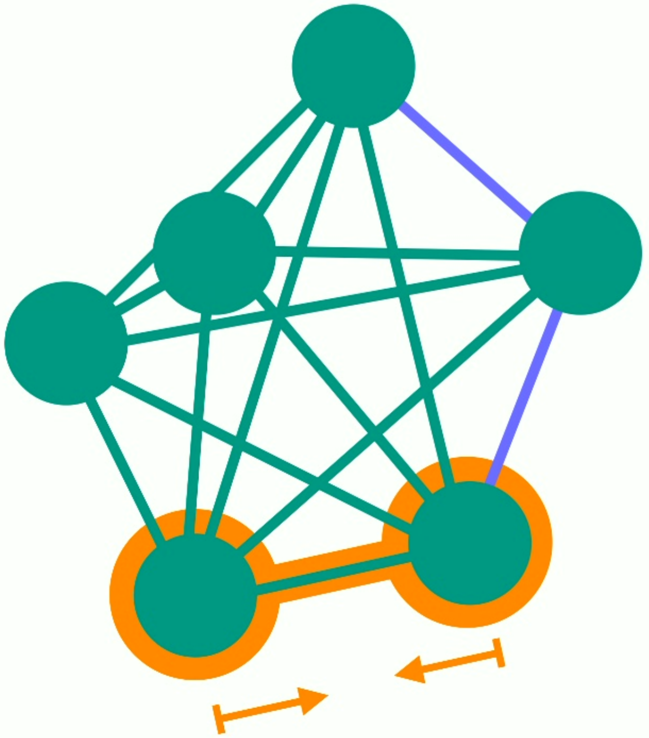
$$\mathbf{m}_{ij}^{(t)} = \mathbf{M}_t\left(\mathbf{h}_i^{(t-1)}, \mathbf{h}_j^{(t-1)}, \mathbf{h}_{ij}^{(t-1)}\right)$$

$$\mathbf{h}_{ij}^{(t)} = \left(\mathbf{v}_{ij}, \mathbf{G}_t\left(\mathbf{h}_{ij}^{(t-1)}, \mathbf{m}_{ij}^{(t)}\right)\right)$$

$$\mathbf{m}_i^{(t)} = \mathrm{Pool}\left(\{\mathbf{m}_{ij}^{(t)} \,|\, j \neq i\}\right)$$

$$\mathbf{h}_i^{(t)} = \left(\mathbf{v}_i, \mathbf{F}_t\left(\mathbf{h}_i^{(t-1)}, \mathbf{m}_i^{(t)}\right)\right)$$

MESSAGE-PASSING NEURAL NETWORK



for $t = 1, \ldots, T$:

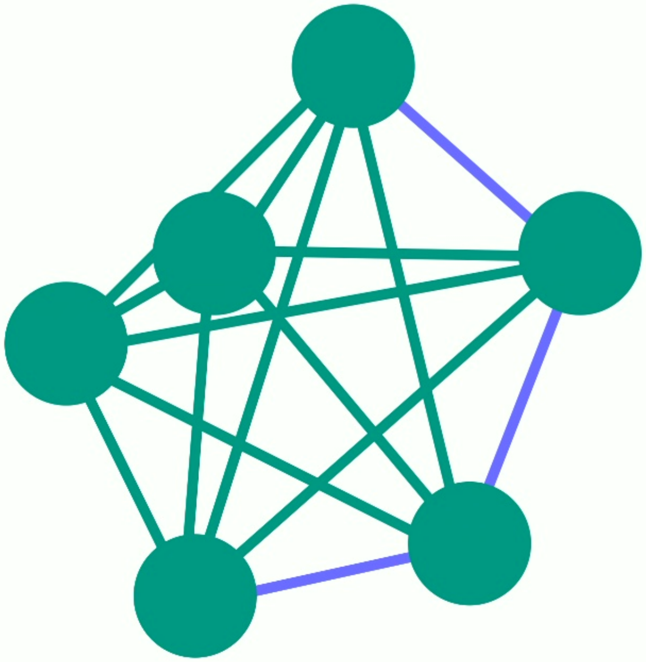
$$\mathbf{m}_{ij}^{(t)} = \mathbf{M}_t\left(\mathbf{h}_i^{(t-1)}, \mathbf{h}_j^{(t-1)}, \mathbf{h}_{ij}^{(t-1)}\right)$$

$$\mathbf{h}_{ij}^{(t)} = \left(\mathbf{v}_{ij}, \mathbf{G}_t\left(\mathbf{h}_{ij}^{(t-1)}, \mathbf{m}_{ij}^{(t)}\right)\right)$$

$$\mathbf{m}_i^{(t)} = \mathrm{Pool}\left(\{\mathbf{m}_{ij}^{(t)} \,|\, j \neq i\}\right)$$

$$\mathbf{h}_i^{(t)} = \left(\mathbf{v}_i, \mathbf{F}_t\left(\mathbf{h}_i^{(t-1)}, \mathbf{m}_i^{(t)}\right)\right)$$

MESSAGE-PASSING NEURAL NETWORK



for $t = 1, \ldots, T$:

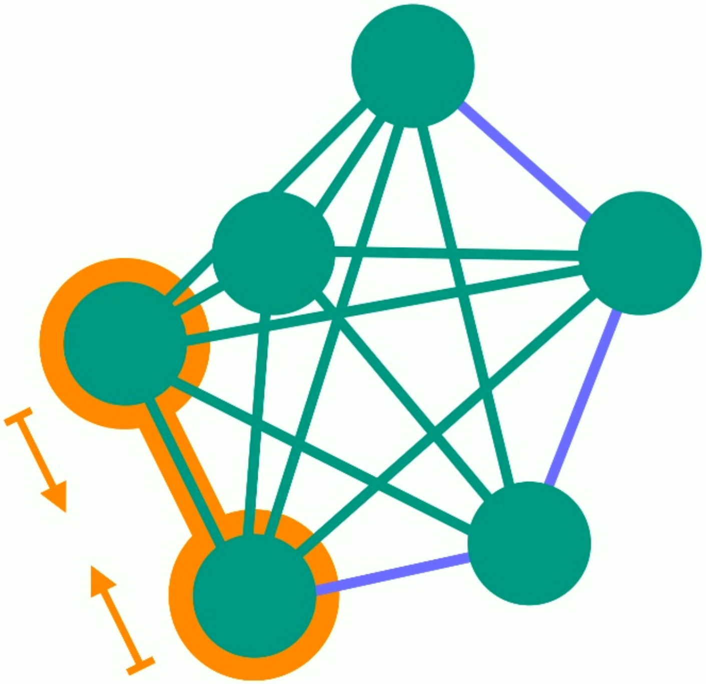
$$\mathbf{m}_{ij}^{(t)} = \mathbf{M}_t\left(\mathbf{h}_i^{(t-1)}, \mathbf{h}_j^{(t-1)}, \mathbf{h}_{ij}^{(t-1)}\right)$$

$$\mathbf{h}_{ij}^{(t)} = \left(\mathbf{v}_{ij}, \mathbf{G}_t\left(\mathbf{h}_{ij}^{(t-1)}, \mathbf{m}_{ij}^{(t)}\right)\right)$$

$$\mathbf{m}_i^{(t)} = \text{Pool}\left(\{\mathbf{m}_{ij}^{(t)} \,|\, j \neq i\}\right)$$

$$\mathbf{h}_i^{(t)} = \left(\mathbf{v}_i, \mathbf{F}_t\left(\mathbf{h}_i^{(t-1)}, \mathbf{m}_i^{(t)}\right)\right)$$

MESSAGE-PASSING NEURAL NETWORK



for $t = 1, \ldots, T$:

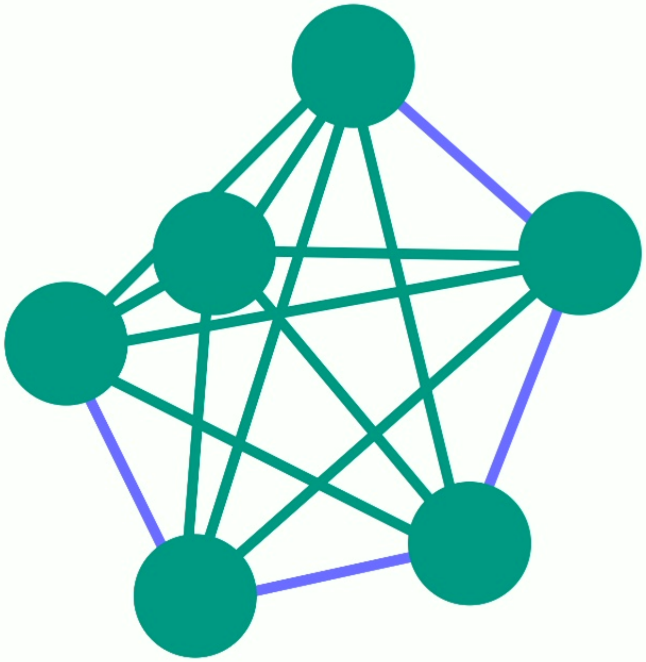
$$\mathbf{m}_{ij}^{(t)} = \mathbf{M}_t \left(\mathbf{h}_i^{(t-1)}, \mathbf{h}_j^{(t-1)}, \mathbf{h}_{ij}^{(t-1)} \right)$$

$$\mathbf{h}_{ij}^{(t)} = \left(\mathbf{v}_{ij}, \mathbf{G}_t \left(\mathbf{h}_{ij}^{(t-1)}, \mathbf{m}_{ij}^{(t)} \right) \right)$$

$$\mathbf{m}_i^{(t)} = \text{Pool} \left(\{ \mathbf{m}_{ij}^{(t)} \, | \, j \neq i \} \right)$$

$$\mathbf{h}_i^{(t)} = \left(\mathbf{v}_i, \mathbf{F}_t \left(\mathbf{h}_i^{(t-1)}, \mathbf{m}_i^{(t)} \right) \right)$$

MESSAGE-PASSING NEURAL NETWORK



for $t = 1, \ldots, T$:

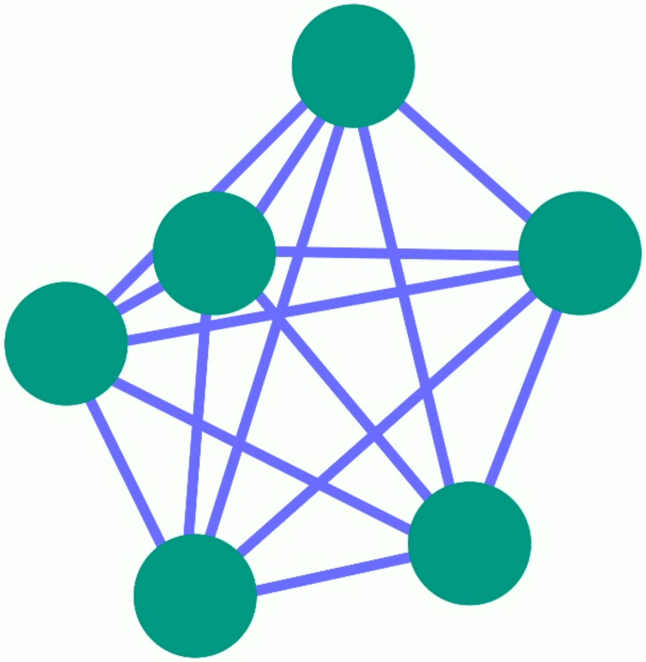
$$\mathbf{m}_{ij}^{(t)} = \mathbf{M}_t\left(\mathbf{h}_i^{(t-1)}, \mathbf{h}_j^{(t-1)}, \mathbf{h}_{ij}^{(t-1)}\right)$$

$$\mathbf{h}_{ij}^{(t)} = \left(\mathbf{v}_{ij}, \mathbf{G}_t\left(\mathbf{h}_{ij}^{(t-1)}, \mathbf{m}_{ij}^{(t)}\right)\right)$$

$$\mathbf{m}_i^{(t)} = \mathrm{Pool}\left(\{\mathbf{m}_{ij}^{(t)} \,|\, j \neq i\}\right)$$

$$\mathbf{h}_i^{(t)} = \left(\mathbf{v}_i, \mathbf{F}_t\left(\mathbf{h}_i^{(t-1)}, \mathbf{m}_i^{(t)}\right)\right)$$

MESSAGE-PASSING NEURAL NETWORK



for $t = 1, \ldots, T$:

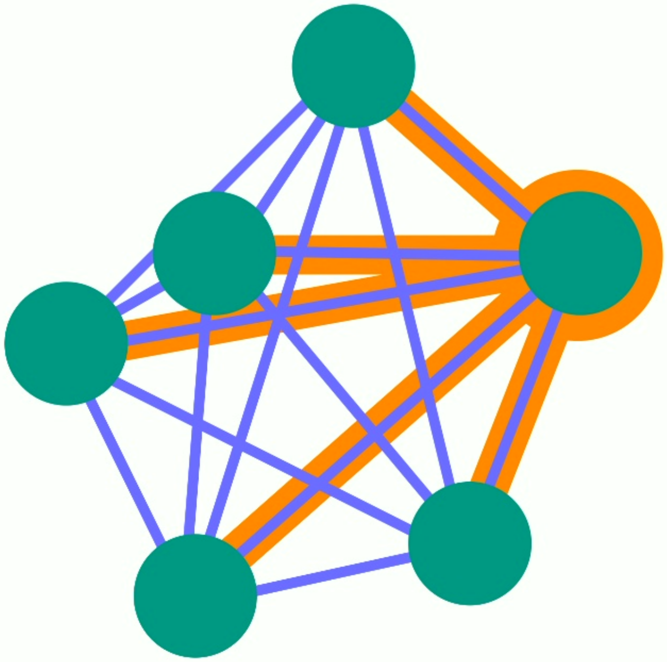
$$\mathbf{m}_{ij}^{(t)} = \mathbf{M}_t \left(\mathbf{h}_i^{(t-1)}, \mathbf{h}_j^{(t-1)}, \mathbf{h}_{ij}^{(t-1)} \right)$$

$$\mathbf{h}_{ij}^{(t)} = \left(\mathbf{v}_{ij}, \mathbf{G}_t \left(\mathbf{h}_{ij}^{(t-1)}, \mathbf{m}_{ij}^{(t)} \right) \right)$$

$$\mathbf{m}_i^{(t)} = \mathrm{Pool} \left(\{ \mathbf{m}_{ij}^{(t)} \,|\, j \neq i \} \right)$$

$$\mathbf{h}_i^{(t)} = \left(\mathbf{v}_i, \mathbf{F}_t \left(\mathbf{h}_i^{(t-1)}, \mathbf{m}_i^{(t)} \right) \right)$$

MESSAGE-PASSING NEURAL NETWORK



for $t = 1, \ldots, T$:

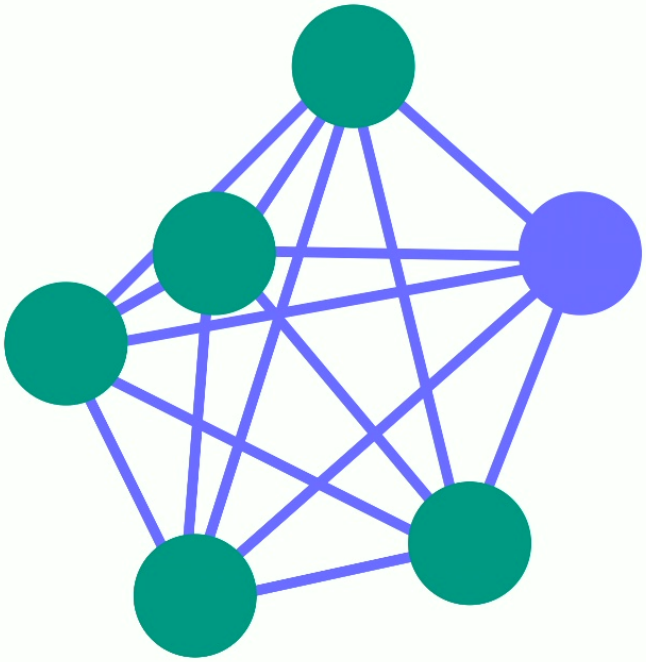
$$\mathbf{m}_{ij}^{(t)} = \mathbf{M}_t\left(\mathbf{h}_i^{(t-1)}, \mathbf{h}_j^{(t-1)}, \mathbf{h}_{ij}^{(t-1)}\right)$$

$$\mathbf{h}_{ij}^{(t)} = \left(\mathbf{v}_{ij}, \mathbf{G}_t\left(\mathbf{h}_{ij}^{(t-1)}, \mathbf{m}_{ij}^{(t)}\right)\right)$$

$$\mathbf{m}_i^{(t)} = \mathrm{Pool}\left(\{\mathbf{m}_{ij}^{(t)} \,|\, j \neq i\}\right)$$

$$\mathbf{h}_i^{(t)} = \left(\mathbf{v}_i, \mathbf{F}_t\left(\mathbf{h}_i^{(t-1)}, \mathbf{m}_i^{(t)}\right)\right)$$

MESSAGE-PASSING NEURAL NETWORK



for $t = 1, \ldots, T$:

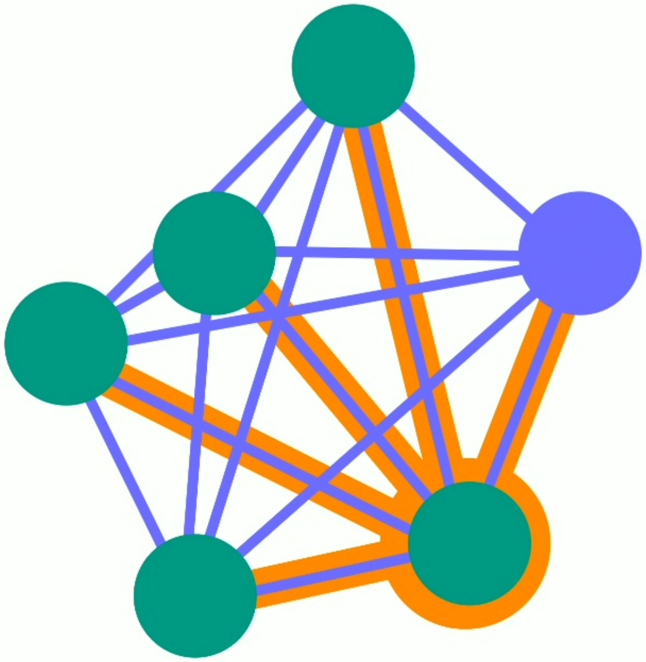
$$\mathbf{m}_{ij}^{(t)} = \mathbf{M}_t \left(\mathbf{h}_i^{(t-1)}, \mathbf{h}_j^{(t-1)}, \mathbf{h}_{ij}^{(t-1)} \right)$$

$$\mathbf{h}_{ij}^{(t)} = \left(\mathbf{v}_{ij}, \mathbf{G}_t \left(\mathbf{h}_{ij}^{(t-1)}, \mathbf{m}_{ij}^{(t)} \right) \right)$$

$$\mathbf{m}_i^{(t)} = \mathrm{Pool} \left(\{ \mathbf{m}_{ij}^{(t)} \,|\, j \neq i \} \right)$$

$$\mathbf{h}_i^{(t)} = \left(\mathbf{v}_i, \mathbf{F}_t \left(\mathbf{h}_i^{(t-1)}, \mathbf{m}_i^{(t)} \right) \right)$$

MESSAGE-PASSING NEURAL NETWORK



for $t = 1, \ldots, T$:

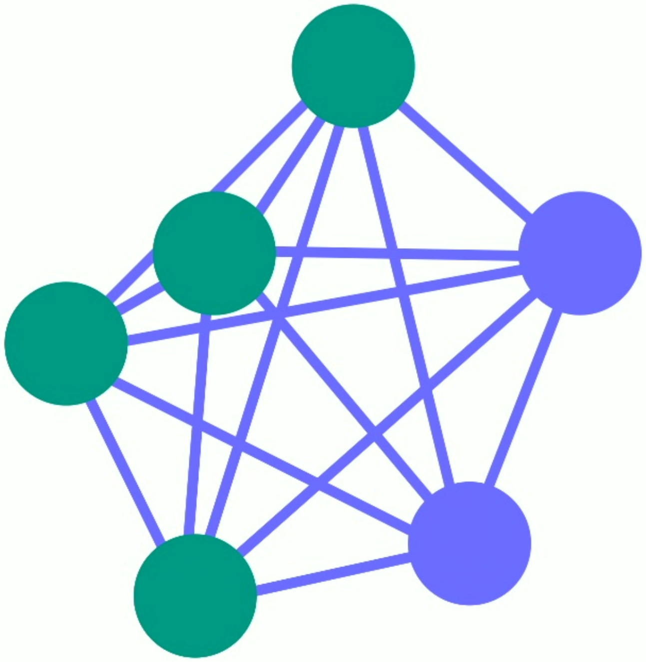
$$\mathbf{m}_{ij}^{(t)} = \mathbf{M}_t\left(\mathbf{h}_i^{(t-1)}, \mathbf{h}_j^{(t-1)}, \mathbf{h}_{ij}^{(t-1)}\right)$$

$$\mathbf{h}_{ij}^{(t)} = \left(\mathbf{v}_{ij}, \mathbf{G}_t\left(\mathbf{h}_{ij}^{(t-1)}, \mathbf{m}_{ij}^{(t)}\right)\right)$$

$$\mathbf{m}_i^{(t)} = \mathrm{Pool}\left(\{\mathbf{m}_{ij}^{(t)} \,|\, j \neq i\}\right)$$

$$\mathbf{h}_i^{(t)} = \left(\mathbf{v}_i, \mathbf{F}_t\left(\mathbf{h}_i^{(t-1)}, \mathbf{m}_i^{(t)}\right)\right)$$

MESSAGE-PASSING NEURAL NETWORK



for $t = 1, \ldots, T$:

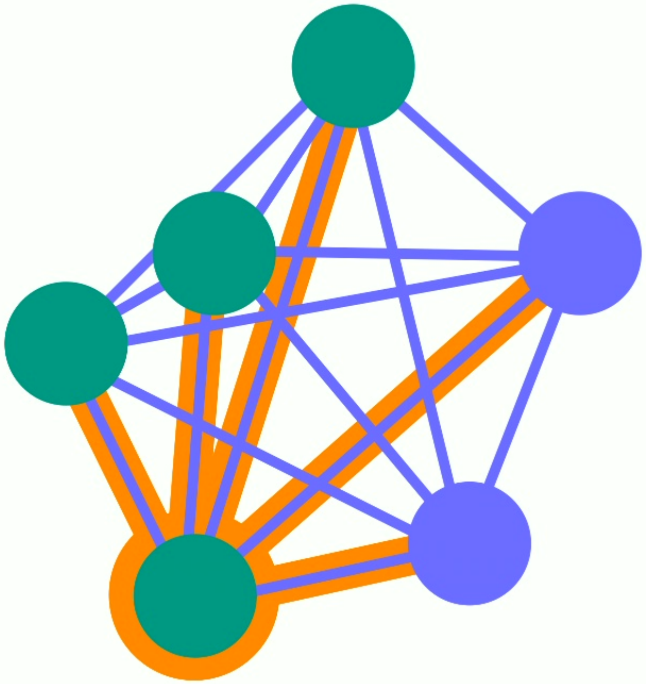
$$\mathbf{m}_{ij}^{(t)} = \mathbf{M}_t\left(\mathbf{h}_i^{(t-1)}, \mathbf{h}_j^{(t-1)}, \mathbf{h}_{ij}^{(t-1)}\right)$$

$$\mathbf{h}_{ij}^{(t)} = \left(\mathbf{v}_{ij}, \mathbf{G}_t\left(\mathbf{h}_{ij}^{(t-1)}, \mathbf{m}_{ij}^{(t)}\right)\right)$$

$$\mathbf{m}_i^{(t)} = \mathrm{Pool}\left(\{\mathbf{m}_{ij}^{(t)} \,|\, j \neq i\}\right)$$

$$\mathbf{h}_i^{(t)} = \left(\mathbf{v}_i, \mathbf{F}_t\left(\mathbf{h}_i^{(t-1)}, \mathbf{m}_i^{(t)}\right)\right)$$

MESSAGE-PASSING NEURAL NETWORK



for $t = 1, \ldots, T$:

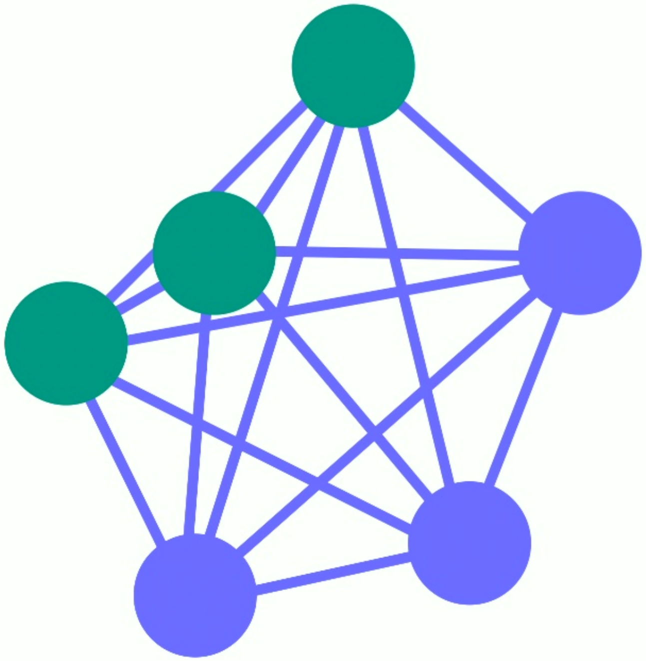
$$\mathbf{m}_{ij}^{(t)} = \mathbf{M}_t\left(\mathbf{h}_i^{(t-1)}, \mathbf{h}_j^{(t-1)}, \mathbf{h}_{ij}^{(t-1)}\right)$$

$$\mathbf{h}_{ij}^{(t)} = \left(\mathbf{v}_{ij}, \mathbf{G}_t\left(\mathbf{h}_{ij}^{(t-1)}, \mathbf{m}_{ij}^{(t)}\right)\right)$$

$$\mathbf{m}_i^{(t)} = \mathrm{Pool}\left(\{\mathbf{m}_{ij}^{(t)} \,|\, j \neq i\}\right)$$

$$\mathbf{h}_i^{(t)} = \left(\mathbf{v}_i, \mathbf{F}_t\left(\mathbf{h}_i^{(t-1)}, \mathbf{m}_i^{(t)}\right)\right)$$

MESSAGE-PASSING NEURAL NETWORK



for $t = 1, \ldots, T$:

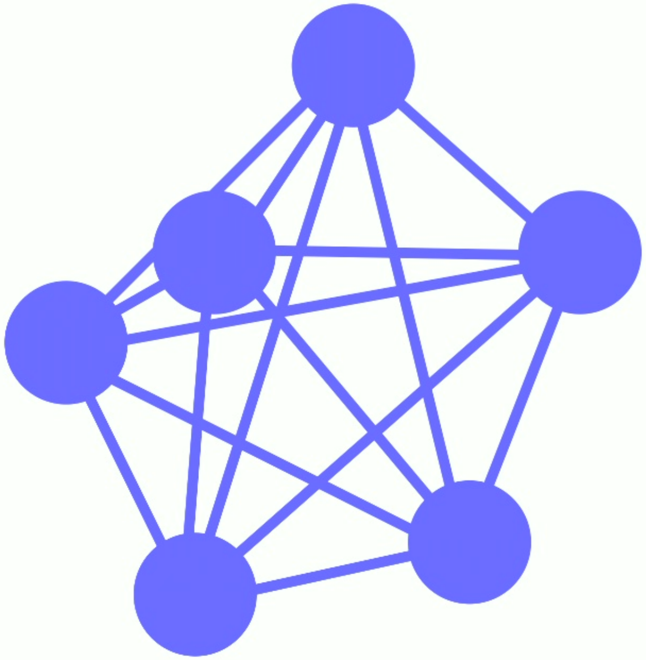
$$\mathbf{m}_{ij}^{(t)} = \mathbf{M}_t \left(\mathbf{h}_i^{(t-1)}, \mathbf{h}_j^{(t-1)}, \mathbf{h}_{ij}^{(t-1)} \right)$$

$$\mathbf{h}_{ij}^{(t)} = \left(\mathbf{v}_{ij}, \mathbf{G}_t \left(\mathbf{h}_{ij}^{(t-1)}, \mathbf{m}_{ij}^{(t)} \right) \right)$$

$$\mathbf{m}_i^{(t)} = \mathrm{Pool} \left(\{ \mathbf{m}_{ij}^{(t)} \,|\, j \neq i \} \right)$$

$$\mathbf{h}_i^{(t)} = \left(\mathbf{v}_i, \mathbf{F}_t \left(\mathbf{h}_i^{(t-1)}, \mathbf{m}_i^{(t)} \right) \right)$$

MESSAGE-PASSING NEURAL NETWORK



for $t = 1, \ldots, T$:

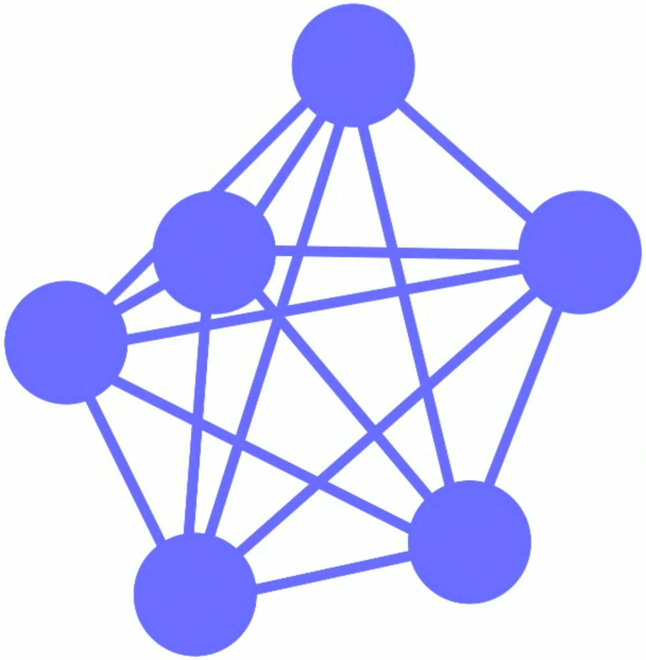
$$\mathbf{m}_{ij}^{(t)} = \mathbf{M}_t\left(\mathbf{h}_i^{(t-1)}, \mathbf{h}_j^{(t-1)}, \mathbf{h}_{ij}^{(t-1)}\right)$$

$$\mathbf{h}_{ij}^{(t)} = \left(\mathbf{v}_{ij}, \mathbf{G}_t\left(\mathbf{h}_{ij}^{(t-1)}, \mathbf{m}_{ij}^{(t)}\right)\right)$$

$$\mathbf{m}_i^{(t)} = \mathrm{Pool}\left(\{\mathbf{m}_{ij}^{(t)} \,|\, j \neq i\}\right)$$

$$\mathbf{h}_i^{(t)} = \left(\mathbf{v}_i, \mathbf{F}_t\left(\mathbf{h}_i^{(t-1)}, \mathbf{m}_i^{(t)}\right)\right)$$

MESSAGE-PASSING NEURAL NETWORK



for $t = 1, \ldots, T$:

$$\mathbf{m}_{ij}^{(t)} = \mathbf{M}_t\left(\mathbf{h}_i^{(t-1)}, \mathbf{h}_j^{(t-1)}, \mathbf{h}_{ij}^{(t-1)}\right)$$

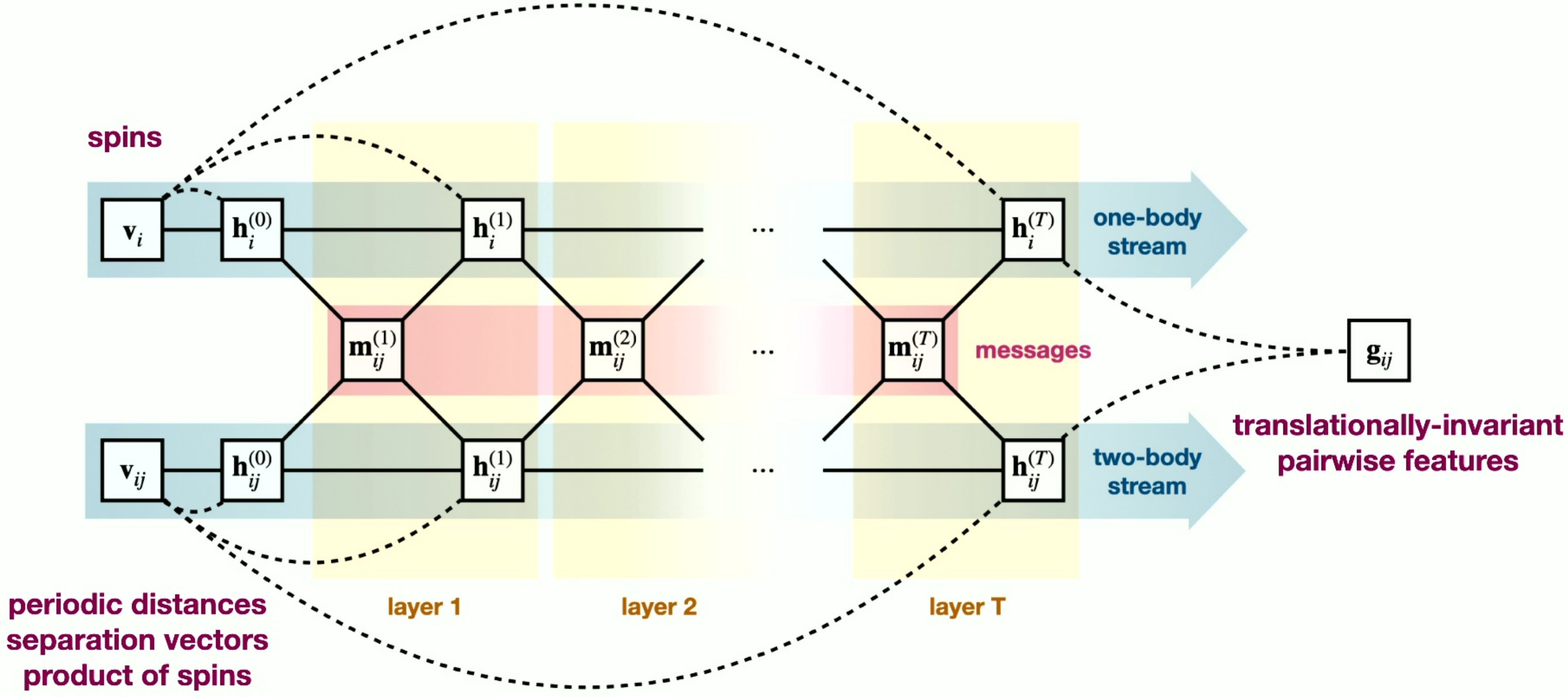
$$\mathbf{h}_{ij}^{(t)} = \left(\mathbf{v}_{ij}, \mathbf{G}_t\left(\mathbf{h}_{ij}^{(t-1)}, \mathbf{m}_{ij}^{(t)}\right)\right)$$

Feedforward neural networks

$$\mathbf{m}_i^{(t)} = \mathrm{Pool}\left(\{\mathbf{m}_{ij}^{(t)} \,|\, j \neq i\}\right)$$

$$\mathbf{h}_i^{(t)} = \left(\mathbf{v}_i, \mathbf{F}_t\left(\mathbf{h}_i^{(t-1)}, \mathbf{m}_i^{(t)}\right)\right)$$

MESSAGE-PASSING NEURAL NETWORK



PFAFFIAN WITH NEURAL BACKFLOW

- Feed output of MPNN into pairing orbital instead of raw data

$$\Phi(X) = \mathrm{pf}\left[\phi(\mathbf{x}_i, \mathbf{x}_j)\right] \;\mapsto\; \Phi(X) = \mathrm{pf}\left[\phi(\mathbf{g}_{ij})\right]$$

- Jastrow correlator is a Deep Set

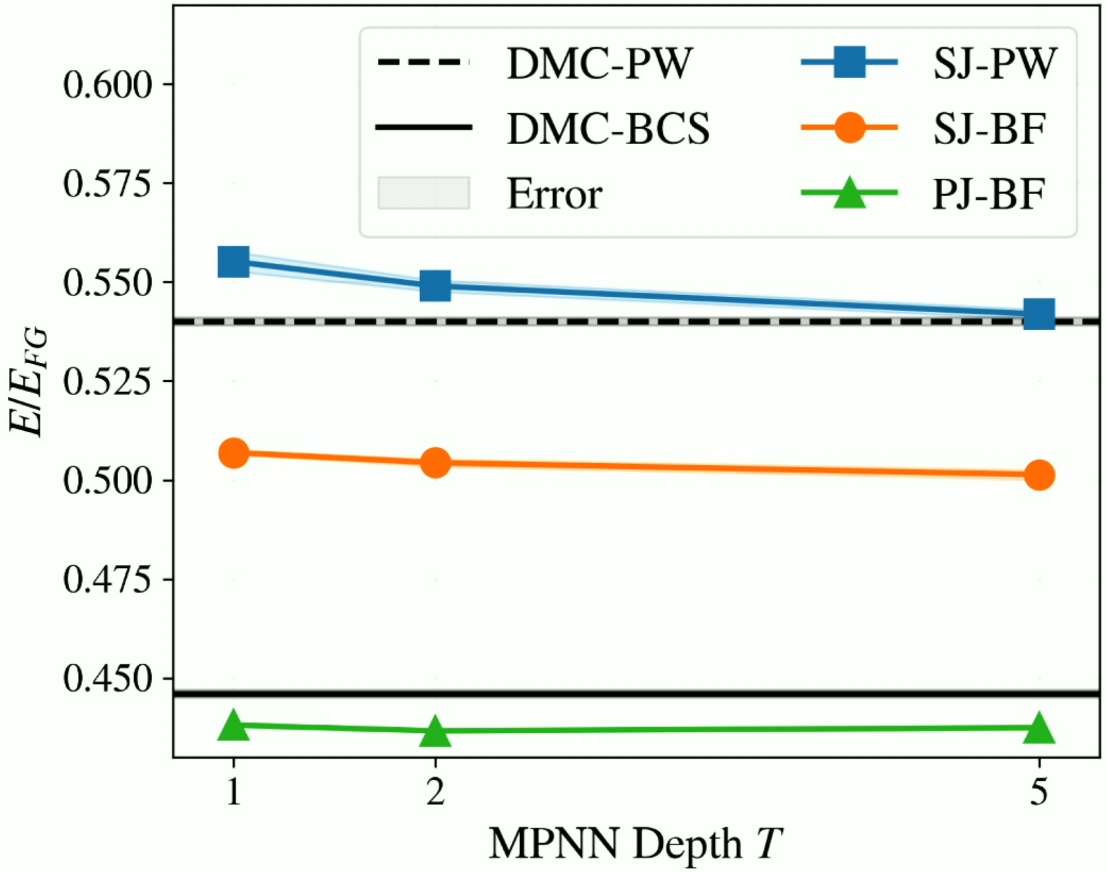
$$J(X) = \underset{\uparrow}{\rho}\left(\mathrm{Pool}\left\{\underset{\uparrow}{\zeta}(\mathbf{g}_{ij})\right\}\right)$$

Feedforward neural networks

- We enforce periodicity, translational invariance, parity and time-reversal symmetries
- First time neural backflow transformations have been applied to the Pfaffian ☺

Deep Sets: Zaheer et al., 31st NeurIPS, Long Beach, CA, USA (2017)

INITIAL COMPARISON: NQS AND DMC



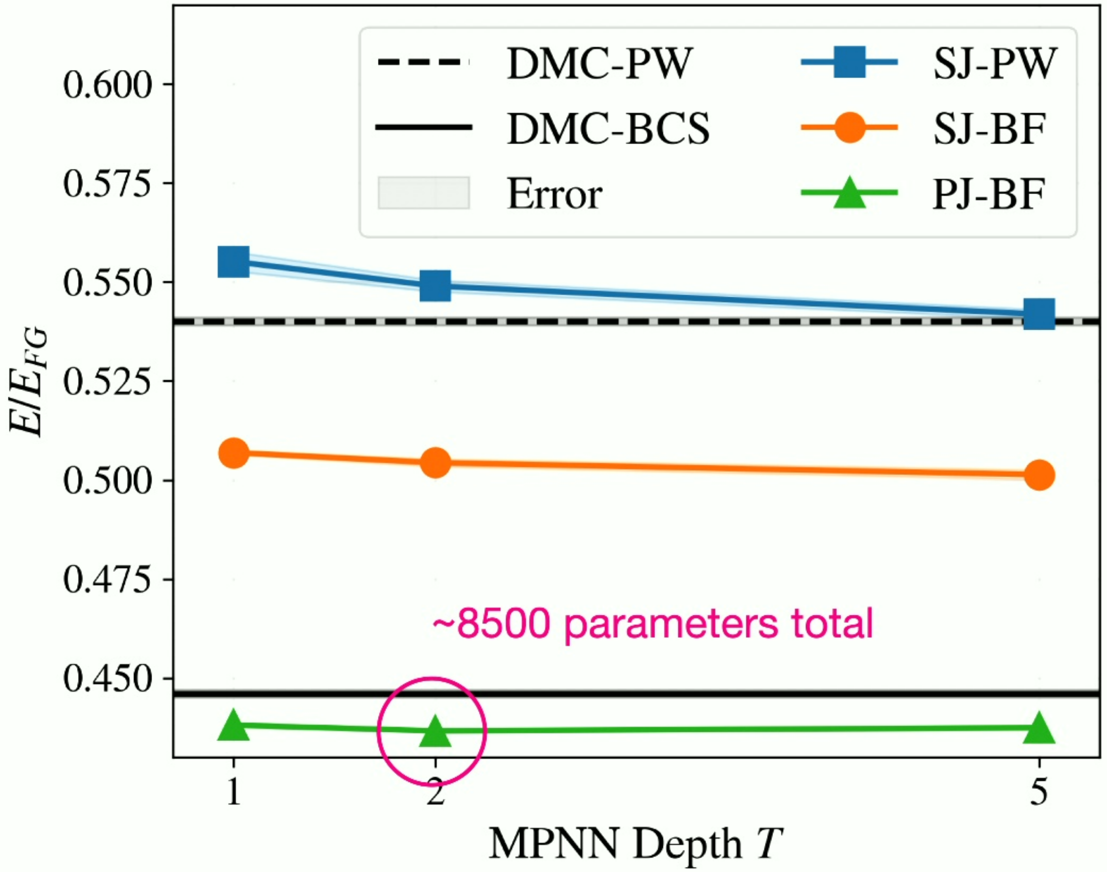
First stage of transfer learning

$k_F r_e = 0.4$

$1/ak_F = 0$

$N = 14$

INITIAL COMPARISON: NQS AND DMC



~8500 parameters total

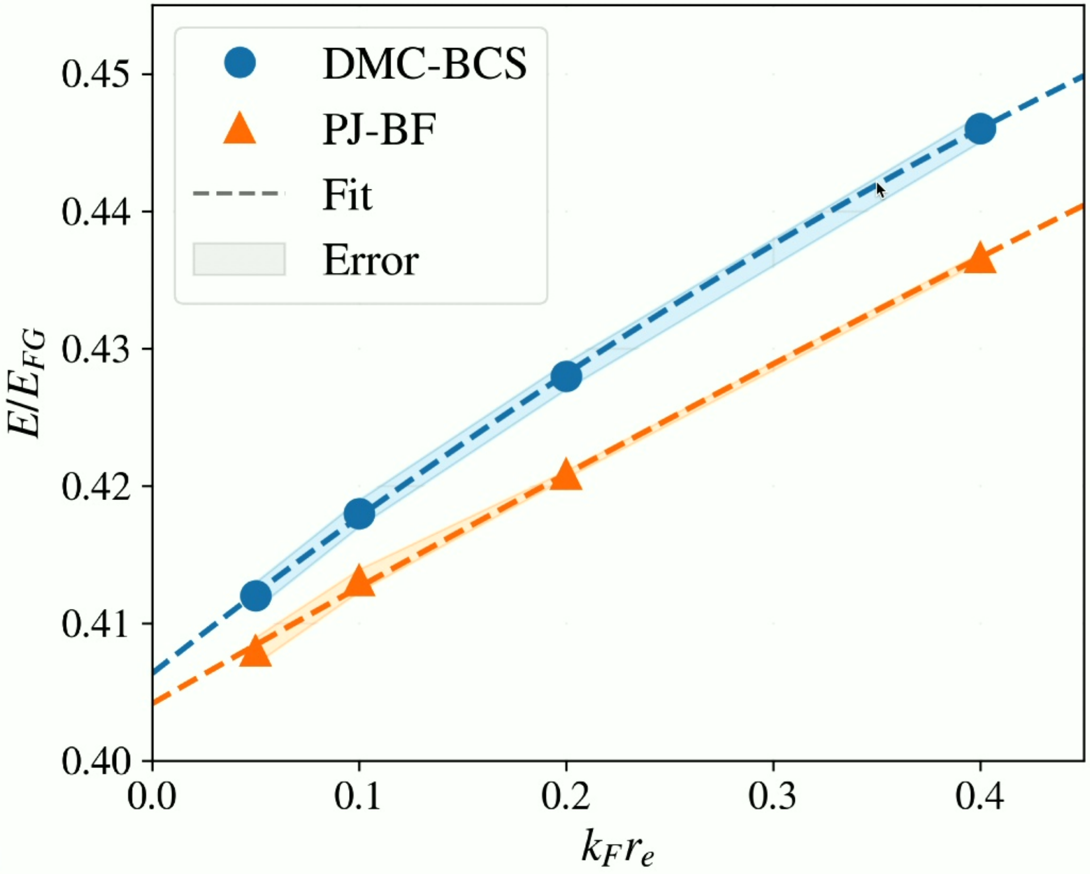
First stage of transfer learning

$k_F r_e = 0.4$

$1/ak_F = 0$

$N = 14$

EXTRAPOLATION TO ZERO EFFECTIVE RANGE

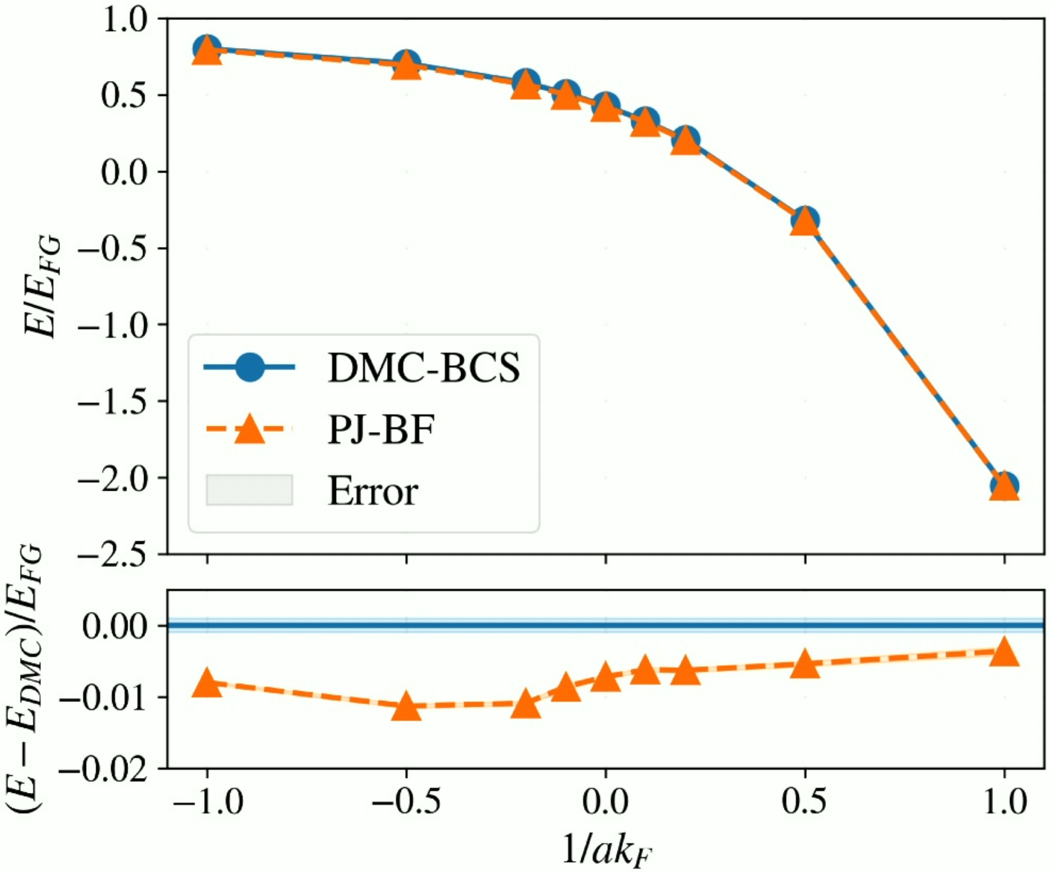


Transfer learning: gradually reduce r_e

$$1/ak_F = 0$$

$$N = 14$$

THE BCS-BEC CROSSOVER

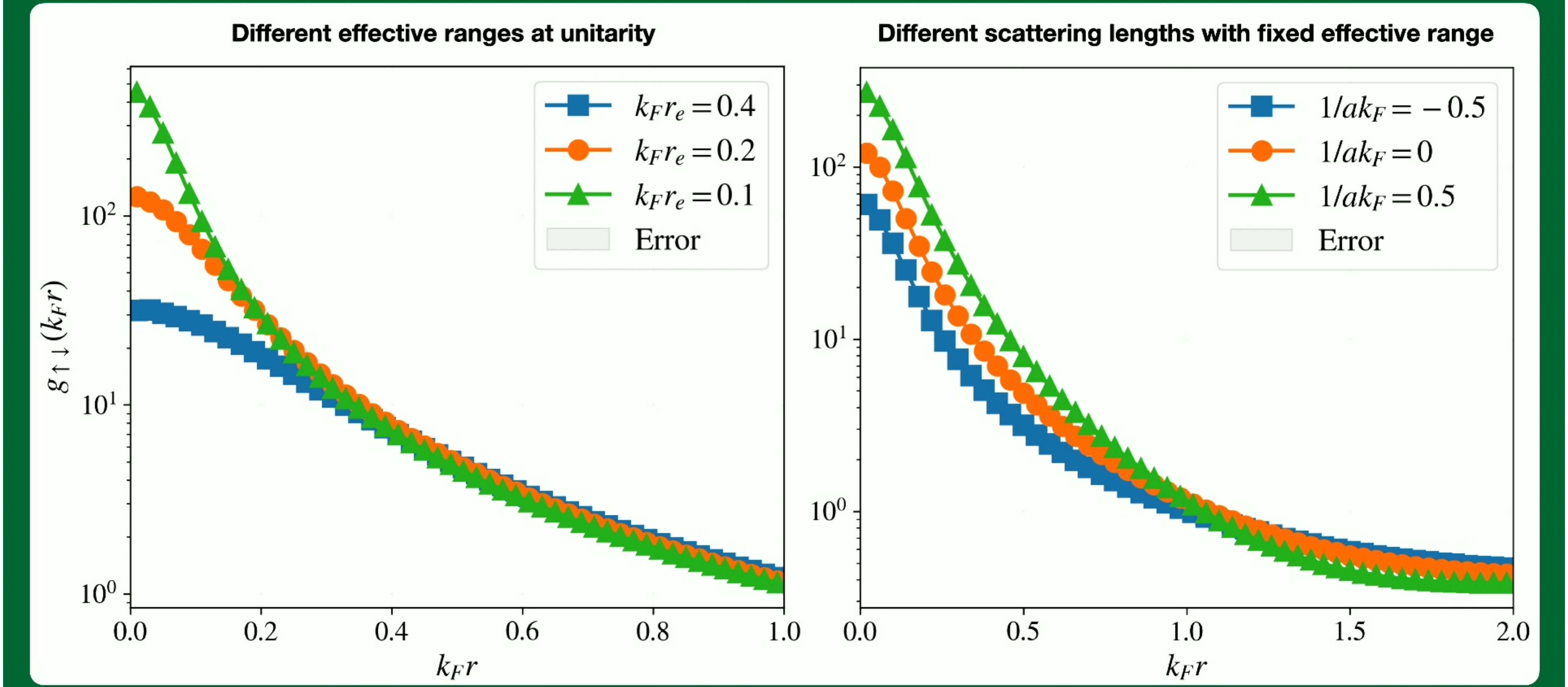


Transfer learning: move away from unitarity

$$k_F r_e = 0.2$$

$$N = 14$$

OPPOSITE-SPIN PAIR DISTRIBUTIONS



Different effective ranges at unitarity

Different scattering lengths with fixed effective range

N-INDEPENDENCE

- The total number of parameters for even N:

$$\big(T(3D+7)+3D+5\big)H^2+\big(T(4d+3D+10)+6d+3D+14\big)H+2$$

Spatial dimension

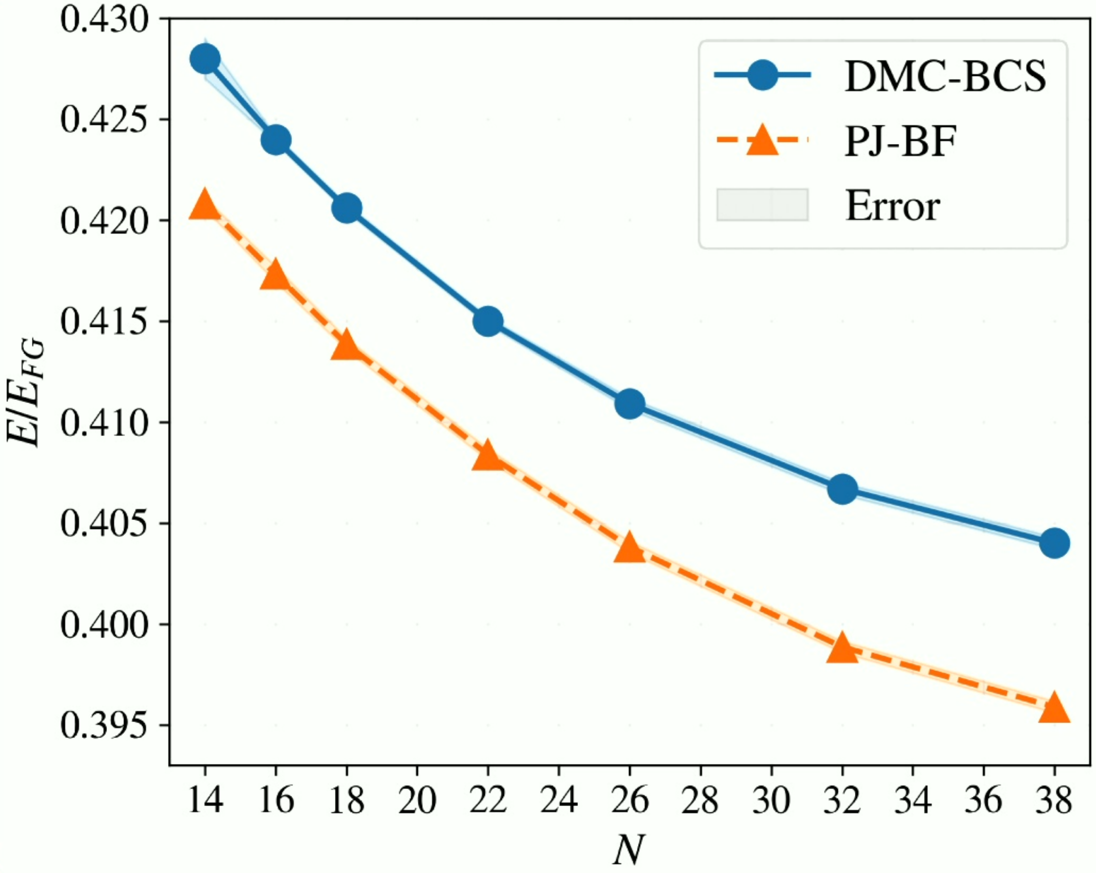
- Hyperparameters:

$T =$ Number of message-passing iterations

$D =$ Number of dense hidden layers in a single feedforward neural network

$H =$ Number of hidden nodes in a single dense layer

TOWARDS THE THERMODYNAMIC LIMIT…



Transfer learning: increase N

$$k_F r_e = 0.2$$

$$1/ak_F = 0$$

ODD N

- Require one additional neural network for the unpaired single-particle orbital

$$\Phi(X) = \mathrm{pf}\begin{bmatrix} 0 & \phi(\mathbf{g}_{12}) & \cdots & \phi(\mathbf{g}_{1N}) & \psi(\mathbf{x}_1) \\ \phi(\mathbf{g}_{21}) & 0 & \cdots & \phi(\mathbf{g}_{2N}) & \psi(\mathbf{x}_2) \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \phi(\mathbf{g}_{N1}) & \phi(\mathbf{g}_{N2}) & \cdots & 0 & \psi(\mathbf{x}_N) \\ -\psi(\mathbf{x}_1) & -\psi(\mathbf{x}_2) & \cdots & -\psi(\mathbf{x}_N) & 0 \end{bmatrix}$$

- Even- and odd-N cases share the same pairing orbital structure (transferrable learning)

- In practice…

$$\psi(\mathbf{x}_i) \mapsto \psi(\mathbf{r}_i, \mathbf{h}_i^{(T)})$$

One-body output of MPNN

PAIRING GAP

- Odd-even staggering: $$\Delta(N) = E(N) - \frac{1}{2}\big(E(N+1) + E(N-1)\big)$$
- Reference values: $$\Delta_{exp} = 0.45(5)\ \varepsilon_F$$

$$\Delta_{DMC-BCS}(66) = 0.50(2)\ \varepsilon_F$$

$\Delta(15)$ →

* = broken translational symmetry

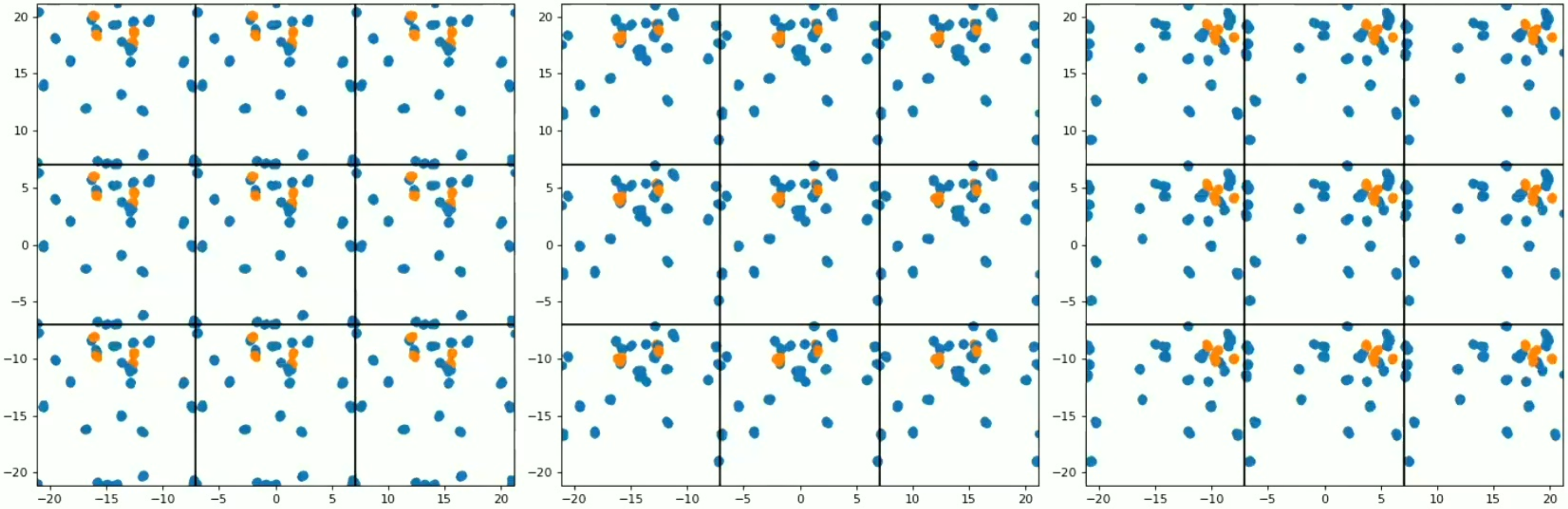
$1/ak_F$	DMC-BCS	PJ-BF	Diff.
-0.5	0.434(6)*	0.426(7)*	-0.008(9)
0	0.577(8)*	0.519(8)*	-0.06(1)
0	0.988(8)	0.918(8)	-0.07(1)
0.5	1.058(6)*	0.962(8)*	-0.10(1)

CONCLUSIONS

- Our Pfaffian wave function is *very* general—even works with Hamiltonians that exchange spin!
- Transfer learning: explore BCS-BEC crossover and thermodynamic limit more efficiently
- Odd- and even-N cases treated in a unified manner
- We require far fewer parameters than other NQS applied to the same problem
- Future work:
    - Direct comparison with AFQMC
    - Partially-spin polarized systems
    - Finite-temperature VMC → find critical temperature?
    - Real-time dynamics → quantized vortices?

APPLICATIONS: DILUTE NUCLEAR MATTER

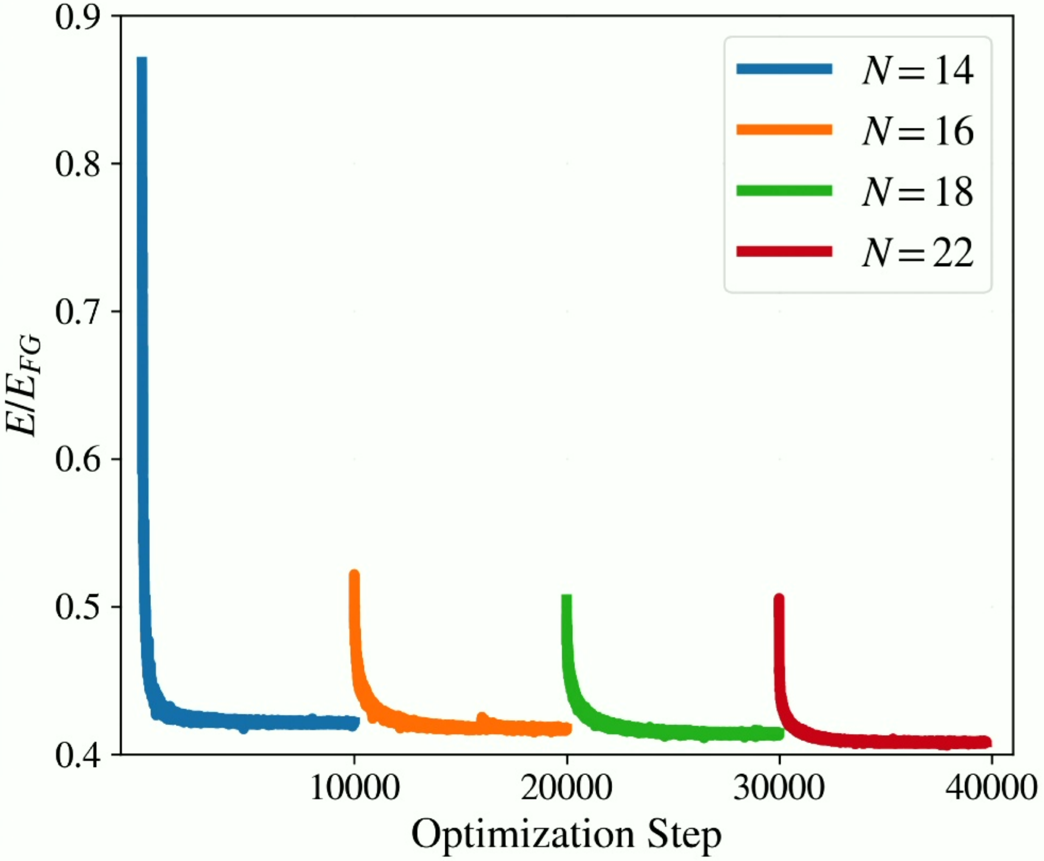
- Very simple pionless-EFT Hamiltonian at leading order
- Proton clusters are visible at low densities $\rho = 0.01 \ \mathrm{fm}^{-3}$
- Can simulate nuclear matter at densities as low as $\rho = 0.001 \ \mathrm{fm}^{-3}$



Credit: Bryce Fore

THANK YOU!

TRANSFER LEARNING



Transfer learning: increase N

$$k_F r_e = 0.2$$

$$1/ak_F = 0$$

COMPUTATIONAL SCALING

Using 4 NVIDIA-A100s for $N \leq 32$

Empirical scaling $\mathcal{O}(N^{1.894})$