Bosonic Helium Clusters within Neural Quantum States

ESNT Workshop — CEA Saclay, DPhN, Orme des Merisiers

Liron H. Avraham

Supervisor: Dr. Betzalel Bazak

The Hebrew University of Jerusalem

The Racah Institute of Physics

2 July, 2026

Outline

1. Motivation
2. Bridging Machine Learning & Physics
3. Bosonic Cluster with Gaussian Potential
4. $^{4}\text{He}_N$ with $V^{\text{Aziz}}_{\text{LM2M2}}(r)$
5. Summary

Motivation

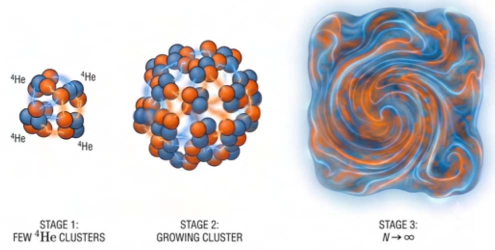
§1 Motivation: Atomic $^4\mathrm{He}_N$ Clusters

1. Why $^4\mathrm{He}_N$ Are Interesting?

- Universality – Dimer, Trimer, Tetramer etc.
- Experimental comparisons
- Quantum liquid (Superfluid)
- Short-range correlations analysis

2. Challenges

- Wavefunction complexity
- Symmetrization $\sim \mathcal{O}(N!)$
- Computational resources
- Severe short-range core repulsion



E. Braaten & H.-W. Hammer, *Phys. Rep.* **428** (2006)
B. Bazak *et al.*, *Phys. Rev. Lett.* **122** (2019) 143001 ; B. Bazak *et al.*, *Phys. Rev. A* **101** (2020) 010501(R)

Bridging Machine Learning & Physics

Deep Neural Network (DNN)

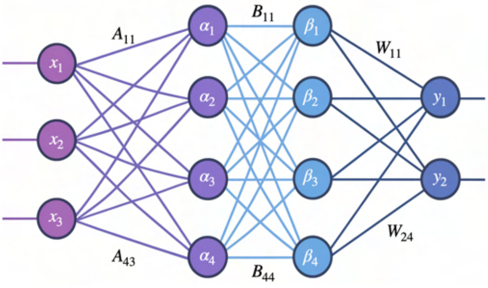
Goal: Approximate a target function f^*

$$\mathbf{y} = f(\mathbf{x};\boldsymbol{\theta}) = f^{(N)} \circ \cdots \circ f^{(1)}(\mathbf{x})$$

$$f^{(i)} \propto (\text{Non-linear}) \circ (\text{Linear})$$

$$\downarrow$$

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \tanh\left\{ \begin{bmatrix} W_{11} & W_{12} & W_{13} & W_{14} \\ W_{21} & W_{22} & W_{23} & W_{24} \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \end{bmatrix} + \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} \right\}$$



Feedforward Neural Network (FNN)

Input layer **Hidden layers** **Output layer**

$$\boxed{\mathbf{x}} \qquad \boxed{\rightarrow\ \alpha\ \rightarrow\ \beta\ \rightarrow} \qquad \boxed{\mathbf{y}}$$

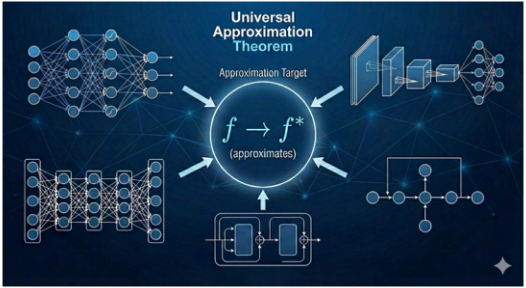
I. Goodfellow *et al.*, *Deep Learning* (MIT Press, 2016)

Universal Approximation Theorem

FNN $\approx f^*$ to **arbitrary accuracy**,
provided sufficient **features**,
architecture **independent**.

However

- Representability > Trainability
- Tailored architectures
- Few/many-body observables



I. Goodfellow *et al.*, *Deep Learning* (MIT Press, 2016)

§2 Briding ML & Physics: Neural Quantum States Framework

NQS Mathematical Formulation

$$\mathcal{H} = \mathrm{span}\left\{|\mathbf{x}\rangle\right\}$$

$$\mathcal{I} = \sum_{\mathbf{x}} |\mathbf{x}\rangle\langle\mathbf{x}|$$

$$\text{VNQS} \equiv \left|\psi_{\boldsymbol{\theta}}\right\rangle = \sum_{\mathbf{x}} \psi_{\boldsymbol{\theta}}(\mathbf{x})\,|\mathbf{x}\rangle$$

$$E\left(\theta\right) = \frac{\langle\psi_\theta|H|\psi_\theta\rangle}{\left\langle\psi_\theta\middle|\psi_\theta\right\rangle} \geq E_0 \quad \text{(Variational Theorem)}$$

Review: M. Medvidović, J. Robledo Moreno, *Eur. Phys. J. Plus* **139** (2024) 631

§2 Briding ML & Physics: Observables Estimation

Quantum Monte Carlo (QMC)

$$\langle \mathcal{H} \rangle_{\psi_{\boldsymbol{\theta}}} = \sum_{\mathbf{x}} p_{\theta}(\mathbf{x})\, \mathcal{H}_L(\mathbf{x}; \boldsymbol{\theta}) \approx \frac{1}{N_s} \sum_{i=1}^{N_s} \mathcal{H}_L(\mathbf{x}_i)$$

$$p_{\theta}(\mathbf{x}) = \frac{\left|\psi_{\boldsymbol{\theta}}(\mathbf{x})\right|^2}{\sum_{\mathbf{x}'} \left|\psi_{\boldsymbol{\theta}}(\mathbf{x}')\right|^2} \quad \text{(PDF)}$$

$$\mathcal{H}_L(\mathbf{x}; \boldsymbol{\theta}) = \frac{\left\langle \mathbf{x} \middle| \mathcal{H} \middle| \psi_{\boldsymbol{\theta}} \right\rangle}{\left\langle \mathbf{x} \middle| \psi_{\boldsymbol{\theta}} \right\rangle} \quad \text{(Local energy function)}$$

Goal

$$\boldsymbol{\theta}_{\text{opt}} = \text{argmin}_{\boldsymbol{\theta}} \left[E(\boldsymbol{\theta}) \right]$$

§2 Briding ML & Physics: Variational Optimization

Ansatz

$\psi_\theta(\mathbf{x}_0) \equiv \langle \mathbf{x}_0 | \psi_{\boldsymbol{\theta}} \rangle$
$\mathbf{x}_0 \equiv$ Initial config.

Sampling

$\mathbf{x}_0 \longrightarrow \mathbf{x}' \sim |\psi_\theta(\mathbf{x})|^2$
(via MCMC)

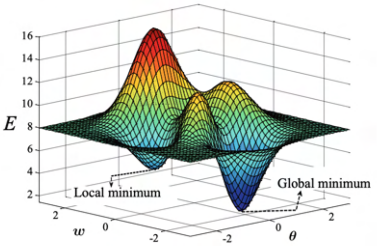
Calculation

$\langle \mathcal{H}_L(\mathbf{x}) \rangle_{\mathbf{x} \sim |\psi_{\boldsymbol{\theta}}|^2}$
$\nabla_{\boldsymbol{\theta}} E(\boldsymbol{\theta})$

Updating

$\theta' \equiv \theta + \delta\theta$
(gradient-based)

Iterating until convergence $\Longrightarrow$ $\boldsymbol{\theta}_{\text{opt}} = \text{argmin}_{\boldsymbol{\theta}} [E(\boldsymbol{\theta})]$



Local and global minima illustration in the variational manifold

Figure: Z.-H. Zhou, Machine Learning (Springer, 2021)

Bosonic Cluster with Gaussian Potential

§3 Gaussian Bosonic Cluster: Toy Model

1. Strategy

- Simple yet interesting potential
- SVM comparable
- Pure neural-network ansatz

2. Hamiltonian - Gaussian Interactions

$$\mathcal{H} = -\frac{\hbar^2}{2m}\sum_i \nabla^2_{\mathbf{x}_i} + V_0 \sum_{i<j} \exp\left(-\frac{|\mathbf{r}_{ij}|^2}{r_0^2}\right)$$

$$m = \hbar = r_0 = 1 \quad ; \quad V_0 = -10$$

3. The Ansatz: Deep Set Architecture

$$\log\left[\Psi(\mathbf{x})\right] = \mathsf{DS}(\mathbf{x}) = \rho\left(\mathsf{Pool}\left[\phi\left(\mathbf{x}\right)\right]\right)$$

NetKet Python library

M. Zaheer *et al.*, arXiv (2017) 1703.06114 ; F. Vicentini *et al.*, *SciPost Phys. Codebases* **7** (2022) ;
D. Hendrycs, K. Gimpel, *arXiv* (2023) 1606.08415

§3 Gaussian Bosonic Cluster: DS($\mathbf{x}$) Architecture

Deep Sets Symmetrization - Permutation Invariance

$$\boxed{\mathrm{DS}(\mathbf{x};\boldsymbol{\theta}) = \rho\left(\mathrm{Pool}\left[\phi(\mathbf{x})\right]\right)}$$

$\phi(\mathbf{x})$:	Data Expansion/Reduction
Pool:	Permutation invariance
ρ:	Mapping data to output

Linear: $\hat{\mathbf{T}}_w(\mathbf{h}_f) \equiv \hat{\mathbf{W}}\mathbf{h}_f + \mathbf{b}_w \quad (\mathbf{h}_f \in \mathbb{R}^f,\ \hat{\mathbf{W}} \in \mathbb{R}^{w\times f},\ \mathbf{b}_w \in \mathbb{R}^w)$

Activation (ex.):

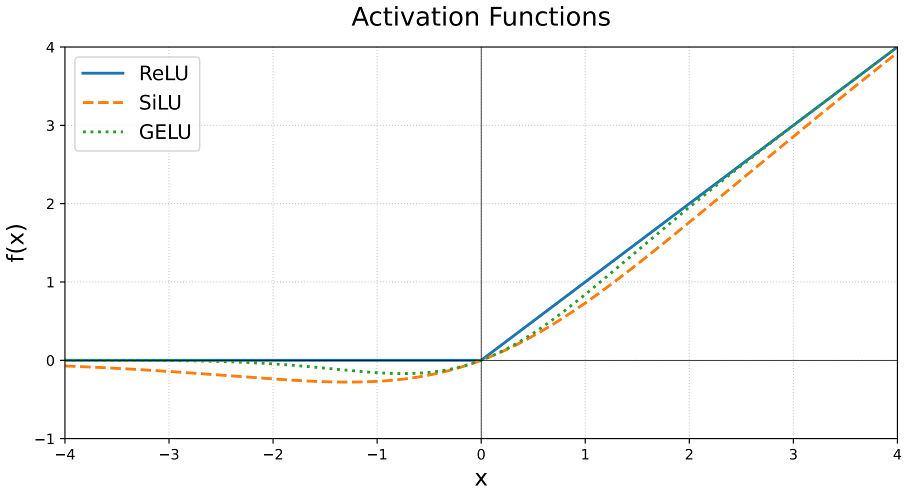
$$\mathrm{ReLU}(x) = \max(0, x)$$

$$\mathrm{GELU}(x) \approx 0.5x\left[1 + \tanh\left(\sqrt{2/\pi}(x + 0.044715x^3)\right)\right]$$

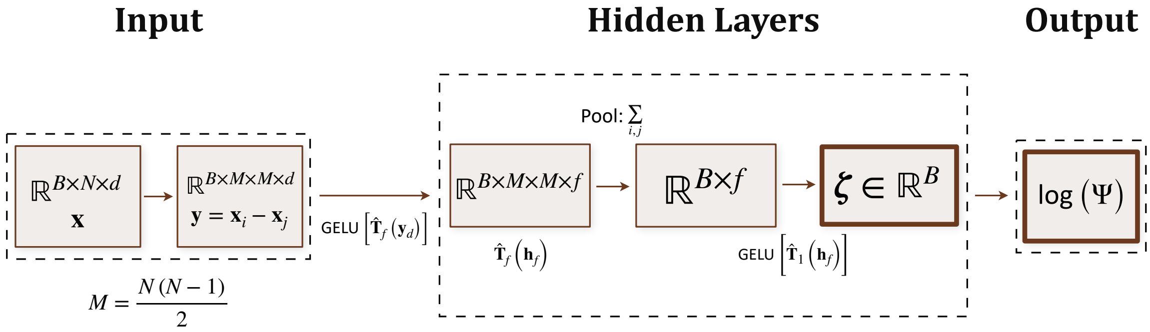
$$\mathrm{SiLU}(x) \equiv x\cdot\sigma(x) = \frac{x}{1+e^{-x}} \quad (\text{"Swish"})$$

Pooling: $\mathrm{Pool}\left[\phi(\mathbf{x})\right] \implies$ Summation

V. Nair and G. E. Hinton, *ICML* (2010) ; D. Hendrycs, K. Gimpel, *arXiv* (2016) 1606.08415 ;
P. Ramachandran *et al.*, arXiv (2017) 1710.05941 ; M. Zaheer et al., arXiv (2017) 1703.06114

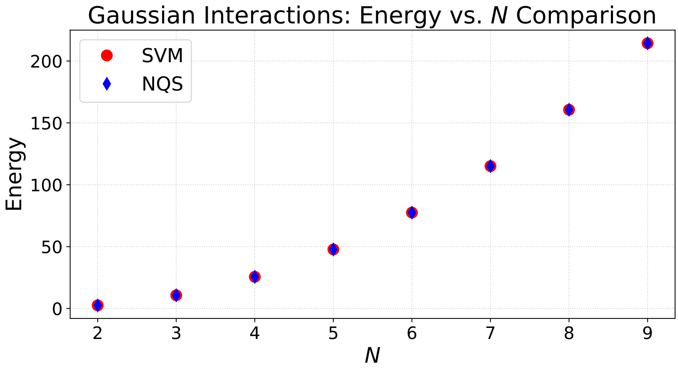

Activation Functions
ReLU
SiLU
GELU
f(x)
x
4
3
2
1
0
−1
−4
−3
−2
−1
0
1
2
3
4

§3 Gaussian Bosonic Cluster: Deep Set Architecture



$$\boxed{\log\left(\Psi\right) = \zeta}$$

§3 Gaussian Bosonic Cluster: SVM vs. NQS Comparison



N	NQS	SVM
2	2.542(3)	2.543
3	10.70(1)	10.713
4	25.607(6)	25.606
5	47.751(5)	47.754
6	77.49(1)	77.489
7	115.064(3)	115.054
8	160.621(4)	160.616
9	214.342(5)	214.327

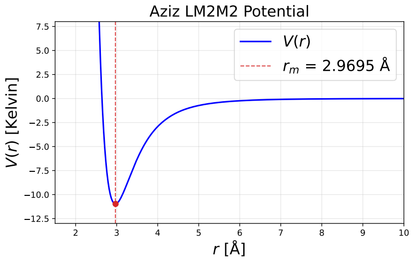
^{4}He$_N$ with $V^{\mathrm{Aziz}}_{\mathrm{LM2M2}}(r)$

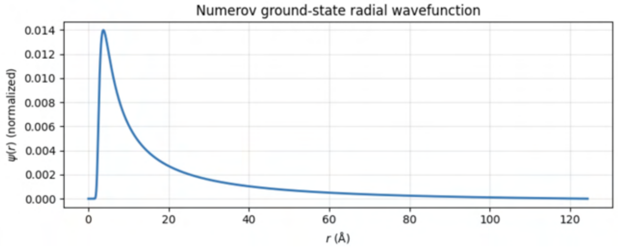
§4 $^4\text{He}_N$ Cluster with LM2M2: System Setup

1. Hamiltonian - LM2M2 Interactions

$$\mathcal{H} = -\frac{\hbar^2}{2mk_B}\sum_i \nabla^2_{\mathbf{r}_i} + \Gamma \cdot \frac{\epsilon}{k_B}\sum_{i<j} V_{ij}$$

$$V_{ij} \equiv V^{\text{Aziz}}_{\text{LM2M2}}(|\mathbf{x}_{ij}|)\ ;\ |\mathbf{x}_{ij}| \equiv x_{ij} = \frac{|\mathbf{r}_{ij}|}{r_m}\ ;\ r_m = 2.9695\ \text{Å}\ ;\ \epsilon = 10.97\ \text{K}\ ;\ 0 < \Gamma \leq 1\ \text{(Adiabatic)}$$





R. A. Aziz, M. J. Slaman, *J. Chem. Phys.* **94** (1991) 8047–8053.

§4 $^4\mathrm{He}_N$ Cluster with LM2M2: System Setup - *cont.*

2. The Ansatz: Jastrow & Deep Set

$$\log\left[\Psi(\mathbf{x};\boldsymbol{\theta})\right] \equiv \mathcal{J} + \zeta = \sum_{i<j} \log\left[f(x_{ij})\right] + \lambda \cdot \mathsf{DS}(\mathbf{x};\boldsymbol{\theta})$$

$$\log\left[f(x_{ij})\right] = -\left[p_0 \log(x_{ij}) + p_1 x_{ij} + \frac{p_2}{x_{ij}^2} + \frac{p_5}{x_{ij}^5}\right]$$

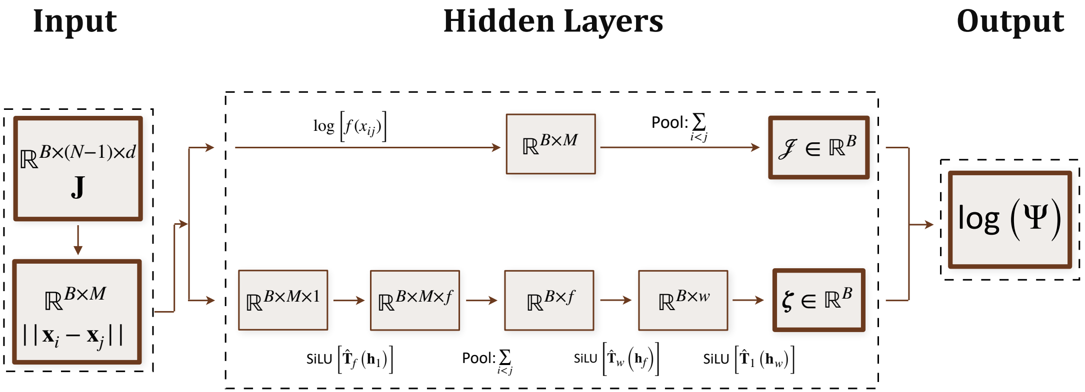
$$\mathsf{DS}(\mathbf{x}) = \rho\left(\sum_{i<j} \phi(x_{ij})\right)$$

The Actual Calculated Observable

$$E_L(\mathbf{x}_0;\boldsymbol{\theta}) = -\frac{\hbar^2}{2m}\sum_{i=1}^{N}\left(\nabla_i^2 \log\Psi(\mathbf{x}_0;\boldsymbol{\theta}) + \left[\nabla_i \log\Psi(\mathbf{x}_0;\boldsymbol{\theta})\right]^2\right) + V(\mathbf{x}_0)$$

S. W. Rick *et al.*, *J. Chem. Phys.* **95** (1991) 3506–3520 ; B. Bazak *et al.*, *Phys. Rev. A* **101** (2020) 010501(R)
M. Zaheer *et al.*, arXiv (2017) 1703.06114 ; D. Hendrycs, K. Gimpel, *arXiv* (2016) 1606.08415

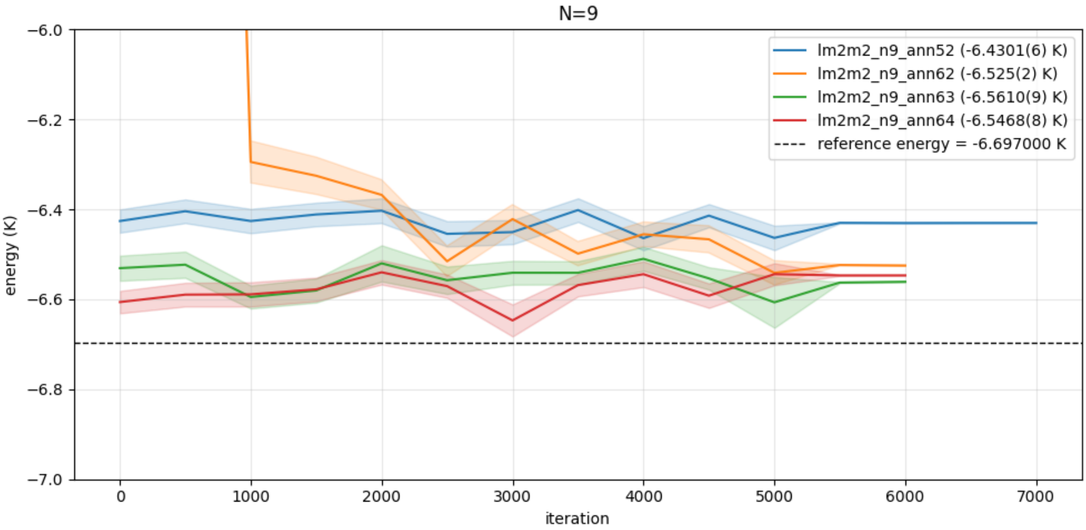
§4 $^4\text{He}_N$ Cluster with LM2M2: DS($\mathbf{x}$) Architecture & Data Flow



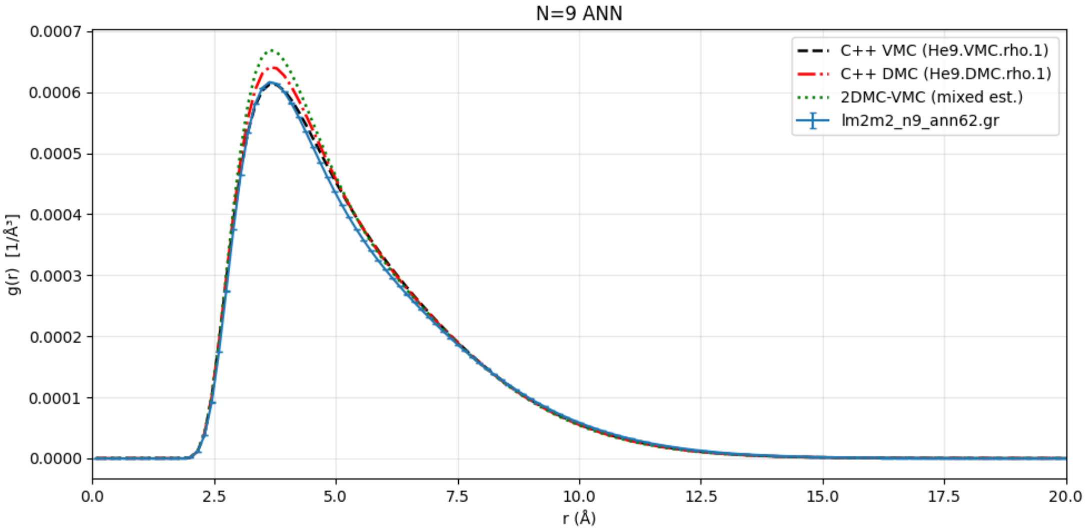
$$M = \frac{N(N-1)}{2}$$

$$\boxed{\begin{aligned}\log(\Psi) &= \mathcal{J} + \xi \\ \xi &= \lambda \cdot \mathrm{DS}(\mathbf{x}, \boldsymbol{\theta})\end{aligned}}$$

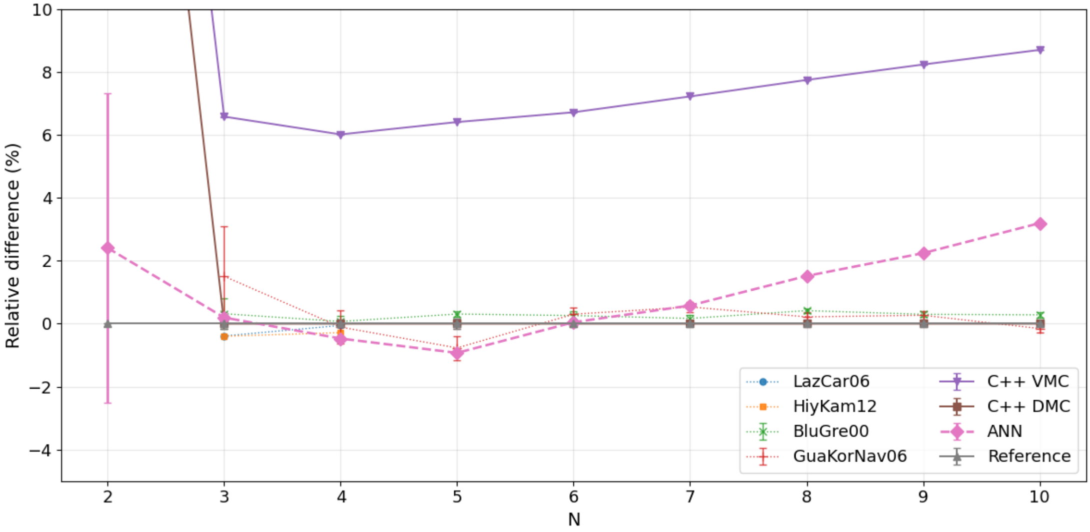
§4 $^4\mathrm{He}_N$ Cluster with LM2M2: Preliminary Results - Convergence



§4 ^{4}He$_N$ Cluster with LM2M2: Preliminary Results - Pair Distribution



§4 $^4\mathrm{He}_N$ Cluster with LM2M2: Comparison to C++



Summary

§4.1 Summary & Looking Forward

Highlights

- Powerful variational numerical method
- Pure NQS isn't enough - Physics must encoded
- Computationally efficient compared to other methods

Next steps

- Refine the model's architecture
- Improve accuracy & minimize variance
- Calculate additional observables
- Extend to other physical systems

§4.2 Questions for Discussion

1. **Resonances & Reactions:**
   - How can we construct an ansatz for resonances and reaction channels?
   - How do we optimize convergence to excited states?
   - Can supervised learning approaches offer advantages over traditional unsupervised methods in capturing resonance states?
2. **Numerical Stability:**
   - **Generally:** Does adiabatic training of the NQS wavefunction mitigate numerical instabilities of hard-core potentials?
   - **Concretely:** Can this approach resolve the specific numerical challenges encountered in the large-cutoff limit in EFTs?
3. **Mathematical and CS Knowledge:**
   - How can we effectively filter the rapid advancements in machine learning for physics-specific applications?
   - Which foundational theorems and neural network architectures are best suited for enforcing physical constraints?

Questions?

Thank you.