Neural quantum states for hypernuclear systems with contact theories.

A. Di Donna[1] A. Lovato[2] F. Pederiva[3] L. Contessi[4]

[1] Faculty of Physics, Unitn
TIFPA

[2] Argonne National Laboratory
TIFPA

[3] University of Trento
TIFPA

[4] IJCLab
Orsay

April 11, 2024

Argonne NATIONAL LABORATORY

TIFPA Trento Institute for Fundamental Physics and Applications

Table of Contents

EoS of Barion Stars and the Hyperon puzzle

- **Equation of State** (EoS) → barion stars' internal structure (BS) → prediction of the maximum stellar mass.

EoS of Barion Stars and the Hyperon puzzle

- **Equation of State** (EoS) → barion stars' internal structure (BS) → prediction of the maximum stellar mass.
- **Hyperons in the core** of the most massive BSs lead to the **hyperon puzzle** → softer EoS (reduce Pauli blocking), reducing the maximum stellar mass.

EoS of Barion Stars and the Hyperon puzzle

- **Equation of State** (EoS) → barion stars' internal structure (BS) → prediction of the maximum stellar mass.
- **Hyperons in the core** of the most massive BSs lead to the **hyperon puzzle** → softer EoS (reduce Pauli blocking), reducing the maximum stellar mass.
- Observational data of the most recent compact object's masses have provided stringent constraints on the EoS that are in contrast with its softening.

EoS of Barion Stars and the Hyperon puzzle

- **Equation of State** (EoS) → barion stars' internal structure (BS) → prediction of the maximum stellar mass.
- **Hyperons in the core** of the most massive BSs lead to the **hyperon puzzle** → softer EoS (reduce Pauli blocking), reducing the maximum stellar mass.
- Observational data of the most recent compact object's masses have provided stringent constraints on the EoS that are in contrast with its softening.
- The main ingredient that determines the stiffness/softness of the EOS after the introduction of hyperons in the core is the **microscopic interaction** between these particles and the nuclear matter.

Key question

- Key question:

Key question

- Key question:
  - → **can Neural Network Quantum States (NNQS) model hypernuclear systems and allow to predict their properties?**
  - → Hypernuclei considered in this work are bound state between an ordinary nucleus with only one hyperon (Λ^0)

$$^{A-1}Z + \Lambda \rightarrow ^{A}_{\Lambda} Z$$

Key question

- Key question:
  - → **can Neural Network Quantum States (NNQS) model hypernuclear systems and allow to predict their properties?**
  - → Hypernuclei considered in this work are bound state between an ordinary nucleus with only one hyperon (Λ^0)

$$^{A-1}Z + \Lambda \rightarrow ^{A}_{\Lambda} Z$$

- **Objectives**
  - → **Parametrize for the wavefunction (variational state) in terms of Neural Networks**,
  - → **Fit an Improved LO interaction** potential derived from $\not{\pi}EFT$.
  - → **Compare NNQS results** with exact few-body techniques over fitted observables.
  - → **Predict Λ-separation energies** for high A systems ($^{7}_{\Lambda}$Li, $^{13}_{\Lambda}$C, $^{16}_{\Lambda}$O and ~~$^{40}_{\Lambda}$Ca~~).

Key question

- Key question:
  - → **can Neural Network Quantum States (NNQS) model hypernuclear systems and allow to predict their properties?**
  - → Hypernuclei considered in this work are bound state between an ordinary nucleus with only one hyperon (Λ^0)

$$^{A-1}Z + \Lambda \rightarrow^{A}_{\Lambda} Z$$

- **Objectives**
  - → **Parametrize for the wavefunction (variational state) in terms of Neural Networks**,
  - → **Fit an Improved LO interaction** potential derived from $\not\pi EFT$.
  - → **Compare NNQS results** with exact few-body techniques over fitted observables.
  - → **Predict** Λ**-separation energies** for high A systems ($^{7}_{\Lambda}$Li, $^{13}_{\Lambda}$C, $^{16}_{\Lambda}$O and ~~$^{40}_{\Lambda}$Ca~~).
- Through NNQS, we have obtained **promising results** with hypernuclear states.

Microscopic Interaction: $\not\pi$EFT Contact potentials

- The interaction potential is derived from the Pionless EFT ($\not\pi$EFT $Q < M_{hi} = m_\pi = 140 MeV$)

Microscopic Interaction: $\not{\pi}$EFT Contact potentials

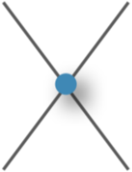
- The interaction potential is derived from the Pionless EFT ($\not{\pi}$EFT $Q < M_{hi} = m_\pi = 140 MeV$)
- Unresolved pions leads to a spin-dependent contact potential, which is described by a set of Low Energy Constants (LECs):

Microscopic Interaction: $\not{\pi}$EFT Contact potentials

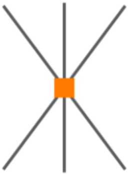
- The interaction potential is derived from the Pionless EFT ($\not{\pi}$EFT $Q < M_{hi} = m_\pi = 140 MeV$)
- Unresolved pions leads to a spin-dependent contact potential, which is described by a set of Low Energy Constants (LECs):

Regulator's cutoffs and LECs are both fitted → **SVM (Suzuki - Varga) and Gaussian Processes**.

- NN potential: Fitted to np/nn scattering lengths and effective ranges.
- 3NF are adjusted to reproduce the ^3H and ^4He BEs.



$$V_{NN}(\mathbf{x}_{ij}) = \left(C_0(\lambda)^s P^{2b}_{S_{tot}=0} + C_0(\lambda)^t P^{2b}_{S_{tot}=1}\right)\delta_\lambda(\mathbf{x}_{ij})$$



$$V_{NNN}(\mathbf{x}_{ij}) = \mathcal{D}_0(\lambda) \sum_{i<j<k} \sum_{cyc} \delta_\lambda(\mathbf{x}_{ik})\delta_\lambda(\mathbf{x}_{ij})$$

Extending $\not{\pi}$EFT to Hyperons

- Extending $\not{\pi}$EFT to Hyperons require the introduction of Λ-hyperons ($m_\Lambda = 1116 MeV$) DoF in the Lagrangian density, $\mathcal{L}$

$$\mathcal{L} = N^\dagger\Big(i\partial_0 + \frac{\nabla^2}{2M_N}\Big)N + \Lambda^\dagger\Big(i\partial_0 + \frac{\nabla^2}{2M_\Lambda}\Big)\Lambda + \mathcal{L}_{2b} + \mathcal{L}_{3b} + \dots$$

Extending $\not{\pi}$EFT to Hyperons

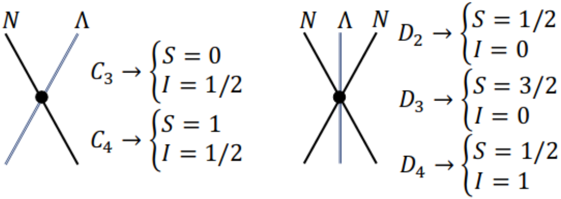
- Extending $\not{\pi}$EFT to Hyperons require the introduction of Λ-hyperons ($m_\Lambda = 1116 MeV$) DoF in the Lagrangian density, $\mathcal{L}$

$$\mathcal{L} = N^\dagger\Big(i\partial_0 + \frac{\nabla^2}{2M_N}\Big)N + \Lambda^\dagger\Big(i\partial_0 + \frac{\nabla^2}{2M_\Lambda}\Big)\Lambda + \mathcal{L}_{2b} + \mathcal{L}_{3b} + \dots$$

- The interaction potential becomes at LO:

$$V_{\Lambda N} = \sum_{IS} C_\lambda^{IS} \sum_{i<j} \mathcal{P}_{IS}(ij)\delta_\lambda(\vec{r}_{ij})$$

$$V_{\Lambda NN} = \sum_{IS} D_\lambda^{IS} \sum_{i<j} \mathcal{Q}_{IS}(ij\Lambda)\delta_\lambda(\vec{r}_{i\Lambda})\delta_\lambda(\vec{r}_{j\Lambda})$$

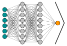


- $\mathcal{P}_{IS}$ ($\mathcal{Q}_{IS}$) are projectors on baryon doublets (triplets) with isospin I and spin S

1. ΛN interaction is fitted to $p\Lambda$ scattering length and effective range.
2. ΛNN interaction is fitted to ${}^3_\Lambda H$, ${}^4_\Lambda H_{S_{tot}=0}$, ${}^4_\Lambda H_{S_{tot}=1}$, and ${}^5_\Lambda He$ binding energies.

The NNQS approach

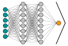
- The Neural Network Quantum States (Carleo et al. 2017) approach is a form of unsupervised learning.
  - **Any type of Variational State** can be represented via the NNQS.

Universal Approximation Theorem $\rightarrow \Psi_{\mathcal{W}}(\mathbf{R},\mathbf{S}) = \langle \mathrm{R},\mathrm{S} |$  $\rangle$

  - **Scalable size neural network** are used for this representation.
  - **The Ground state** $\rightarrow$ unprecedented precision with a variational approach
  - **Hidden-Nuclons Wavefunction introduce dynamical correlations** between particles and preserve statistical correlations.

The NNQS approach

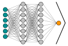
- The Neural Network Quantum States (Carleo et al. 2017) approach is a form of unsupervised learning.
  - **Any type of Variational State** can be represented via the NNQS.

Universal Approximation Theorem $\rightarrow \Psi_{\mathcal{W}}(\mathbf{R}, \mathbf{S}) = \langle \mathrm{R}, \mathrm{S} |$  $\rangle$

  - **Scalable size neural network** are used for this representation.
  - **The Ground state** $\rightarrow$ unprecedented precision with a variational approach
  - **Hidden-Nuclons Wavefunction introduce dynamical correlations** between particles and preserve statistical correlations.

The NNQS approach

- The Neural Network Quantum States (Carleo et al. 2017) approach is a form of unsupervised learning.
    - **Any type of Variational State** can be represented via the NNQS.

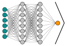
Universal Approximation Theorem → $\Psi_{\mathcal{W}}(\mathbf{R},\mathbf{S}) = \left\langle \mathrm{R},\mathrm{S} \right|$  $\left.\right\rangle$

    - **Scalable size neural network** are used for this representation.
    - **The Ground state** → unprecedented precision with a variational approach
    - **Hidden-Nuclons Wavefunction introduce dynamical correlations** between particles and preserve statistical correlations.
- Advantages:
    1. Easily applicable to different systems
        - **Approximation error is no longer Ψ-dependent** for a given interaction.

The NNQS approach

- The Neural Network Quantum States (Carleo et al. 2017) approach is a form of unsupervised learning.
  - **Any type of Variational State** can be represented via the NNQS.

$$\text{Universal Approximation Theorem} \rightarrow \Psi_{\mathcal{W}}(\mathbf{R}, \mathbf{S}) = \langle \mathrm{R}, \mathrm{S} | \ \rangle$$

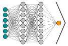


  - **Scalable size neural network** are used for this representation.
  - **The Ground state** → unprecedented precision with a variational approach
  - **Hidden-Nuclons Wavefunction introduce dynamical correlations** between particles and preserve statistical correlations.
- Advantages:
  1. Easily applicable to different systems
     - **Approximation error is no longer Ψ-dependent** for a given interaction.
  2. Less time- and computational-consuming than other methods for $A \geq 5$.
  3. Computational cost scales polinomially with $\sim \alpha A^{5 \div 6}$

The NNQS approach

- The Neural Network Quantum States (Carleo et al. 2017) approach is a form of unsupervised learning.
  - **Any type of Variational State** can be represented via the NNQS.

$$\text{Universal Approximation Theorem} \rightarrow \Psi_{\mathcal{W}}(\mathbf{R}, \mathbf{S}) = \langle \mathrm{R}, \mathrm{S} | \quad \rangle$$



  - **Scalable size neural network** are used for this representation.
  - **The Ground state** → unprecedented precision with a variational approach
  - **Hidden-Nuclons Wavefunction introduce dynamical correlations** between particles and preserve statistical correlations.
- Advantages:
  1. Easily applicable to different systems
     - **Approximation error is no longer Ψ-dependent** for a given interaction.
  2. Less time- and computational-consuming than other methods for $A \geq 5$.
  3. Computational cost scales polinomially with $\sim \alpha A^{5 \div 6}$
- Disadvantages:
  1. EFT contact potentials $\propto \delta_{\lambda}(\mathbf{x}_{ij})$ introduces **irreducible error at large cut-offs** in Monte Carlo integration due to small statistics in the interaction range →Is better to keep the cutoff small!!.

Description of the NNQS

NNQS is a Variational Approach:

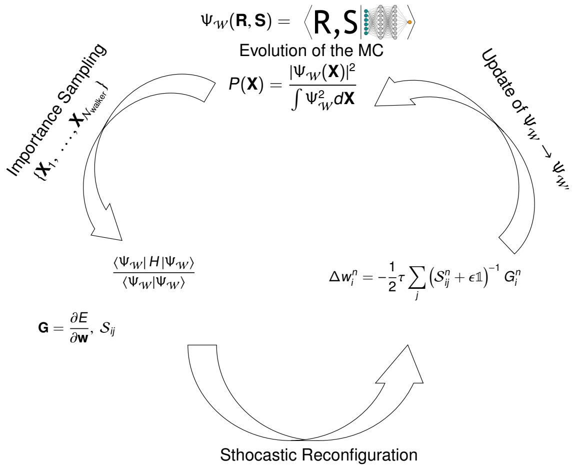
1. Apply the V.P.

$$E_V = \frac{\langle \Psi_{\mathcal{W}} | H | \Psi_{\mathcal{W}} \rangle}{\langle \Psi_{\mathcal{W}} | \Psi_{\mathcal{W}} \rangle} \geq E_0$$

2. Iterative optimization till

$$\frac{\delta E_V(\{\mathcal{W}\})}{\delta \{\mathcal{W}\}} = 0$$

- The determination of $E_V \sim E_0$ and Ψ_0:
  1. Sampling the Wavefunction using $\mathcal{N}$ Markov Chains in parallel.
  2. Evaluating observables and gradients' expectation values with Importance Sampling.
  3. Updating parameters using Stochastic Reconfiguration.



NNQS - Hidden Nucleon Wavefunction

Step 0: modellization of the wavefunction with Hidden Nucleon approach:

Fermionic wave functions from neural-network constrained hidden states - Carleo, Moreno, Georges, Stokes PNAS Vol. 119 | No. 32 (2022)

1. Fermionic systems $\rightarrow$ an anti-symmertic wavefunction (Pauli exclusion) with a **single extended Slater Determinant**.

NNQS - Hidden Nucleon Wavefunction

Step 0: modellization of the wavefunction with Hidden Nucleon approach:

Fermionic wave functions from neural-network constrained hidden states - Carleo, Moreno, Georges, Stokes PNAS Vol. 119 | No. 32 (2022)

1. Fermionic systems → an anti-symmertic wavefunction (Pauli exclusion) with a **single extended Slater Determinant**.
2. It gives a systematic and extendible approach.

NNQS - Hidden Nucleon Wavefunction

Step 0: modellization of the wavefunction with Hidden Nucleon approach:

Fermionic wave functions from neural-network constrained hidden states - Carleo, Moreno, Georges, Stokes PNAS Vol. 119 | No. 32 (2022)

1. Fermionic systems → an anti-symmertic wavefunction (Pauli exclusion) with a **single extended Slater Determinant**.
2. It gives a systematic and extendible approach.

→ **We introduce A_h "Hidden" DoFs → not real partices**:

- Represented by the **hidden orbitals** χ_i
- Each virtual particle coordinate $\tilde{x}_j = f(\{X\})$ **is a bosonic function** of all the real particles coordinates

NNQS - Hidden Nucleon Wavefunction

Step 0: modellization of the wavefunction with Hidden Nucleon approach:

Fermionic wave functions from neural-network constrained hidden states - Carleo, Moreno, Georges, Stokes PNAS Vol. 119 | No. 32 (2022)

1. Fermionic systems → an anti-symmertic wavefunction (Pauli exclusion) with a **single extended Slater Determinant**.
2. It gives a systematic and extendible approach.

→ **We introduce A_h "Hidden" DoFs → not real partices**:

- Represented by the **hidden orbitals** χ_i
- Each virtual particle coordinate $\tilde{x}_j = f(\{X\})$ **is a bosonic function** of all the real particles coordinates

$$\det \begin{vmatrix} \phi_v(\mathbf{X}) & \phi_v(f(\{\mathbf{X}\})) \\ \chi_h(\mathbf{X}) & \chi_h(f(\{\mathbf{X}\})) \end{vmatrix} = \det \begin{bmatrix} \phi_1(x_1) & \phi_1(x_2) & \phi_1(f(\{X\})) \\ \phi_2(x_1) & \phi_2(x_2) & \phi_2(f(\{X\})) \\ \chi_3(x_1) & \chi_3(x_2) & \chi_3(f(\{X\})) \end{bmatrix}$$

Visible orbitals on visible coordinates | Visible orbitals on hidden coordinates

Hidden orbitals on visible coordinates | Hidden orbitals on hidden coordinates

What about statistical correlations?

NNQS - Hidden Nucleon Wavefunction

$$\phi_1(f(x)) \cdot \left(\phi_2(x_1)\chi_3(x_2) - \chi_3(x_1)\phi_2(x_2)\right)$$
$$- \phi_2(f(x)) \cdot \left(\phi_1(x_1)\chi_3(x_2) - \chi_3(x_1)\phi_1(x_2)\right)$$
$$+ \chi_3(f(x)) \cdot \left(\phi_1(x_1)\phi_2(x_2) - \phi_2(x_1)\phi_1(x_2)\right)$$

NNQS - Hidden Nucleon Wavefunction

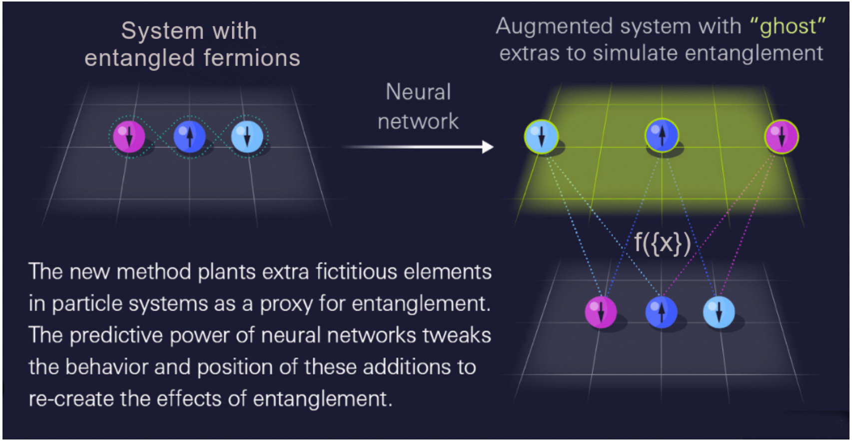
$$
\begin{aligned}
\phi_1(f(x)) \cdot \left(\phi_2(x_1)\chi_3(x_2) - \chi_3(x_1)\phi_2(x_2)\right) \\
- \phi_2(f(x)) \cdot \left(\phi_1(x_1)\chi_3(x_2) - \chi_3(x_1)\phi_1(x_2)\right) \\
+ \chi_3(f(x)) \cdot \left(\phi_1(x_1)\phi_2(x_2) - \phi_2(x_1)\phi_1(x_2)\right)
\end{aligned}
$$

it is an equivalent form of the C.I. expansion with only one excited states. (but it's more than this!)

$$
|\Phi\rangle = C_0|\Psi_0\rangle + \sum_{ra} C_a^r |\Psi_a^r\rangle + \sum_{\substack{a<b \\ r<s}} C_{ab}^{rs} |\Psi_{ab}^{rs}\rangle + \sum_{\substack{a<b<c \\ r<s<t}} C_{abc}^{rst} |\Psi_{abc}^{rst}\rangle + \dots \tag{1}
$$

NNQS - Hidden Nucleon Wavefunction

- Introduce dynamical correlations between particles by means of functions of bosonic nature $f(\{\mathbf{x}\})$ which accounts for Jastrow correlators for all the particles in the system:



HN wavefunction in the Visible coordinate sector

- The hidden coordinate sector is represented through a phase - amplitude modulation of the wave function

$$\det\left[\begin{array}{cc|c} \phi_1(x_1) & \phi_1(x_2) & \phi_1(f(\{X\})) \\ \phi_2(x_1) & \phi_2(x_2) & \phi_2(f(\{X\})) \\ \hline \chi_3(x_1) & \chi_3(x_2) & \chi_3(f(\{X\})) \end{array}\right]$$

$$\phi_\alpha(x_i) = e^{u^\alpha_\phi(x_i)} \tanh[v^\alpha_\phi(x_i)] \sim e^{u^\alpha_\phi(x_i) + j \cdot v^\alpha_\phi(x_i)}$$

$$\chi_\alpha(x_i) = e^{u^\alpha_\chi(x_i)} \tanh[v^\alpha_\chi(x_i)] \sim e^{u^\alpha_\chi(x_i) + j \cdot v^\alpha_\chi(x_i)}$$

HN wavefunction in the Visible coordinate sector

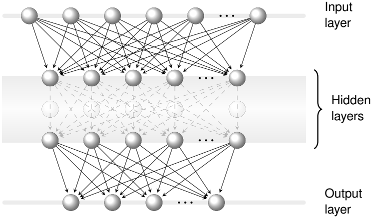
- The hidden coordinate sector is represented through a phase - amplitude modulation of the wave function

$$\det \left[\begin{array}{cc|c} \phi_1(x_1) & \phi_1(x_2) & \phi_1(f(\{X\})) \\ \phi_2(x_1) & \phi_2(x_2) & \phi_2(f(\{X\})) \\ \hline \chi_3(x_1) & \chi_3(x_2) & \chi_3(f(\{X\})) \end{array} \right]$$

$$\phi_\alpha(x_i) = e^{u^\alpha_\phi(x_i)} \tanh[v^\alpha_\phi(x_i)] \sim e^{u^\alpha_\phi(x_i) + j\cdot v^\alpha_\phi(x_i)}$$

$$\chi_\alpha(x_i) = e^{u^\alpha_\chi(x_i)} \tanh[v^\alpha_\chi(x_i)] \sim e^{u^\alpha_\chi(x_i) + j\cdot v^\alpha_\chi(x_i)}$$

- u and v are both FF Neural Networks with only one hidden layer (Let's call this a "standard network" for this work.).



A standard Single Particle FFNN has

- Input Nodes = 5 $\rightarrow [\mathbb{R}^3, s_z, t_z]$
- Output Nodes = 1 $\rightarrow \phi(x_i)$

NNQS Hidden coordinate sector - Deep Sets Architecture

- The hidden coordinate sector is represented through a phase - amplitude modulation of the wave function

$$\det\left[\begin{array}{cc|c} \phi_1(x_1) & \phi_1(x_2) & \phi_1(f(\{X\})) \\ \phi_2(x_1) & \phi_2(x_2) & \phi_2(f(\{X\})) \\ \hline \chi_3(x_1) & \chi_3(x_2) & \chi_3(f(\{X\})) \end{array}\right]$$

$$\phi_i(f(\{\mathbf{X}\})) = e^{\mathcal{U}^i_\phi(\{\mathbf{X}\})} \tanh[\mathcal{V}^i_\phi(\{\mathbf{X}\})] \sim e^{\mathcal{U}^i_\phi(\{\mathbf{X}\}) + j \cdot \mathcal{V}^i_\phi(\{\mathbf{X}\})}$$

$$\chi_i(f(\{\mathbf{X}\})) = e^{\mathcal{U}^i_\chi(\{\mathbf{X}\})} \tanh[\mathcal{V}^i_\chi(\{\mathbf{X}\})] \sim e^{\mathcal{U}^i_\chi(\{\mathbf{X}\}) + j \cdot \mathcal{V}^i_\chi(\{\mathbf{X}\})}$$

NNQS Hidden coordinate sector - Deep Sets Architecture

- The hidden coordinate sector is represented through a phase - amplitude modulation of the wave function

$$\det\begin{bmatrix} \phi_1(x_1) & \phi_1(x_2) & \phi_1(f(\{X\})) \\ \phi_2(x_1) & \phi_2(x_2) & \phi_2(f(\{X\})) \\ \hline \chi_3(x_1) & \chi_3(x_2) & \chi_3(f(\{X\})) \end{bmatrix}$$

$$\phi_i(f(\{\mathbf{X}\})) = e^{\mathcal{U}^i_\phi(\{\mathbf{X}\})}\tanh[\mathcal{V}^i_\phi(\{\mathbf{X}\})] \sim e^{\mathcal{U}^i_\phi(\{\mathbf{X}\})+j\cdot\mathcal{V}^i_\phi(\{\mathbf{X}\})}$$

$$\chi_i(f(\{\mathbf{X}\})) = e^{\mathcal{U}^i_\chi(\{\mathbf{X}\})}\tanh[\mathcal{V}^i_\chi(\{\mathbf{X}\})] \sim e^{\mathcal{U}^i_\chi(\{\mathbf{X}\})+j\cdot\mathcal{V}^i_\chi(\{\mathbf{X}\})}$$

- **Deep Sets architecture:** any permutation invariant neural network ($\mathcal{U}$ and $\mathcal{V}$) can be sum-decomposed in the following way:

$$\mathcal{F}(\{\mathbf{X}\}) = \rho_{\mathcal{F}}\Big[\sum_{i\neq j}\phi_{\mathcal{F}}([\mathbf{x}_i,\mathbf{x}_j])\Big] \qquad \mathcal{F} = \mathcal{U},\mathcal{V}$$

- ϕ and ρ are both Neural Network, the sum operation destroy the order dependence of the network's inputs.

NNQS Hidden coordinate sector - Deep Sets Architecture

- The hidden coordinate sector is represented through a phase - amplitude modulation of the wave function

$$\det\left[\begin{array}{cc|c} \phi_1(x_1) & \phi_1(x_2) & \phi_1(f(\{X\})) \\ \phi_2(x_1) & \phi_2(x_2) & \phi_2(f(\{X\})) \\ \hline \chi_3(x_1) & \chi_3(x_2) & \chi_3(f(\{X\})) \end{array}\right]$$

$$\phi_i(f(\{\mathbf{X}\})) = e^{\mathcal{U}^i_\phi(\{\mathbf{X}\})}\tanh[\mathcal{V}^i_\phi(\{\mathbf{X}\})] \sim e^{\mathcal{U}^i_\phi(\{\mathbf{X}\})+j\cdot\mathcal{V}^i_\phi(\{\mathbf{X}\})}$$

$$\chi_i(f(\{\mathbf{X}\})) = e^{\mathcal{U}^i_\chi(\{\mathbf{X}\})}\tanh[\mathcal{V}^i_\chi(\{\mathbf{X}\})] \sim e^{\mathcal{U}^i_\chi(\{\mathbf{X}\})+j\cdot\mathcal{V}^i_\chi(\{\mathbf{X}\})}$$

- **Deep Sets architecture:** any permutation invariant neural network ($\mathcal{U}$ and $\mathcal{V}$) can be sum-decomposed in the following way:

$$\mathcal{F}(\{\mathbf{X}\}) = \rho_{\mathcal{F}}\left[\sum_{i\neq j}\phi_{\mathcal{F}}([\mathbf{x}_i,\mathbf{x}_j])\right] \qquad \mathcal{F} = \mathcal{U},\mathcal{V}$$

- ϕ and ρ are both Neural Network, the sum operation destroy the order dependence of the network's inputs.
- This decomposition is called **Sum-Pooling**, since the aggregation function is $g(\cdot) = \sum_{x\in\mathfrak{X}}$.

Application of NNQS to hypernuclear variational states

- Variational state of a hypernucleus (only one Λ):

$$\Psi_{HN} = \phi_\Lambda(\mathbf{x}_\Lambda, g(\{\mathbf{x}_1, \ldots, \mathbf{x}_{A-1}, \mathbf{x}_\Lambda\})) \cdot \det_{A-1_Z} \begin{vmatrix} \phi_v(\mathbf{X}) & \phi_v(f(\{\mathbf{X}\})) \\ \chi_h(\mathbf{X}) & \chi_h(f(\{\mathbf{X}\})) \end{vmatrix}$$

Application of NNQS to hypernuclear variational states

- Variational state of a hypernucleus (only one Λ):

$$\Psi_{HN} = \phi_\Lambda(\mathbf{x}_\Lambda, g(\{\mathbf{x}_1, \dots, \mathbf{x}_{A-1}, \mathbf{x}_\Lambda\})) \cdot \det_{A-1 Z} \begin{vmatrix} \phi_v(\mathbf{X}) & \phi_v(f(\{\mathbf{X}\})) \\ \chi_h(\mathbf{X}) & \chi_h(f(\{\mathbf{X}\})) \end{vmatrix}$$

- In order to represent ΛN correlations related with the ${}^3_\Lambda H$ mixed asymmetry spin state, the aggregator function for non-identical particles has to be modified

$$g(\mathcal{S}[\mathbf{x}_1, \dots, \mathbf{x}_{A-1}], \mathbf{x}_\Lambda) = \left[\mathbf{x}_\Lambda, \sum_{i=1}^{A-1} \phi(\mathbf{r}_i, \mathbf{s}_i)\right]$$

The coordinate of the Lambda are concatenated with the sum pooling → the ΛN permutation invariance is not introduced

Application of NNQS to hypernuclear variational states

- Variational state of a hypernucleus (only one Λ):

$$\Psi_{HN} = \phi_\Lambda(\mathbf{x}_\Lambda, g(\{\mathbf{x}_1, \dots, \mathbf{x}_{A-1}, \mathbf{x}_\Lambda\})) \cdot \det_{A-1Z} \begin{vmatrix} \phi_v(\mathbf{X}) & \phi_v(f(\{\mathbf{X}\})) \\ \chi_h(\mathbf{X}) & \chi_h(f(\{\mathbf{X}\})) \end{vmatrix}$$

- In order to represent ΛN correlations related with the ${}^3_\Lambda H$ mixed asymmetry spin state, the aggregator function for non-identical particles has to be modified

$$g(\mathcal{S}[\mathbf{x}_1, \dots, \mathbf{x}_{A-1}], \mathbf{x}_\Lambda) = \left[\mathbf{x}_\Lambda, \sum_{i=1}^{A-1} \phi(\mathbf{r}_i, \mathbf{s}_i)\right]$$

The coordinate of the Lambda are concatenated with the sum pooling → the ΛN permutation invariance is not introduced

- ϕ_Λ has already mentioned phase-amplitude modulation

$$\phi_\Lambda\Big(g(\mathcal{S}(\{\mathbf{x}_i\}), \mathbf{x}_\Lambda)\Big) = \tanh\left[\rho_p\Big(g(\mathcal{S}(\{\mathbf{x}_i\}), \mathbf{x}_\Lambda)\Big)\right] \cdot \exp\left[\rho_a\Big(g(\mathcal{S}(\{\mathbf{x}_i\}), \mathbf{x}_\Lambda)\Big)\right]$$

Pre-processing (backflow) of coordinates via MPNN

What if we want to use one only one hidden state state?

- We can introduce the following backflow transformation for the single particle's coordinates:

Pre-processing (backflow) of coordinates via MPNN

What if we want to use one only one hidden state state?

- We can introduce the following backflow transformation for the single particle's coordinates:

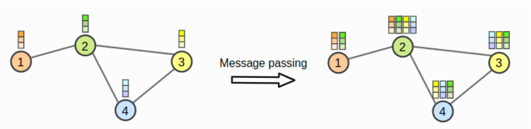
The backflow transformation modifies the coordinates of the particles so that the "effective" position of a particle is a function not only of its "bare" position but also of the positions of the other particles.

Pre-processing (backflow) of coordinates via MPNN

What if we want to use one only one hidden state state?

- We can introduce the following backflow transformation for the single particle's coordinates:

The backflow transformation modifies the coordinates of the particles so that the "effective" position of a particle is a function not only of its "bare" position but also of the positions of the other particles.



In the NN language it consist in the application of a minimal MPNN:
an all-toall connected graph, encoding effective particle positions (x_i-nodes) and their interactions (m_{ij}-edges)

Backflow Λ-coordinate

$$m_{Lj} = \phi([x_{\text{lam}}, x_j])$$
$$m_L = \frac{1}{N_{\text{part}}} \sum_{j=1}^{N_{\text{part}}} m_{Lj}$$
$$\mathbf{X}_L = [x_{lam}, m_L]$$

Backflow Ns coordinates

$$m_{ij} = \chi([x_i, x_j])$$
$$m_i = \frac{1}{N_{\text{part}}} \sum_{j=1}^{N_{\text{part}}} m_{ij}$$
$$\mathbf{X}_i = [x_i, m_i]$$

→ MPNN NQS for the Homogeneous Electron Gas - Lovato, Carleo, Kim, Pescia, Nys (2023) arXiv:2305.07240v3

Hidden Nucleon generalization for $n - \Lambda$-hypernuclear states

Is possible to include more than one Λs?

Hidden Nucleon generalization for $n - \Lambda$-hypernuclear states

Is possible to include more than one Λs?
Putting all the previous ideas together...

$$\Psi_{HN} = \det_{\Lambda} \begin{vmatrix} \phi_{\text{v}}(x_{lam}) & \phi_{\text{v}}(f(\{\mathbf{X}_L\})) \\ \chi_{\text{h}}(x_{lam}) & \chi_{\text{h}}(f(\{\mathbf{X}_L\})) \end{vmatrix} \cdot \det_{^{A-1}Z} \begin{vmatrix} \phi_{\text{v}}(\mathbf{X}_i) & \phi_{\text{v}}(f(\{\mathbf{X}_i\})) \\ \chi_{\text{h}}(\mathbf{X}_i) & \chi_{\text{h}}(f(\{\mathbf{X}_i\})) \end{vmatrix}$$

Observables evaluation

Step 1: Importance Sampling and Metropolis

- Expectation values are computed through **importance sampling**. By rewriting the multidimensional integral for a generic observable $\mathcal{O}$:

$$\mathcal{O}_V = \frac{\langle\Psi|\mathcal{O}|\Psi\rangle}{\langle\Psi|\Psi\rangle} = \frac{\int d\mathbf{R}\,\langle\Psi|\mathbf{R},\mathbf{S}\rangle\langle\mathbf{R},\mathbf{S}|\mathcal{O}|\Psi\rangle\frac{\langle\mathbf{R},\mathbf{S}|\Psi\rangle}{\langle\mathbf{R},\mathbf{S}|\Psi\rangle}}{\int d\mathbf{R}\,\langle\Psi|\mathbf{R},\mathbf{S}\rangle\langle\mathbf{R},\mathbf{S}|\Psi\rangle} =$$

$$\overset{\mathbf{X}=\{\mathbf{R},\mathbf{S}\}}{=} \frac{\int d\mathbf{R}|\Psi(\mathbf{X})|^2\mathcal{O}_L(\mathbf{X})}{\int d\mathbf{R}|\Psi(\mathbf{X})|^2} = \int d\mathbf{R}\,P(\mathbf{X})\mathcal{O}_L(\mathbf{X})$$

with $\mathcal{O}_L(\mathcal{O}) = \frac{\mathcal{O}\Psi(\mathbf{X})}{\Psi(\mathbf{X})}$ and $P(\mathbf{X}) = \frac{|\Psi(\mathbf{X})|^2}{\int d\mathbf{R}|\Psi(\mathbf{X})|^2} \rightarrow \{\mathbf{X}_1,\ldots,\mathbf{X}_{\mathcal{N}_{\text{walker}}}\}$

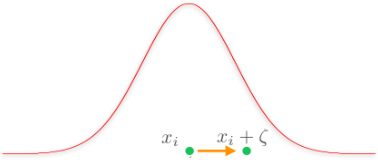
$$\mathcal{O}_V = \frac{1}{\mathcal{N}_{\text{walker}}}\sum_{s=1}^{\mathcal{N}_{\text{walker}}}\mathcal{O}_L(\mathbf{X}_s) \qquad \sigma_{\mathcal{O}_V} = \sqrt{\frac{1}{\mathcal{N}_{\text{walker}}-1}\sum_{s=1}^{\mathcal{N}_{\text{walker}}}\left(\mathcal{O}_L(\mathbf{X}_s)-\mathcal{O}_V\right)^2}$$

Metropolis Hastings Algorithm

Step 2: MH Our goal is to sample the probability distribution described by $P(\mathbf{X}) = \frac{|\Psi_T(\mathbf{X})|^2}{\int d\mathbf{R}|\Psi(\mathbf{X})|^2}$.
The $M(RT)^2$ algorithm is based on the idea of random walk: **a Markov Chain with transition matrix Π.**

$$\mathcal{P}(X_{i+1} = x_{i+1}|X_0 = x_0, \cdots, X_i = x_i) = \Pi(x_i, x_{i+1}) = q(x_{i+1}|x_i)r(x_{i+1}|x_i)$$

- $x_{i+1} = x_i + \zeta_{N(\mu=0,\sigma=1)}$
- $r(x_{i+1}|x_i) = \min\left(1, \frac{P(x_{i+1})\cancel{q(x_i|x_{i+1})}}{P(x_i)\cancel{q(x_{i+1}|x_i)}}\right)$

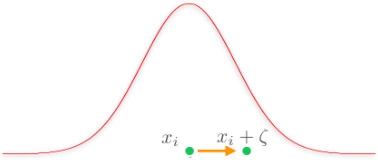


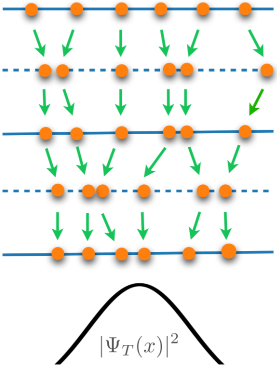
Metropolis Hastings Algorithm

Step 2: MH Our goal is to sample the probability distribution described by $P(\mathbf{X}) = \frac{|\Psi_T(\mathbf{X})|^2}{\int d\mathbf{R}|\Psi(\mathbf{X})|^2}$.
The $M(RT)^2$ algorithm is based on the idea of random walk: **a Markov Chain with transition matrix Π.**

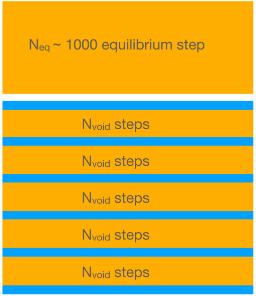
$$\mathcal{P}(X_{i+1} = x_{i+1}|X_0 = x_0, \cdots, X_i = x_i) = \Pi(x_i, x_{i+1}) = q(x_{i+1}|x_i)r(x_{i+1}|x_i)$$

- $x_{i+1} = x_i + \zeta_{N(\mu=0,\sigma=1)}$
- $r(x_{i+1}|x_i) = \min\left(1, \frac{P(x_{i+1})\cancel{q(x_i|x_{i+1})}}{P(x_i)\cancel{q(x_{i+1}|x_i)}}\right)$





1. Wait for thermalization
2. Evaluate $\mathcal{N}_{avg}$ spaced by $\mathcal{N}_{void}$ step.



The Stochastic Reconfiguration algorithm

Step 3: Optimization of the Parameters:

- Stochastic-Reconfiguration (S. Sorella 2005), in the context of VMC, is equivalent to performing **imaginary-time evolution in the parameters space**, and it is related to the Natural Gradient descent method (Amari et al.).

The Stochastic Reconfiguration algorithm

Step 3: Optimization of the Parameters:

- Stochastic-Reconfiguration (S. Sorella 2005), in the context of VMC, is equivalent to performing **imaginary-time evolution in the parameters space**, and it is related to the Natural Gradient descent method (Amari et al.).
- Imaginary-time evolution $e^{iH(i\tau)} \approx (1 - H\tau)$ in the parameter space reads

$$(1 - H\tau)\left|\Psi_T(\mathcal{W})\right\rangle = \left|\Psi_T(\mathcal{W} + \Delta\mathcal{W})\right\rangle = \Delta\mathcal{W}_0\left|\Psi_T(\mathcal{W})\right\rangle + \sum_j \Delta\mathcal{W}_j\mathcal{O}^j\left|\Psi_T(\mathcal{W})\right\rangle$$

where $\mathcal{O}^i\left|\Psi_T(\mathcal{W})\right\rangle = \left|\frac{\partial}{\partial \mathcal{W}_i}\Psi(\mathcal{W})\right\rangle$. Multiplying from left by

$$\frac{\left\langle\Psi(\mathcal{W})\right|}{\left\langle\Psi_T(\mathcal{W})\middle|\Psi_T(\mathcal{W})\right\rangle} \quad \text{and} \quad \frac{\left\langle\Psi(\mathcal{W})\right|\mathcal{O}^i}{\left\langle\Psi_T(\mathcal{W})\middle|\Psi_T(\mathcal{W})\right\rangle}$$

gives the two following equations

$$\begin{cases}\langle(1 - H\tau)\rangle = \Delta\mathcal{W}_0 + \sum_i \Delta\mathcal{W}_i\langle\mathcal{O}^i\rangle \\ \langle\mathcal{O}^i(1 - H\tau)\rangle = \Delta\mathcal{W}_0\langle\mathcal{O}^i\rangle + \sum_j \Delta\mathcal{W}_j\langle\mathcal{O}^i\mathcal{O}^j\rangle\end{cases}$$

The Stochastic Reconfiguration Algorithm

- solving the first for $\Delta\mathcal{W}_0$ and exploiting it in the second gives

$$\underbrace{\Big(\langle H\rangle\langle O^i\rangle - \langle HO^i\rangle\Big)}_{-\frac{1}{2}G_i}\tau = \sum_j \Delta\mathcal{W}_j \underbrace{\Big(\langle O^i O^j\rangle - \langle O^i\rangle\langle O^j\rangle\Big)}_{\mathcal{S}_{ij}}$$

It naturally leads to the updating rule

$$\mathcal{W}_i^{n+1} = \mathcal{W}_i^n + \Delta\mathcal{W}_i^n = \mathcal{W}_i^n - \frac{1}{2}\tau\sum_j\Big(\underbrace{\mathcal{S}_{ij}^n + \epsilon\mathbb{1}_{ij}}_{\text{avoid saddle points}}\Big)^{-1} G_j^n$$

where the quantities of interests are all evaluated through importance sampling

$$G_i = \frac{\partial E(\mathcal{W})}{\partial \mathcal{W}_i} = 2\Big(\langle O_i H\rangle - E_V\langle O_i\rangle\Big) \qquad \mathcal{S}_{ij} = \langle O_i O_j\rangle - \langle O_i\rangle\langle O_j\rangle$$

The QGT looks like a variance..

In classical Information Theory the Riemannian structure of the parameter space $\mathcal{P}$ of a statistical model $P(\mathbf{x}|\mathbf{w})$, which depends on some parameters $\mathbf{w}$, is defined by the **Fisher information**:

$$g_{ij}(\mathbf{w}) = \mathbb{E}_{p(\mathbf{x}|\mathbf{w})}\left[\frac{\partial \log(p(\mathbf{x}|\mathbf{w}))}{\partial w_i}\frac{\partial \log(p(\mathbf{x}|\mathbf{w}))}{\partial w_j}\bigg|\mathbf{w}\right] = -\mathbb{E}_{p(\mathbf{x}|\mathbf{w})}\left[\frac{\partial^2}{\partial w_i \partial w_j}\log\left(p(\mathbf{x}|\mathbf{w})\right)\bigg|\mathbf{w}\right]$$

Fisher Information is:

1. **The Variance of the derivative of the log-likelihood (Score Function)**
2. **The curvature of the log-likelihood**

Peaked Loglikelihood $= \partial^2_{\mathbf{w_i}}$ high curvature $=$ sample brings high information about $\mathbf{w}_i$

we suppress the variation of such a parameter $G_i^n \to \sum_j (\mathcal{S}_{ij}^n)^{-1} G_j^n$

In Quantum Information Theory:

$$d_{FS}(\Psi,\Phi) = \arccos\left(|\langle \Psi_{\text{norm}}, \Phi_{\text{norm}}\rangle|^{\frac{1}{2}}\right) \to d_{FS}[\Psi_V(\mathbf{X}|\mathbf{w}), \Psi_V(\mathbf{X}|\mathbf{w}+d\mathbf{w})]^2 = \sum_{ij} \mathcal{S}_{ij}(\mathbf{w}) dw_i dw_j$$

If Ψ is defined over a basis $|\psi(\mathbf{X}|\mathbf{w})\rangle \approx \sum_{\mathbf{X}\in\{\mathbf{X}\}} \sqrt{p(\mathbf{X}|\mathbf{w})}\,|\mathbf{X}\rangle$ can be shown that $\mathcal{S}_{ij}(\mathbf{w}) = \frac{1}{4} g_{ij}(\mathbf{w})$.

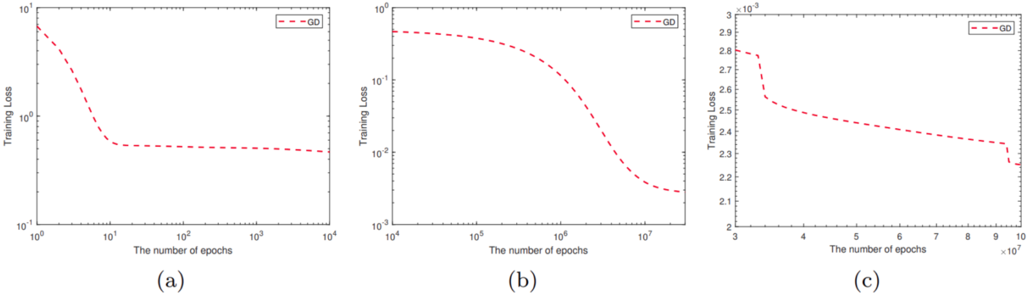
Quantum Natural Gradient - Carleo et al. (Quantum May 2020), page 10.

The Natural Gradient Descent

Amari (S. I. Amari, Neural Computation 10, 251 (1998).) → steepest descent direction of a cost function $L(w)$ in a Riemannian space is given by the **Natural Gradient**.

$$-\tilde{\nabla} L(\mathbf{w}) = -g^{-1}(\mathbf{w})\nabla L(\mathbf{w})$$

- The NG flatten the metric of the parameter space!
- Using the Natural Gradient the plateau phenomenon might disappear.



SR with RMSProp

- In the RMSProp the gradient express the **acceleration** in the parameter space. The $\mathcal{S}$ matrix is regularized with the running averages of the squared gradients.

$$\mathbf{m}^{n+1} = \beta \mathbf{m}^{n} + (1-\beta)\mathbf{G}^{n} \odot \mathbf{G}^{n}$$

$$\mathcal{S}_{ij}^{n} + \epsilon \cdot \text{diag}(1) \rightarrow \mathcal{S}_{ij}^{n} + \epsilon \cdot \text{diag}(\sqrt{\mathbf{m}^{n}} + 10^{-8})$$

- The ϵ hyperparameter add an L^2 penality term to the solution of the system from which $\Delta\mathcal{W}$ are determined $\rightarrow ||\Delta\mathcal{W}\mathcal{S} - \mathbf{G}|| = ||\epsilon\Delta\mathcal{W}||$

Now let's see some results..

Interaction Potential:

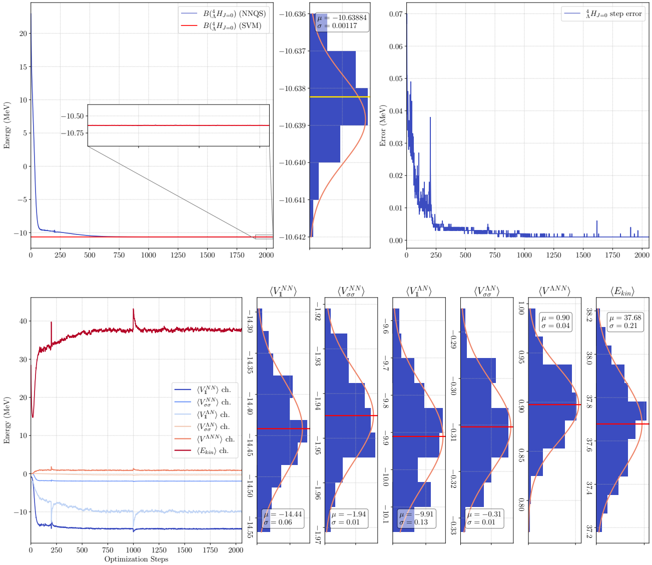
Idea of the improved LO interaction → Two- and three-nucleon contact interactions and ground-state energies of light- and medium-mass nuclei, Phys. Rev. C 103, 054003 2021, Schiavilla et al. - model "o"

- Coulomb interaction derived from the same paper.

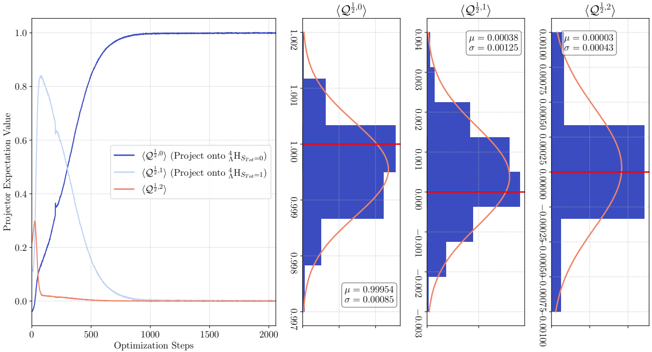
Gaussian regulators: $\delta_\lambda(x) = \frac{\Lambda^3}{8\pi^{\frac{3}{2}}} e^{-\frac{\Lambda^2 r^2}{4}}$

- NN and NNN interaction:
  1. NN fitted to pn/nn: $a_0^{s,t}$, and $r_0^{s,t}$ with variable phase method.
  2. NNN fitted to ^{3}H and ^{4}He BE with Gaussian Processes
- ΛN and ΛNN interaction:
  1. ΛN fitted to pΛ $a_0^{s,t}$, and $r_0^{s,t}$.
  2. ΛNN fitted to the $^3_\Lambda$H, $^4_\Lambda$H$_{S_{tot}=0}$, $^4_\Lambda$H$_{S_{tot}=1}$, and $^5_\Lambda$He BE.

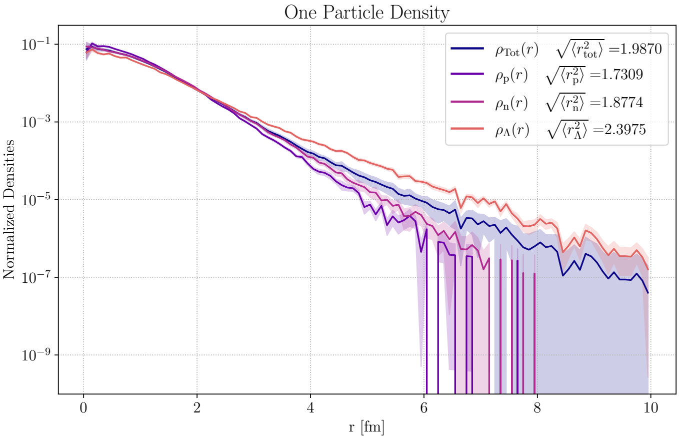
${}^{4}_{\Lambda}H_{S_{tot}=0}$ with Fitted Potential



${}^{4}_{\Lambda}H_{S_{tot}=0}$ Projectors

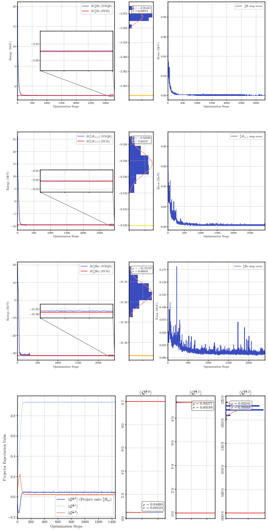


${}^{4}_{\Lambda}H_{S_{tot}=0}$ Density

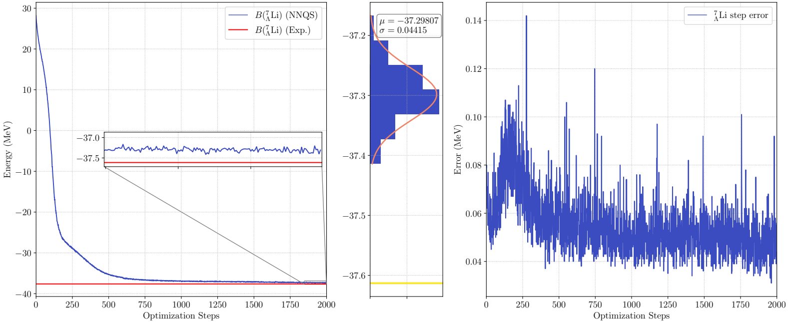


Work in progress:

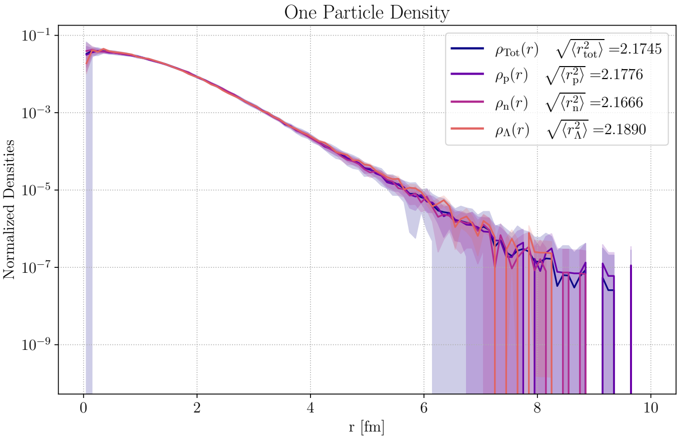
- In several systems the total spin is not a good quantum number.
- The main effect is a fluctuation of the energy around the convergence region for the target energies of the fit with an average accuracy of $5 \div 60 KeV$.



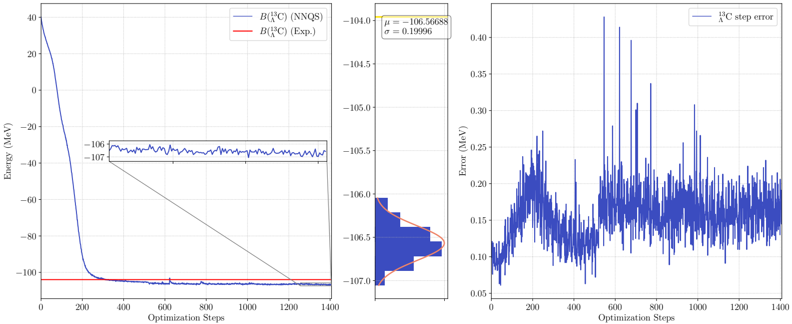
$^{7}_{\Lambda}Li$ with Fitted Potential



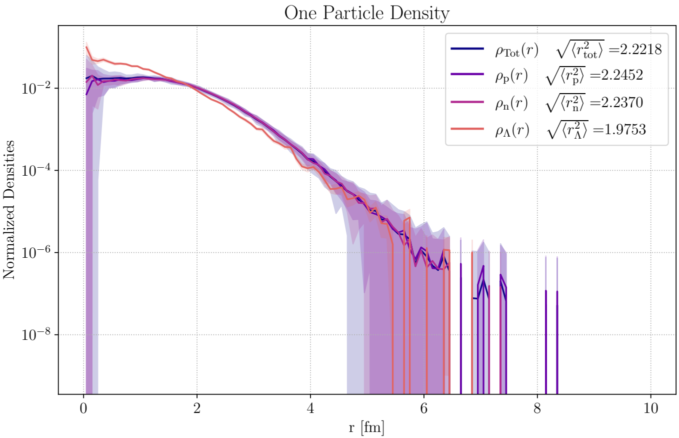
$^{7}_{\Lambda}Li$ with Fitted Potential



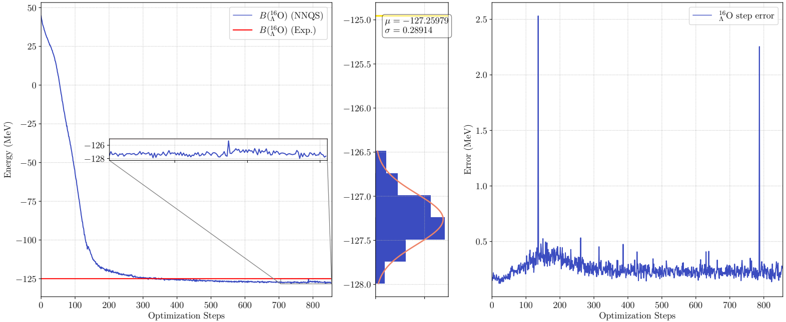
$^{13}_{\Lambda}C$ with Fitted Potential



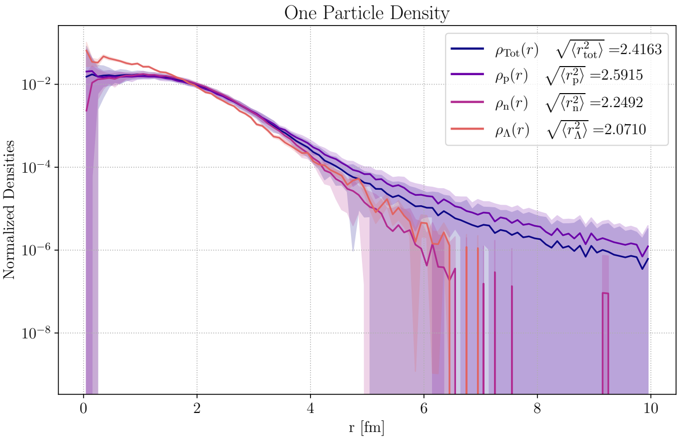
$^{13}_{\Lambda}C$ with Fitted Potential

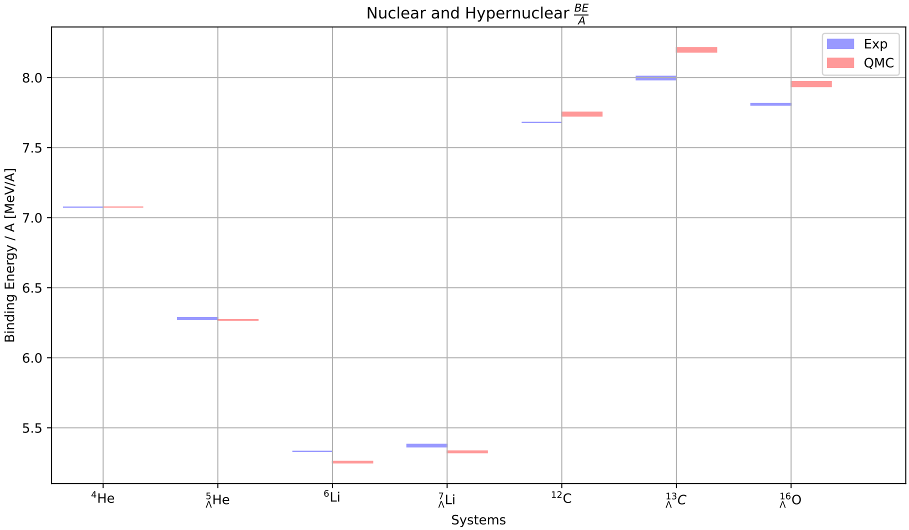


$^{16}_{\Lambda}O$ with Fitted Potential



$^{16}_{\Lambda}O$ with Fitted Potential




Nuclear and Hypernuclear $\frac{BE}{A}$
Exp
QMC
Binding Energy / A [MeV/A]
8.0
7.5
7.0
6.5
6.0
5.5
^{4}He
$^{5}_{\Lambda}$He
^{6}Li
$^{7}_{\Lambda}$Li
^{12}C
$^{13}_{\Lambda}C$
$^{16}_{\Lambda}$O
Systems

Conclusions

- NNQS shows great flexibility in modeling hypernuclear bound states.
  - NQS Variational State → match with accuracy the binding energy of hypernuclear systems.
  - Fitted interaction: Improved LO guarantes the desired accuracy.
  - Still improvements needs to be done: Obtain a wavefunction with good quantum numbers for the total Spin.
- **Outlook**:
  1. Moving to NLO and restore cutoff dependence.
  2. Extension of calculations to larger mass hypernuclei.
  3. Hypernuclear matter → EoS

Conclusions

- NNQS shows great flexibility in modeling hypernuclear bound states.
  - NQS Variational State → match with accuracy the binding energy of hypernuclear systems.
  - Fitted interaction: Improved LO guarantes the desired accuracy.
  - Still improvements needs to be done: Obtain a wavefunction with good quantum numbers for the total Spin.
- **Outlook**:
  1. Moving to NLO and restore cutoff dependence.
  2. Extension of calculations to larger mass hypernuclei.
  3. Hypernuclear matter → EoS
- **Question**: Which hypernuclear systems might be more interesting to analise?

Conclusions

- NNQS shows great flexibility in modeling hypernuclear bound states.
  - NQS Variational State → match with accuracy the binding energy of hypernuclear systems.
  - Fitted interaction: Improved LO guarantes the desired accuracy.
  - Still improvements needs to be done: Obtain a wavefunction with good quantum numbers for the total Spin.
- **Outlook**:
  1. Moving to NLO and restore cutoff dependence.
  2. Extension of calculations to larger mass hypernuclei.
  3. Hypernuclear matter → EoS
- **Question**: Which hypernuclear systems might be more interesting to analise?

Thank you!

Neural-Network Quantum States

The Neural Network Quantum States (NNQS) approach can be classified as a form of **unsupervised machine learning** but actually occupy a middle ground between supervised and unsupervised learning, as it incorporates elements from both approaches:

Approach	**Input**	**Learning Procedure**	**Output**
Supervised Learning	Labeled input-output pairs	Minimize the cost function through gradient-descent → *Backpropagation*	Model that map inputs to corresponding output labels
NNQS	Quantum system coordinates Unlabeled $(\mathbf{x}, \mathbf{s}, \mathbf{t})$	Variational optimization (Energy-gradient informed cost function → Backprop.)	Amplitude of the ground state wavefunction
Unsupervised Learning	Unlabeled data	Training with unlabeled data (Clustering- Dimensionality red. ...)	Model that represent patterns or structure in the data

Similarly to Supervised Learning (GD, SGD, RMSProp, Momentum, AdaGrad...) in NNQS we use first- and second-order derivatives information (gradient, hessian) to improve convergence and performance.

The Quantum Geometric Tensor

The Riemannian structure of the parameter space $\mathcal{P}$ of a statistical model $p(\mathbf{x}|\mathbf{w})$, which depends on some parameters $\mathbf{w}$ is defined by the Fisher information:

$$g_{ij}(\mathbf{w}) = \mathbb{E}_{p(\mathbf{x}|\mathbf{w})}\left[\frac{\partial \log(p(\mathbf{x}|\mathbf{w}))}{\partial w_i}\frac{\partial \log(p(\mathbf{x}|\mathbf{w}))}{\partial w_j}\bigg|\mathbf{w}\right] = -\mathbb{E}_{p(\mathbf{x}|\mathbf{w})}\left[\frac{\partial^2}{\partial w_i \partial w_j}\log\left(p(\mathbf{x}|\mathbf{w})\right)\bigg|\mathbf{w}\right]$$

An intuition about this property of the Fisher information in clarified by noting that the Kullback-Liebler divergence

$$D_{KL}(p(\mathbf{x},\mathbf{w}), p(\mathbf{x},\mathbf{w}')) = \int d\mathbf{x} p(\mathbf{x},\mathbf{w}) \log\left(\frac{p(\mathbf{x},\mathbf{w})}{p(\mathbf{x},\mathbf{w}')}\right)$$

evaluated between $p(\mathbf{w})$ and $p(\mathbf{w}')$, with $\mathbf{w}' \to \mathbf{w}$, is Taylor expanded to

$$\begin{aligned} D_{KL}\big(p(\mathbf{x}|\mathbf{w}), p(\mathbf{x}|\mathbf{w}')\big) &\approx \frac{1}{2}\sum_{i,j}\left(\frac{\partial^2}{\partial w'_i \partial w'_j} D_{KL}\big(p(\mathbf{x}|\mathbf{w}), p(\mathbf{x}|\mathbf{w}')\big)\right) dw_i dw_j \\ &= -\frac{1}{2}\sum_{i,j}\int d\mathbf{x} p(\mathbf{x}|\mathbf{w})\left[\frac{\partial^2}{\partial w'_i \partial w'_j}\log\left(p(\mathbf{x}|\mathbf{w}')\right)\right]_{\mathbf{w}'=\mathbf{w}} dw_i dw_j \\ &= \frac{1}{2}\sum_{i,j} g_{ij}(\mathbf{w}) dw_i dw_j \end{aligned}$$

The Quantum Geometric Tensor

We can observe the relation between the Quantum Geometric Tensor $\mathcal{S}_{ij}$ and the fisher information by means of the Fubini-Study metric, which is the natural metric for the two-level pure quantum-machanical system defined by the Block Sphere

$$d_{FS}(\Psi, \Phi) = \arccos\left(|\langle \Psi_{\text{norm}}, \Phi_{\text{norm}} \rangle|^{\frac{1}{2}}\right)$$

and considering its infinitesimal form, which correspond to the Quantum Geometric Tensor

$$d_{FS}[\Psi_V(\mathbf{X}|\mathbf{w}), \Psi_V(\mathbf{X}|\mathbf{w} + d\mathbf{w})]^2 = \sum_{ij} \mathcal{S}_{ij}(\mathbf{w}) dw_i dw_j$$

defined over a basis $|\psi(\mathbf{X}|\mathbf{w})\rangle \approx \sum_{\mathbf{X}\in\{\mathbf{X}\}} \sqrt{p(\mathbf{X}|\mathbf{w})}\,|\mathbf{X}\rangle$ can be easily shown that $\mathcal{S}_{ij}$ is proportional to the Fisher Information $\mathcal{S}_{ij}(\mathbf{w}) = \frac{1}{4} g_{ij}(\mathbf{w})$.

When the Fisher information (The variance of the gradient of the log-likelyhood) is high, it implies that the variance of the score function is also high, meaning small changes in the parameter $\mathcal{W}$ lead to large changes in the log-likelihood.

Quantum Natural Gradient - Carleo et al. (Quantum May 2020), page 10.

Metropolis Algorithm

Our goal is to sample the probability distribution described by $P(\mathbf{X}) = \frac{|\Psi(\mathbf{X})|^2}{\int d\mathbf{R}|\Psi(\mathbf{X})|^2}$. The $M(RT)^2$ algorithm is based on the idea of random walk, namely a Markov Chain with transition matrix Π.

$$\mathcal{P}(X_{i+1} = x_{i+1}|X_0 = x_0, \cdots, X_i = x_i) = \Pi(x_i, x_{i+1})$$

We can split $\Pi(x_i, x_{i+1})$ in two terms:

$$\Pi(x_{i+1}|x_i) = q(x_{i+1}|x_i)r(x_{i+1}|x_i)$$

This means that if you move to x_{i+1} from x_i

1. x_{i+1} is proposed it with probability $q(x_{i+1}|x_i)$
2. x_{i+1} is accepted with probability $r(x_{i+1}|x_i)$ otherwise x_i is kept.
   The acceptance probability is given by

$$r(x_{i+1}|x_i) = \min\left(1, \frac{P(x_{i+1})\cancel{q(x_i|x_{i+1})}}{P(x_i)\cancel{q(x_{i+1}|x_i)}}\right) \quad \rightarrow \text{Works for non-normalized P(x)}$$

This ensures that the fraction of time spent in each state is proportional to $P(x_i)$.

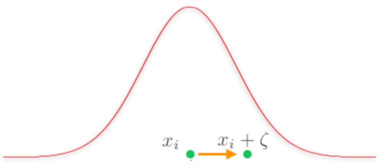
- If $r(x_{i+1}|x_i) = 0.9$ we draw a point z from $U(0,1)$
- if $z_U < 0.9$ we accept the new point, otherwise we reject x_{i+1} and keep x_i.

Metropolis Algorithm

If the proposal is a Normal distribution at each step of the propagation the walkers are moved as follows:

$$q(x_{i+1}|x_i) = N(x_{i+1}|x_i, \mu = 0, \sigma = 1)$$

$$x_{i+1} = x_i + \zeta_{N(\mu=0,\sigma=1)}$$



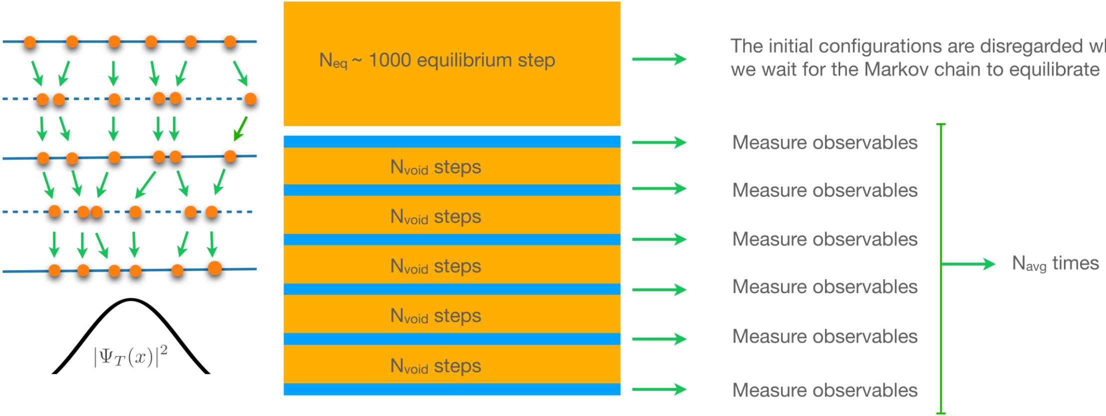
$\mathcal{N}$ MC in parallel

The transfer matrix Π of the MH algorithm ensure that the detailed balance principle is valid and consequently that the Markov chain will converge to the desired target density P after thermalization

$$P(x_i)\Pi(x_{i+1}|x_i) = P(x_{i+1})\Pi(x_i|x_{i+1})$$

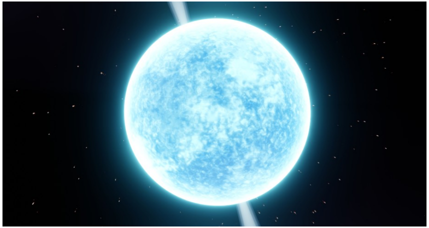
This says that the flow from i to $i+1$ must equal the flow from $i+1$ to i.

Metropolis Algorithm



EoS of Nucleon Stars and the Hyperon puzzle

- One of the most relevant solutions to the Hyperon puzzle relies on the determination of the EoS through **microscopic approaches** which applies a three-body *YNN* interaction potential to compensate the attractive *YN* interaction.



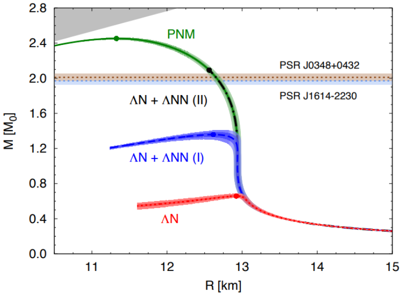
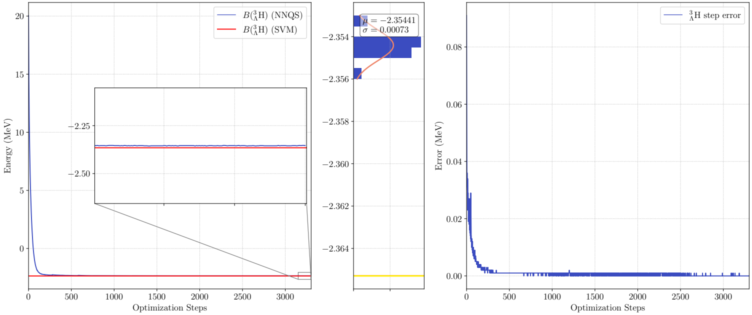
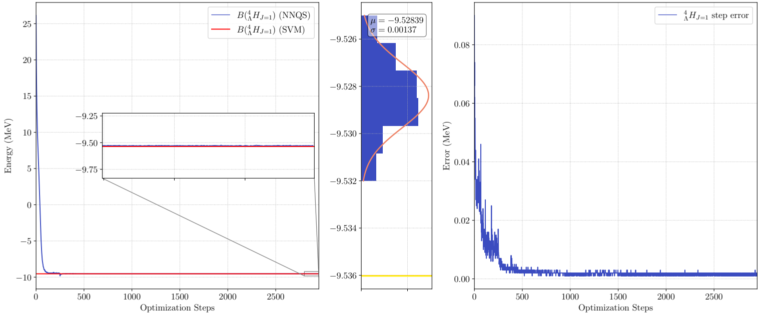
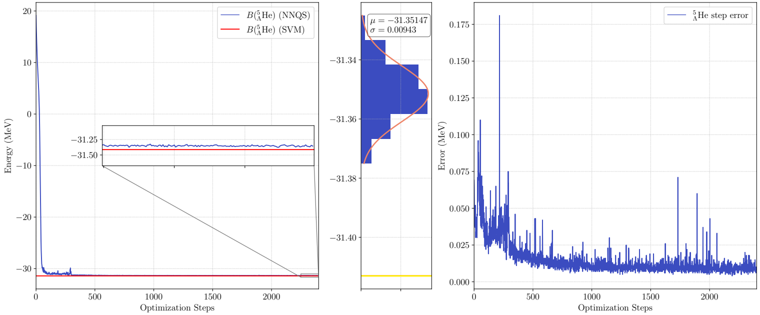


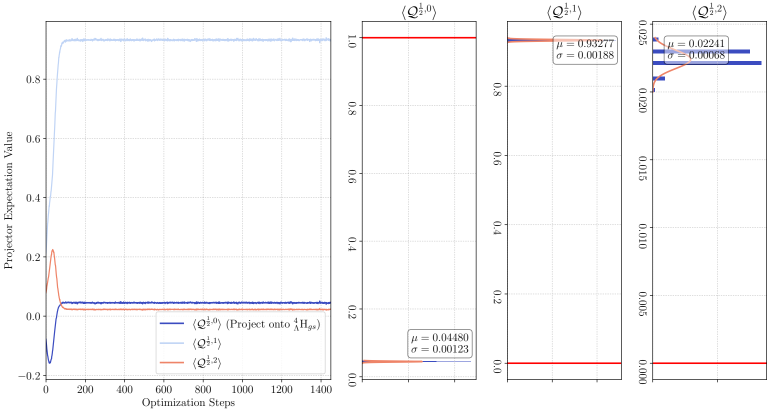
Figure: D. Lonardoni et al., PRL 114 (2015)


$B(^3_\Lambda H)$ (NNQS)
$B(^3_\Lambda H)$ (SVM)
Energy (MeV)
Optimization Steps
$\mu = -2.35441$
$\sigma = 0.00073$
$^3_\Lambda$H step error
Error (MeV)
Optimization Steps


$B(^4_\Lambda H_{J=1})$ (NNQS)
$B(^4_\Lambda H_{J=1})$ (SVM)
Energy (MeV)
Optimization Steps
$\mu = -9.52839$
$\sigma = 0.00137$
$^4_\Lambda H_{J=1}$ step error
Error (MeV)
Optimization Steps



$B(^5_\Lambda\text{He})$ (NNQS)
$B(^5_\Lambda\text{He})$ (SVM)
Energy (MeV)
Optimization Steps
$\mu = -31.35147$
$\sigma = 0.00943$
$^5_\Lambda$He step error
Error (MeV)
Optimization Steps


Projector Expectation Value
Optimization Steps
$\langle \mathcal{Q}^{\frac{1}{2},0} \rangle$ (Project onto ${}^4_\Lambda H_{gs}$)
$\langle \mathcal{Q}^{\frac{1}{2},1} \rangle$
$\langle \mathcal{Q}^{\frac{1}{2},2} \rangle$
$\mu = 0.04480$
$\sigma = 0.00123$
$\mu = 0.93277$
$\sigma = 0.00188$
$\mu = 0.02241$
$\sigma = 0.00068$