Neural Quantum States advancements

for Hypernuclear physics

Andrea Di Donna
June 26, 2024

andrea.didonna@unitn.it
TIFPA - Trento University

Outline

Argonne NATIONAL LABORATORY

TIFPA Trento Institute for Fundamental Physics and Applications

Open questions in Hypernuclear Physics:

UNIVERSITÀ DI TRENTO

- **Hyperon-Nucleon (YN) and Hyperon-Hyperon (YY) interactions are less well-understood than Nuclear forces. → Still these are relevant . . .**

- **At high densities the onset of hyperons in Neutron Stars → Hyperon puzzle**
  - Softening of the EoS due to reduced Pauli blocking.
  - Observational data of 2 $M_{\odot}$ neutron stars are in constrast with softening!
  - To explain M/r ratio we have to introduce strongly repulsive *YNN* → **microscopic interaction**.

- **Interpretation of Astrophysical Observations**
  - We want to understand how Hyperons affect the properties of the emitted gravitational waves, influencing observables such as **tidal deformabilities** during the merger of two NSs.

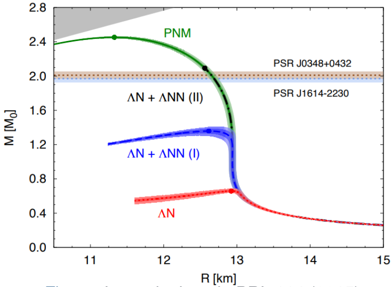


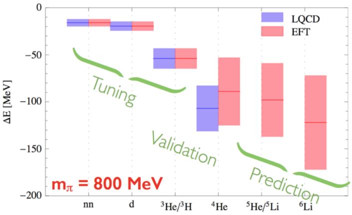
Figure: Lonardoni et al., PRL 114 (2015)

- **Datas from new experiments may allow to constrain more efficiently the microscopic *YN* interaction at higher dentisies.**
- **To tackle these problems we need more accurate microscopic interaction models and computational tools.**

UNIVERSITÀ DI TRENTO

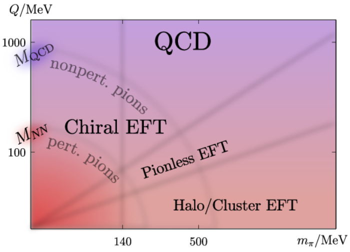
Ab Initio description of Interactions

- *Ab initio* approaches derived from QCD →
  1. Relies on scarce experimental data to be constrained: Scattering parameters and BE to constrain two- and three-body potentials.
  2. Provide controlled estimations of theoretical uncertainties expanding order by order.
  3. Disadvantages: After renormalization unresolved contact interactions must be extrapolated at $\lambda \to \infty$



X

Figure: $\not{\pi}$EFT at LO, Barnea et al, PRL 2015



- **Predictivity:** *Ab initio* approach can predict properties of systems that have not yet been observed.
- E.g.: Extending applicability of LQCD
  1. Two- and three-body LECs from LQCD at fixed pion masses → predict four-body and higher-body interactions with EFT.

UNIVERSITÀ DI TRENTO

- **Hypernuclei are natural laboratories to study these interactions**:
  - To determine the hyperon-nucleon and hyperon-hyperon interaction potentials we need precise calculations of hypernuclear BE.
  - Advancment in the computational approaches plays a significant role.

- **Objectives**
  - **Parametrize** the wavefunction (variational state) using NNQS for single lambda hypernuclei:

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

  - **Fit** on few-body systems an **improved** leading-order (LO) interaction potential derived from $\not\pi$ EFT.
  - **Predict** Λ-separation and other observables for higher-A systems (${}^{7}_{\Lambda}$Li, ${}^{13}_{\Lambda}$C, ${}^{16}_{\Lambda}$O).

$$B_{\Lambda} = E(^{A-1}Z) - E({}^{A}_{\Lambda}Z)$$

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

UNIVERSITÀ DI TRENTO

- **Pionless Effective Field Theory ($\not\pi$EFT):**
  - Interaction potential derived from $\not\pi$EFT, valid for energies $Q < M_{hi} = m_\pi = 140\,\text{MeV}$.
  - Unresolved pions lead to spin-dependent contact potentials described by Low Energy Constants (LECs).

$$C_{IS}(\lambda)\delta_\lambda(\mathbf{x}_{ij}) = C_{IS}(\lambda)e^{-\frac{\lambda^2\mathbf{x}_{ij}^2}{4}}$$

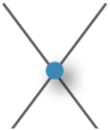
- **Fitting Procedure:**
  - Regulator cutoffs and LECs are fitted using SVM (Suzuki-Varga) and Gaussian Processes.

NN Potential:

- Fitted to neutron-proton (np) and neutron-neutron (nn) scattering lengths and effective ranges.

Three-Nucleon Force (3NF):

- Adjusted to reproduce the binding energies (BEs) of ^{3}H and ^{4}He.



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

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

Extending $\not\pi$EFT to Hyperons

UNIVERSITÀ DI TRENTO

Introduction of Λ-Hyperons → Modify LO Interaction Potential:

- Extending Pionless Effective Field Theory ($\not\pi$EFT) to hyperons requires introducing Λ-hyperon degrees of freedom ($m_\Lambda = 1116\,\text{MeV}$) into the Lagrangian density $\mathcal{L}$.

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

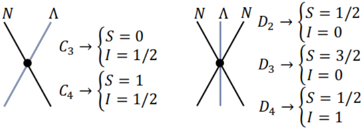
$$V_{\Lambda N} = \sum_{IS} C_\lambda^{IS} \sum_{i<j} \mathcal{P}_{IS}(ij)\delta_\lambda(\vec{r}_{ij})$$

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

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

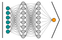
Fitting Procedure:

1. ΛN interaction is fitted to $p\Lambda$ low enregy scattering lengths and effective ranges *Alexander et al.*
2. ΛNN interaction is fitted to binding energies of ${}^3_\Lambda$H, ${}^4_\Lambda\text{H}_{S_\text{tot}=0}$, ${}^4_\Lambda\text{H}_{S_\text{tot}=1}$, and ${}^5_\Lambda$He.



Overview:

- The Neural Network Quantum States (NNQS) approach (Carleo et al., 2017) is a form of unsupervised learning.
- **Universal Approximation:** Any type of variational state can be represented via NNQS.

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

- Utilizes **scalable neural networks** for representation.
- **Hidden Nucleon wavefunctions introduce dynamical correlations** between particles and preserve statistical correlations through peculiar Network architectures.

Advantages:

- Easily applicable to different systems.
- **Approximation error is no longer wavefunction-dependent** for a given interaction.
- Less time-consuming and computationally demanding than other exact methods for $N \geq 5$.
- Computational cost scales polynomially with $\sim \alpha N^{5\text{–}6}$.
- GPU parallelization offer linear scaling with the number of GPU (Tested up to 32 GPUs $\approx$ 2s/Opt. step for ${}^{16}_{\Lambda}O$)

Disadvantages:

- EFT contact potentials $\delta_\lambda(\mathbf{x}_{ij})$ introduce **errors at large cut-offs** in Monte Carlo evaluation of observables due to limited statistics in the interaction range.

UNIVERSITÀ DI TRENTO

NNQS as a Variational Method

1. **Variational Principle:**

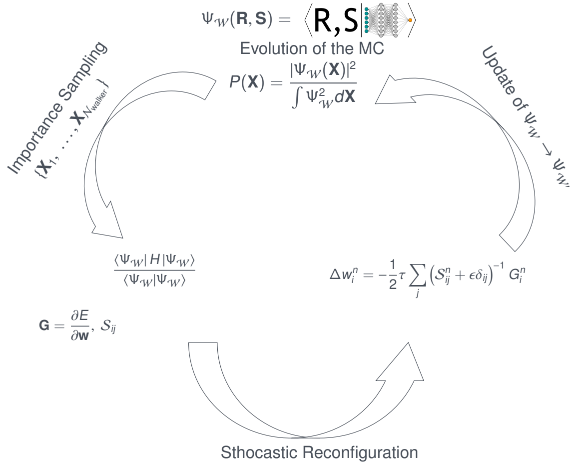
$$E_V = \frac{\langle \Psi_{\mathcal{W}} | H | \Psi_{\mathcal{W}} \rangle}{\langle \Psi_{\mathcal{W}} | \Psi_{\mathcal{W}} \rangle} \geq E_0$$

2. **Iterative Optimization:**

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

Defining an iterative procedure:

1. **Sampling the Wavefunction:** Use $\mathcal{N}$ Markov Chains in parallel.
2. **Evaluating Observables and Gradients:** Calculate expectation values using importance sampling.
3. **Parameter Update:** Apply stochastic reconfiguration.



UNIVERSITÀ DI TRENTO

- **Modeling the Wavefunction with the Hidden Nucleon Approach for Fermionic Systems:**
  - Provides a systematic and extensible method for modeling **antisymmetric wavefunctions** through a **single extended Slater determinant**.
  - Reference: "Fermionic wave functions from neural-network constrained hidden states" by Carleo, Moreno et al. 2022

Constructing the Wavefunction:

1. Introduces A_h virtual particles **"hidden" DoFs** in contrast to the real ones **"visible" DoFs**.
2. Hidden DoFs are represented by adding **hidden orbitals** $\chi_i(\tilde{x}_j)$.
3. Each virtual particle coordinate $\tilde{x}_j$ is a bosonic function of all real particle coordinates: $\tilde{x}_j = f(\{\mathbf{X}\})$ where $\{X\} = \{\mathbf{x}_1, \ldots, \mathbf{x}_A\}$.

$$\Psi_{\mathrm{HN}}(\mathbf{X}) = \det\begin{pmatrix} \phi_{\mathrm{v}}(\mathbf{X}) & \phi_{\mathrm{v}}(f(\{\mathbf{X}\})) \\ \chi_{\mathrm{h}}(\mathbf{X}) & \chi_{\mathrm{h}}(f(\{\mathbf{X}\})) \end{pmatrix} = \det\begin{bmatrix} \phi_1(x_1) & \phi_1(x_2) & \phi_1(f(\{X\})) \\ \phi_2(x_1) & \phi_2(x_2) & \phi_2(f(\{X\})) \\ \chi_3(x_1) & \chi_3(x_2) & \chi_3(f(\{X\})) \end{bmatrix}$$

Visible orbitals on visible coordinates | Visible orbitals on hidden coordinates

Hidden orbitals on visible coordinates | Hidden orbitals on hidden coordinates

UNIVERSITÀ DI TRENTO

Question: Does extending the determinant preserve its antisymmetry under particle permutations?
Expansion of the Determinant:

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

Comparison with Configuration Interaction (CI) Expansion:

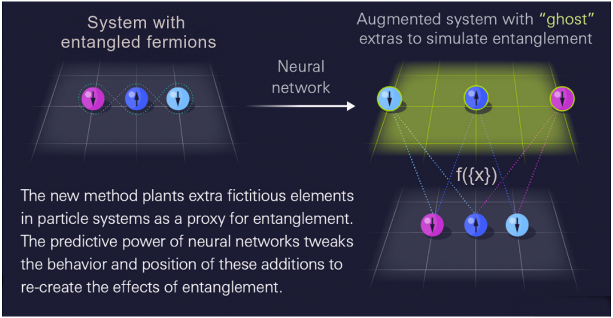
$$|\Phi\rangle = C_0 |\Psi_0\rangle + \sum_{ra} C_a^r |\Psi_a^r\rangle + \sum_{\substack{a<b \\ r<s}} C_{ab}^{rs} |\Psi_{ab}^{rs}\rangle + \sum_{\substack{a<b<c \\ r<s<t}} C_{abc}^{rst} |\Psi_{abc}^{rst}\rangle + \dots$$

- Looks like a CI expansion with only one excited state (but it's more than that*).
- If each virtual coordinate $\tilde{x}_i = f(x)$ is a **permutation-invariant function (Network)** of the visible particle coordinates, the **antisymmetry** of the wavefunction is preserved.

UNIVERSITÀ DI TRENTO

Incorporating Dynamical Correlations with $f(\{\mathbf{x}\})$

- The function $f(\{\mathbf{x}\})$ accounts for **Jastrow correlators**, introducing dynamical correlations into the wavefunction (**Isospin, Spin, Coulomb correlations** $\mathbf{x}_i = [\mathbf{r}_i, s_{z_i}, t_{z_i}]$).



Neural Network Implementation:

- **Permutation Invariance:** Introduced using the **Deep Sets architecture**.
- **Enhancing Correlations:** Improved through a **backflow transformation** of the coordinates $\{\mathbf{x}\}$.

UNIVERSITÀ DI TRENTO

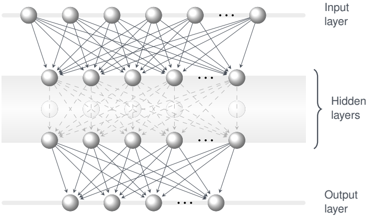
- **Phase-Amplitude representation of the Visible Coordinate orbitals:**
  - The Visible coordinate sector involves both visible and hidden orbitals. The functions $\phi_\alpha(x_i)$ and $\chi_\alpha(x_i)$ are defined as:

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

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

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

  - Here, u and v are both feedforward neural networks (FFNNs) with only one hidden layer.



Single Particle FFNN Details:

- **Input Nodes:** 5
  $\rightarrow [\mathbf{r}_i, s_z, t_z]$ where $\mathbf{r}_i \in \mathbb{R}^3$
- **Non-linear Activation:** Applied in hidden layers
- **Output Nodes:** 1
  $\rightarrow \phi_\alpha(x_i)$

UNIVERSITÀ DI TRENTO

- **Phase-Amplitude representation of the Hidden Coordinate orbitals:**
  - The Hidden coordinate sector involves both visible and hidden orbitals **evaluated on hidden coordinates**. The functions $\phi_\alpha(x_i)$ and $\chi_\alpha(x_i)$ are defined as:

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

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

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

- **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 $g(\cdot) = \sum_{x\in\mathfrak{X}}$.

UNIVERSITÀ DI TRENTO

- **Variational State of a Hypernucleus (single Λ orbital):**

$$\Psi_{HN} = \phi_\Lambda(\mathbf{x}_\Lambda, f(\{\mathbf{x}_1, \dots, \mathbf{x}_{A-1}, \mathbf{x}_\Lambda\})) \cdot \det_{A-1_Z} \begin{vmatrix} \phi_\mathrm{v}(\mathbf{X}) & \phi_\mathrm{v}(f(\{\mathbf{X}\})) \\ \chi_\mathrm{h}(\mathbf{X}) & \chi_\mathrm{h}(f(\{\mathbf{X}\})) \end{vmatrix} \quad \rightarrow \textbf{Not sufficient!}$$

UNIVERSITÀ DI TRENTO

- **Variational State of a Hypernucleus (single Λ orbital):**

$$\Psi_{HN} = \phi_\Lambda(\mathbf{x}_\Lambda, f(\{\mathbf{x}_1, \ldots, \mathbf{x}_{A-1}, \mathbf{x}_\Lambda\})) \cdot \det_{A-1_Z} \begin{vmatrix} \phi_v(\mathbf{X}) & \phi_v(f(\{\mathbf{X}\})) \\ \chi_h(\mathbf{X}) & \chi_h(f(\{\mathbf{X}\})) \end{vmatrix} \rightarrow \textbf{Not sufficient!}$$

  - **EXAMPLE:** To represent ΛN correlations in the ${}^3_\Lambda$H mixed asimmetry spin state, the aggregator function for non-identical particles must be modified:

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

    Here, the coordinate of the Λ particle is concatenated with the sum pooling, so ΛN permutation invariance is not introduced.

  - **Phase-Amplitude representaion of ϕ_Λ:**

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

UNIVERSITÀ DI TRENTO

- **Extension to more than one Λs:**

By combining the previous ideas, we generalize the Hidden Nucleon wavefunction to include multiple Λ particles:

$$\Psi_{\mathrm{HN}} = \det_{\Lambda}\begin{pmatrix} \phi_{\mathrm{v}}(\mathbf{X}_\Lambda) & \phi_{\mathrm{v}}\left(f\left(\{\mathbf{X}_\Lambda\}\right)\right) \\ \chi_{\mathrm{h}}(\mathbf{X}_\Lambda) & \chi_{\mathrm{h}}\left(f\left(\{\mathbf{X}_\Lambda\}\right)\right) \end{pmatrix} \cdot \det_{A-1_Z}\begin{pmatrix} \phi_{\mathrm{v}}(\mathbf{X}_i) & \phi_{\mathrm{v}}\left(f\left(\{\mathbf{X}_i\}\right)\right) \\ \chi_{\mathrm{h}}(\mathbf{X}_i) & \chi_{\mathrm{h}}\left(f\left(\{\mathbf{X}_i\}\right)\right) \end{pmatrix}$$

Where:

- $\det_\Lambda$ corresponds to the Λ hyperons.
- $\det_{A-1_Z}$ corresponds to the nucleons (protons and neutrons).
- $\mathbf{X}_\Lambda$ and $\mathbf{X}_i$ are the backflow-transformed coordinates for the Λ particles and nucleons, respectively.

UNIVERSITÀ DI TRENTO

Observables Evaluation

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

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

- **Definitions:**
  - **Local observable:** $\mathcal{O}_L(\mathbf{X}) = \frac{\mathcal{O}\Psi(\mathbf{X})}{\Psi(\mathbf{X})}$
  - **Probability density:** $P(\mathbf{X}) = \frac{|\Psi(\mathbf{X})|^2}{\int d\mathbf{R}|\Psi(\mathbf{X})|^2}$ → Sampled trough Metropolis Hastings Algorithm

Monte Carlo Estimation:

$$\mathcal{O}_V = \frac{1}{\mathcal{N}_{\text{walker}}}\sum_{s=1}^{\mathcal{N}_{\text{walker}}}\mathcal{O}_L(\mathbf{X}_s)$$

Statistical Uncertainty:

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

UNIVERSITÀ DI TRENTO

Objective: To sample the probability distribution: $P(\mathbf{X}) = \frac{|\Psi_T(\mathbf{X})|^2}{\int d\mathbf{R} |\Psi_T(\mathbf{X})|^2}$

Metropolis-Hastings Algorithm:

- Based on the concept of a **random walk** forming a **Markov chain** with transition matrix Π.
- The transition probability is given by:

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

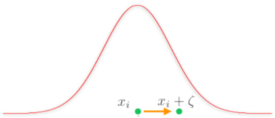
Proposal Step:

- Generate a new state:

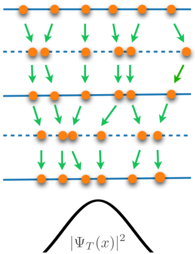
$$x_{i+1} = x_i + \zeta, \quad \zeta \sim N(0, \sigma^2)$$

- Compute the acceptance ratio:

$$r(x_{i+1}|x_i) = \min\left(1, \frac{P(x_{i+1}) q(x_i|x_{i+1})}{P(x_i) q(x_{i+1}|x_i)}\right)$$

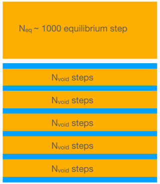


Multiple Particle Diffusion:



Sampling Procedure:

1. Allow the system to **thermalize**.
2. Collect N_{avg} samples, spaced by N_{void} steps: **reduce autocorrelation**



UNIVERSITÀ DI TRENTO

Overview:

- **Stochastic Reconfiguration (SR):** Introduced by S. Sorella (2005).
- In the context of Variational Monte Carlo (VMC) SR is equivalent to performing an **imaginary-time evolution** in parameter space.
- It is related to the **Natural Gradient Descent** method (Amari et al.).

Imaginary-Time Evolution in Parameter Space:

- The imaginary-time evolution operator is approximated as ($\tau \approx 0$):

$$e^{-H\tau} \approx 1 - H\tau$$

- Acting on the trial wavefunction $|\Psi_T(\mathcal{W})\rangle$, we have:

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

- Where: $O^i\,|\Psi_T(\mathcal{W})\rangle = \left|\frac{\partial}{\partial \mathcal{W}_i}\Psi_T(\mathcal{W})\right\rangle$

Deriving the SR Equations:

- Multiply both sides from the left by $\frac{\langle\Psi_T(\mathcal{W})|}{\langle\Psi_T(\mathcal{W})|\Psi_T(\mathcal{W})\rangle}$ and $\frac{\langle\Psi_T(\mathcal{W})|O^i}{\langle\Psi_T(\mathcal{W})|\Psi_T(\mathcal{W})\rangle}$,

Solving for Parameter Updates:

- Solving the first equation for $\Delta \mathcal{W}_0$ and substituting into the second equation yields:

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

- **Defining quantities Evaluated via Importance Sampling:**
  - **Gradient:**

$$G_i = 2\left(\langle HO^i\rangle - \langle H\rangle\langle O^i\rangle\right) = -2\frac{\partial E(\mathcal{W})}{\partial \mathcal{W}_i}$$

  - **Quantum Geometric Tensor (QGT):**

$$\mathcal{S}_{ij} = \langle O^i O^j\rangle - \langle O^i\rangle\langle O^j\rangle$$

  - The update rule for the parameters becomes:

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

  - The term $\epsilon\delta_{ij}$ is added to avoid saddle points and improve numerical stability.
- Note: The QGT resembles a covariance matrix.

Fisher Information and Quantum Geometric Tensor

Classical Information Theory:

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

- **Covariance of the derivative of the log-likelihood** (Score Function)

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

- **Curvature of the log-likelihood**

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

Quantum Information Theory:

- The Fubini-Study metric defines the distance between infitesimally close quantum states:

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

- If Ψ is defined over a basis: $|\Psi(\mathbf{X}|\mathbf{w})\rangle = \sum_{\mathbf{X}} \sqrt{P(\mathbf{X}|\mathbf{w})}\,|\mathbf{X}\rangle$ Then it 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

Interpretations:

- A sharp Fisher Information (high curvature $\partial^2_{\mathbf{w}_i}$) $\rightarrow \Psi$ (or its walker's sample) is relevantly determined from parameter $\mathbf{w}_i$.
- We suppress the variation of such parameters: $G_i^n \rightarrow \sum_j (\mathcal{S}_{ij}^n)^{-1} G_j^n$.

Stochastic Reconfiguration with RMSProp

Combining SR with RMSProp:

- In RMSProp, the gradient expresses the **acceleration** in the parameter space.
- The $\mathcal{S}$ matrix (Quantum Geometric Tensor) is regularized using running averages of the squared gradients to improve numerical stability.

Update Rules:

Compute running averages:

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

Regularize $\mathcal{S}$ matrix:

$$\mathcal{S}_{ij}^{n} + \epsilon\,\delta_{ij} \to \mathcal{S}_{ij}^{n} + \epsilon\,\mathrm{diag}\left(\sqrt{\mathbf{m}^{n}} + 10^{-8}\right)$$

Explanation:

- The hyperparameter ϵ adds an L^2 penalty term to the solution of the linear system used to determine $\Delta\mathcal{W}$:

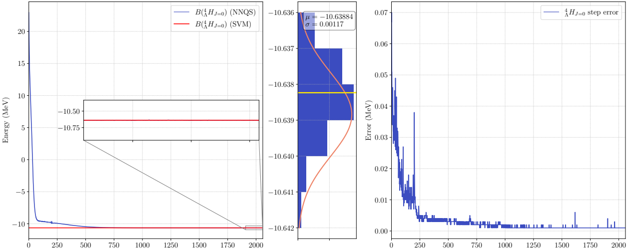
$$||\Delta\mathcal{W}\,\mathcal{S} - \mathbf{G}|| = ||\epsilon\,\Delta\mathcal{W}||$$

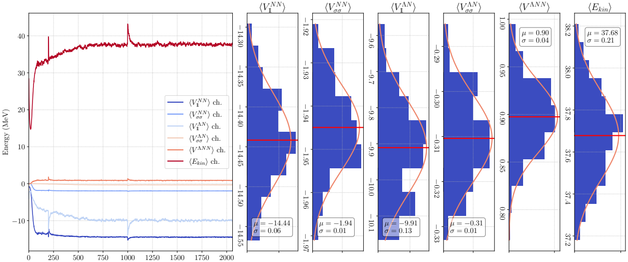
- This regularization helps to avoid issues with ill-conditioned $\mathcal{S}$ matrices and prevents large, unstable parameter updates.

UNIVERSITÀ DI TRENTO

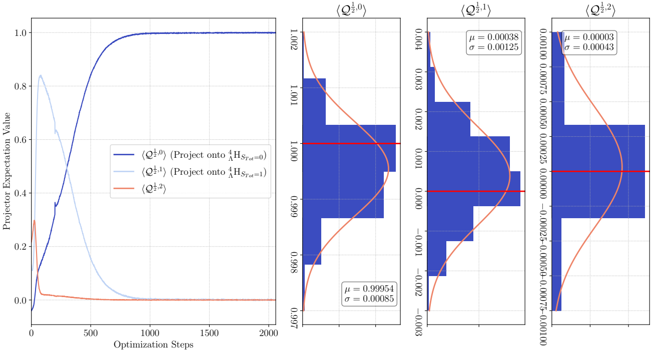
Let's see some results. . .

UNIVERSITÀ DI TRENTO



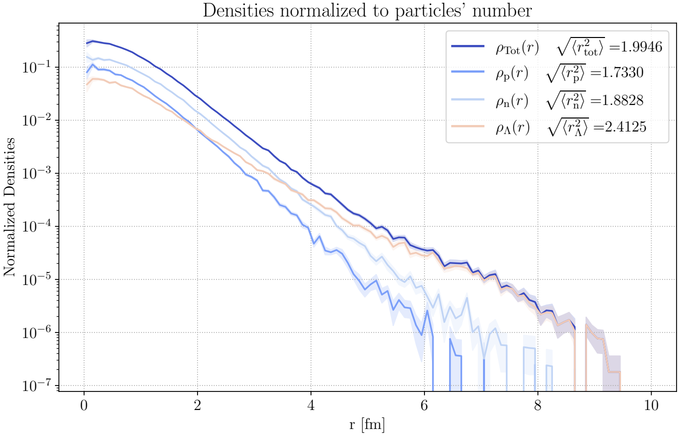


${}^{4}_{\Lambda}H_{S_{tot}=0}$ Projectors

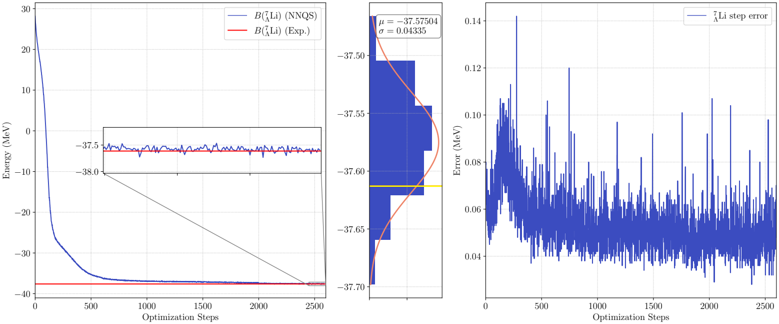


${}^{4}_{\Lambda}H_{S_{tot}=0}$ Density

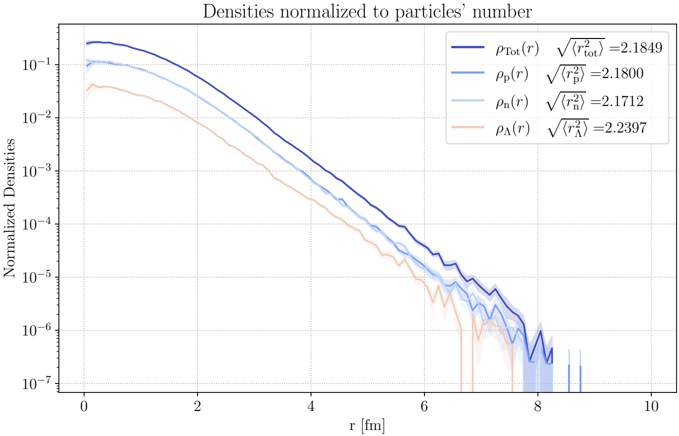
UNIVERSITÀ DI TRENTO



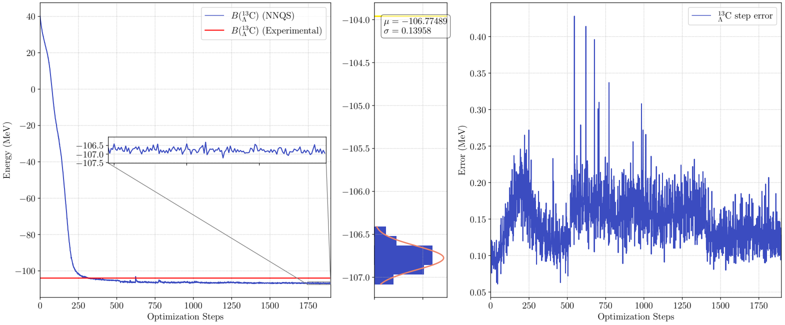
$^{7}_{\Lambda}Li$ with Fitted Potential



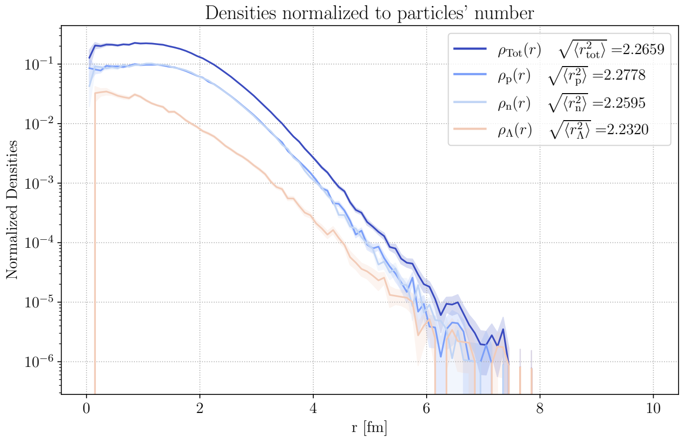
${}^{7}_{\Lambda}Li$ with Fitted Potential



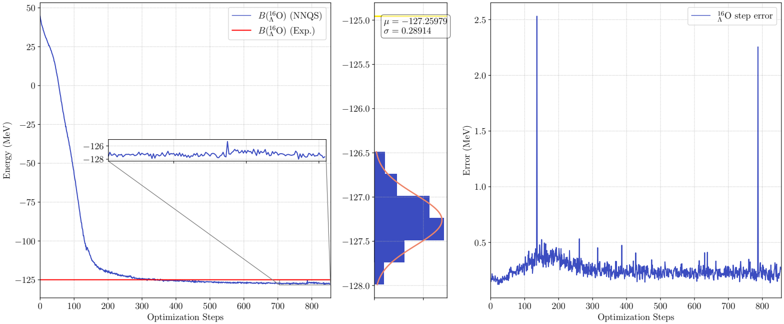
${}^{13}_{\Lambda}C$ with Fitted Potential



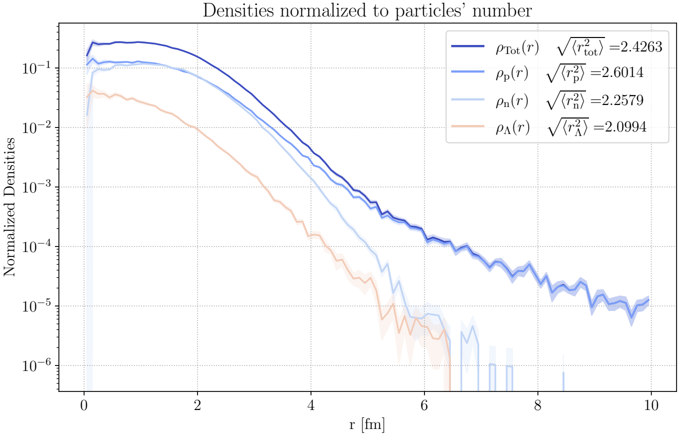
$^{13}_{\Lambda}C$ with Fitted Potential

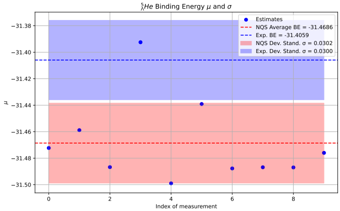


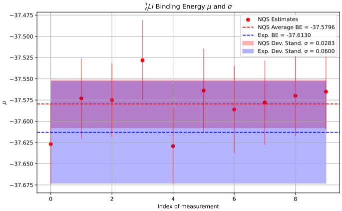
$^{16}_{\Lambda}O$ with Fitted Potential

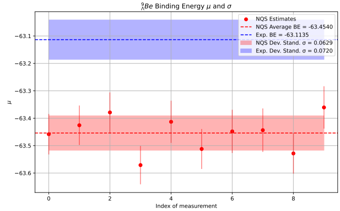


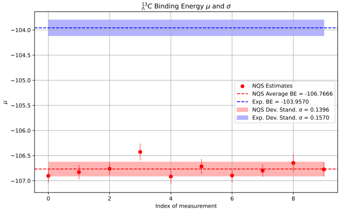
$^{16}_{\Lambda}O$ with Fitted Potential



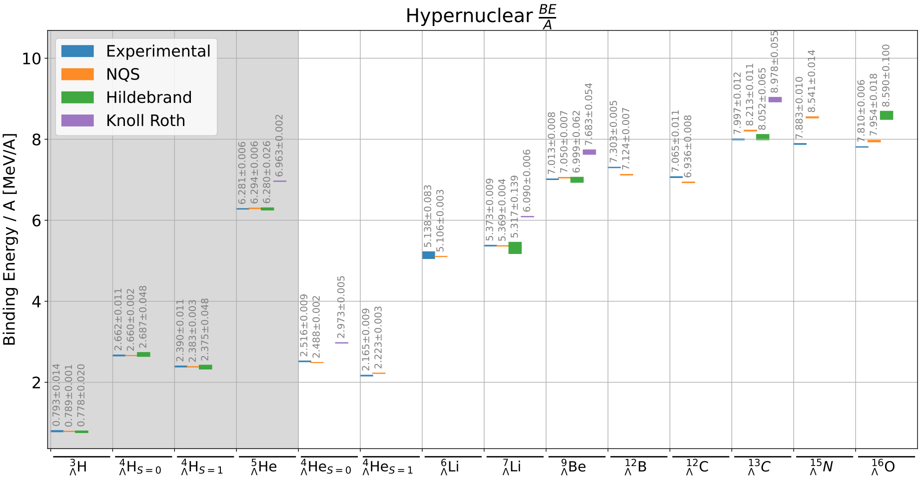


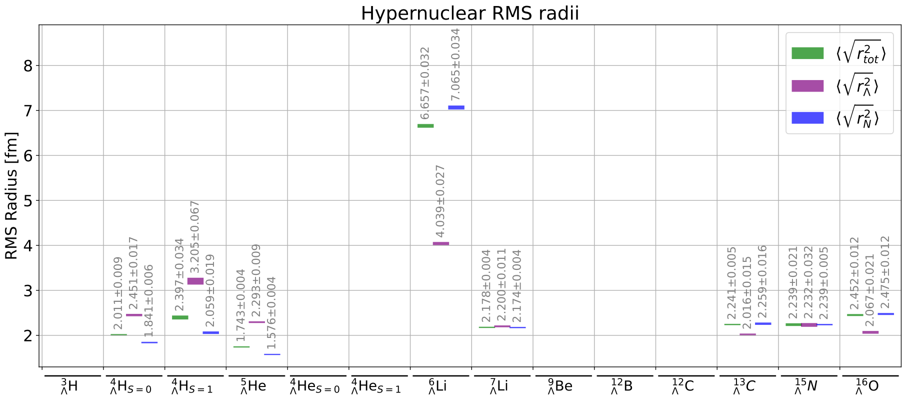




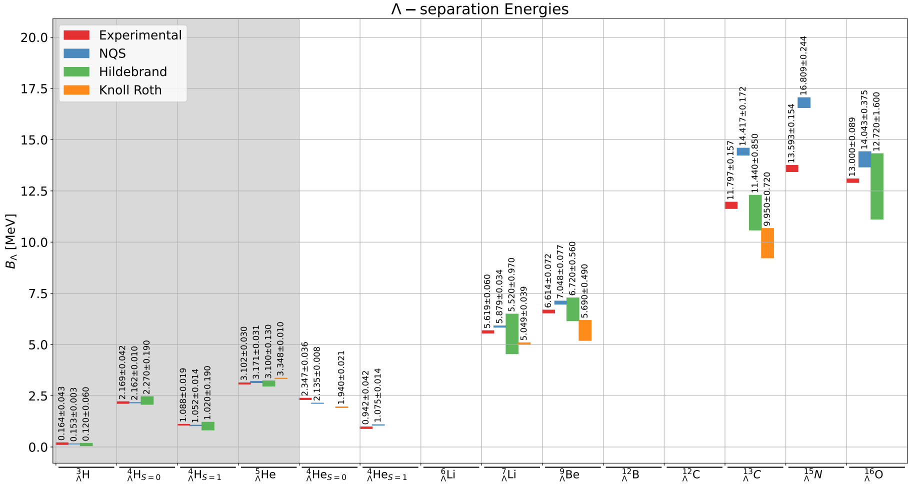


To test the stability of the ΛN-ΛNN interaction each BE is evaluated for 10 different potentials


Hypernuclear $\frac{BE}{A}$
Binding Energy / A [MeV/A]
Experimental
NQS
Hildebrand
Knoll Roth
0.793±0.014
0.789±0.001
0.778±0.020
2.662±0.011
2.660±0.002
2.687±0.048
2.390±0.011
2.383±0.003
2.375±0.048
6.281±0.006
6.294±0.006
6.280±0.026
6.963±0.002
2.516±0.009
2.488±0.002
2.973±0.005
2.165±0.009
2.223±0.003
5.138±0.083
5.106±0.003
5.373±0.009
5.369±0.004
5.317±0.139
6.090±0.006
7.013±0.008
7.050±0.007
6.999±0.062
7.683±0.054
7.303±0.005
7.124±0.007
7.065±0.011
6.936±0.008
7.997±0.012
8.213±0.011
8.052±0.065
8.978±0.055
7.883±0.010
8.541±0.014
7.810±0.006
7.954±0.018
8.590±0.100
$^{3}_{\Lambda}$H
$^{4}_{\Lambda}$H$_{S=0}$
$^{4}_{\Lambda}$H$_{S=1}$
$^{5}_{\Lambda}$He
$^{4}_{\Lambda}$He$_{S=0}$
$^{4}_{\Lambda}$He$_{S=1}$
$^{6}_{\Lambda}$Li
$^{7}_{\Lambda}$Li
$^{9}_{\Lambda}$Be
$^{12}_{\Lambda}$B
$^{12}_{\Lambda}$C
$^{13}_{\Lambda}C$
$^{15}_{\Lambda}N$
$^{16}_{\Lambda}$O



Hypernuclear RMS radii
RMS Radius [fm]
⟨√r²_tot⟩
⟨√r²_Λ⟩
⟨√r²_N⟩
2.011±0.009
2.451±0.017
1.841±0.006
2.397±0.034
3.205±0.067
2.059±0.019
1.743±0.004
2.293±0.009
1.576±0.004
6.657±0.032
4.039±0.027
7.065±0.034
2.178±0.004
2.200±0.011
2.174±0.004
2.241±0.005
2.016±0.015
2.259±0.016
2.239±0.021
2.232±0.032
2.239±0.005
2.452±0.012
2.067±0.021
2.475±0.012
3ΛH
4ΛH S=0
4ΛH S=1
5ΛHe
4ΛHe S=0
4ΛHe S=1
6ΛLi
7ΛLi
9ΛBe
12ΛB
12ΛC
13ΛC
15ΛN
16ΛO


Λ − separation Energies
Experimental
NQS
Hildebrand
Knoll Roth
B_Λ [MeV]
20.0
17.5
15.0
12.5
10.0
7.5
5.0
2.5
0.0
0.164±0.043
0.153±0.003
0.120±0.060
2.169±0.042
2.162±0.010
2.270±0.190
1.088±0.019
1.052±0.014
1.020±0.190
3.102±0.030
3.171±0.031
3.100±0.130
3.348±0.010
2.347±0.036
2.135±0.008
1.940±0.021
0.942±0.042
1.075±0.014
5.619±0.060
5.879±0.034
5.520±0.970
5.049±0.039
6.614±0.072
7.048±0.077
6.720±0.560
5.690±0.490
11.797±0.157
14.417±0.172
11.440±0.850
9.950±0.720
13.593±0.154
16.809±0.244
13.000±0.089
14.043±0.375
12.720±1.600
$^{3}_{\Lambda}$H
$^{4}_{\Lambda}$H$_{S=0}$
$^{4}_{\Lambda}$H$_{S=1}$
$^{5}_{\Lambda}$He
$^{4}_{\Lambda}$He$_{S=0}$
$^{4}_{\Lambda}$He$_{S=1}$
$^{6}_{\Lambda}$Li
$^{7}_{\Lambda}$Li
$^{9}_{\Lambda}$Be
$^{12}_{\Lambda}$B
$^{12}_{\Lambda}$C
$^{13}_{\Lambda}C$
$^{15}_{\Lambda}N$
$^{16}_{\Lambda}$O

Conclusions

- **NNQS shows great flexibility in modeling hypernuclear bound states.**
  - NQS Variational State → match with accuracy the binding energy and quantum numbers of studied hypernuclear systems.
  - Fitted interaction: Improved LO guarantees the desired accuracy.
- **Other advancements in the field: (A. Lovato, A. Gnech)**
  - Shell structure is emerging even if not encoded in the system https://arxiv.org/pdf/2308.16266
- **Outlook**:
  1. Moving to NLO and restore cutoff dependence.
  2. Extension of calculations to larger mass hypernuclei.
  3. Hypernuclear matter → EoS

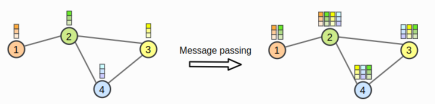
Conclusions

- **NNQS shows great flexibility in modeling hypernuclear bound states.**
  - NQS Variational State → match with accuracy the binding energy and quantum numbers of studied hypernuclear systems.
  - Fitted interaction: Improved LO guarantes the desired accuracy.
- **Other advancements in the field: (A. Lovato, A. Gnech)**
  - Shell structure is emerging even if not encoded in the system https://arxiv.org/pdf/2308.16266
- **Outlook**:
  1. Moving to NLO and restore cutoff dependence.
  2. Extension of calculations to larger mass hypernuclei.
  3. Hypernuclear matter → EoS

Thank you!

UNIVERSITÀ DI TRENTO

- **Backflow Transformation for Single Particle Coordinates:**
  - Modifies particle coordinates so that the **effective position** of a particle depends on its own position and the positions of all other particles.



Neural Network Interpretation:

- Application of a minimal **Message Passing Neural Network (MPNN)**.
- An **all-to-all connected graph** encoding:
  - **Nodes** (x_i)**:** Particle positions.
  - **Edges** (m_{ij})**:** Interactions between particles.

Backflow for Λ Coordinate

$$m_{Lj} = \phi\left([x_\Lambda, x_j]\right)$$

$$m_L = \frac{1}{N_{\text{part}}} \sum_{j=1}^{N_{\text{part}}} m_{Lj}$$

$$\mathbf{X}_\Lambda = [x_\Lambda, m_L]$$

Backflow for Nucleon Coordinates

$$m_{ij} = \chi\left([x_i, x_j]\right)$$

$$m_i = \frac{1}{N_{\text{part}}} \sum_{j=1}^{N_{\text{part}}} m_{ij}$$

$$\mathbf{X}_i = [x_i, m_i]$$

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

Riferimenti I

Gal, A., Hungerford, E. V., & Millener, D. J. (2016).
Strangeness in nuclear physics.
Reviews of Modern Physics, 88(3), 035004.

Lonardoni, D., Gandolfi, S., & Pederiva, F. (2015).
Hyperon Puzzle: Hints from Quantum Monte Carlo Calculations.
Physical Review C, 87(4), 041303(R).

Abbott, B. P., et al. (2017).
GW170817: Observation of Gravitational Waves from a Binary Neutron Star Inspiral.
Physical Review Letters, 119(16), 161101.

Haidenbauer, J., et al. (2013).
Hyperon-nucleon interaction at next-to-leading order in chiral effective field theory.
Nuclear Physics A, 915, 24-58.

Lonardoni, D., et al. (2014).
Hyperon-Nucleon Interactions from Quantum Monte Carlo Calculations.
Physical Review Letters, 113(13), 142301.

Hiyama, E., & Yamamoto, Y. (2012).
Recent progress in studies of baryon-baryon interactions and hypernuclei.
Progress in Particle and Nuclear Physics, 63(2), 339-395.

Epelbaum, E., Hammer, H.-W., & Meißner, U.-G. (2009).
Modern theory of nuclear forces.
Reviews of Modern Physics, 81(4), 1773.